

ISIPTA '13

Allowing for probability zero in...

credal networks under epistemic irrelevance

...using sets of desirable gambles

Jasper De Bock & Gert de Cooman

3 July 2013, Compiègne, France





Research group
SYSTeMS



Jasper De Bock



Gert de Cooman



Stavros Lopatzidis



Arthur Van Camp

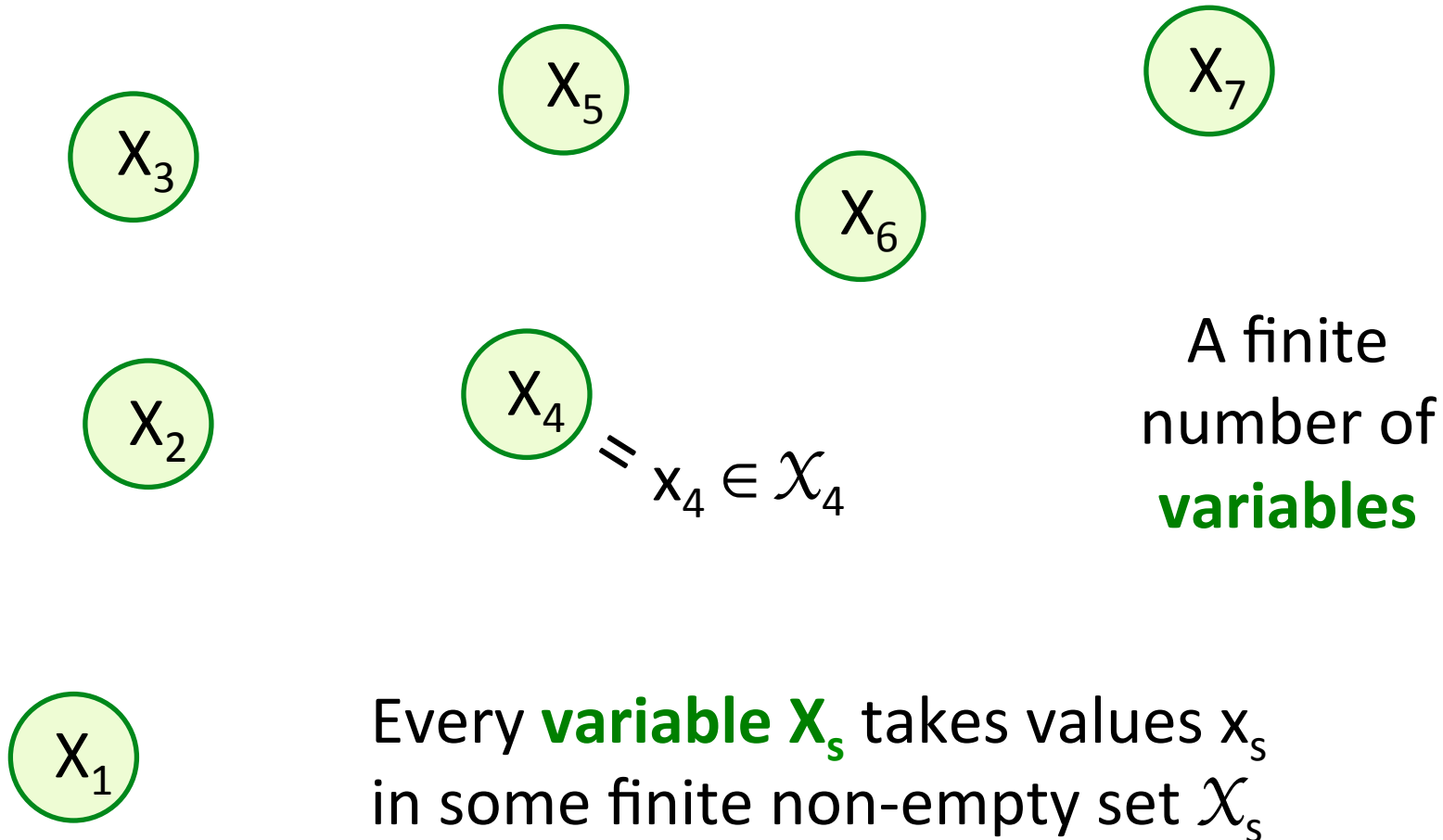


Erik Quaeghebeur

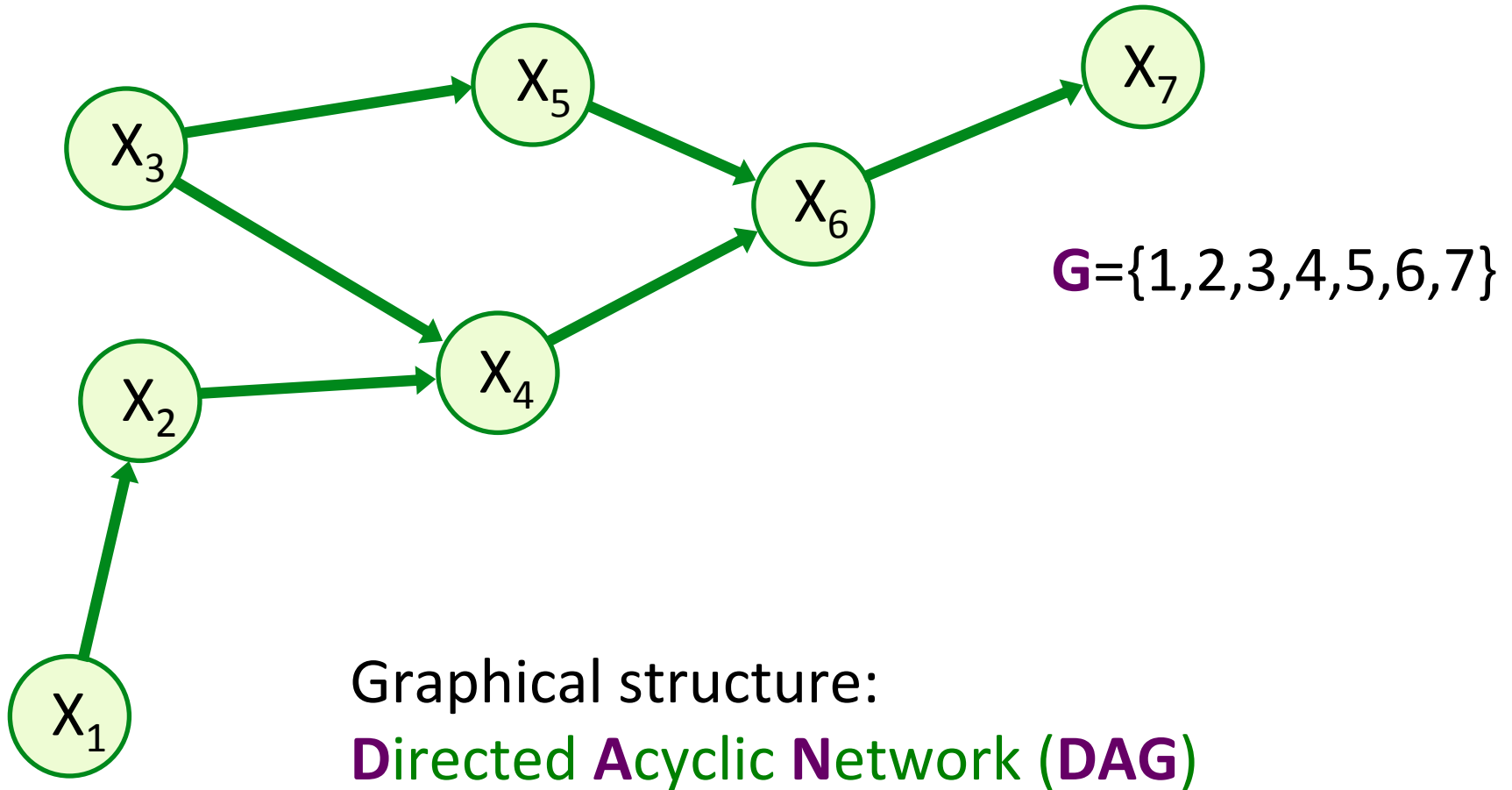


Márcio A. Diniz

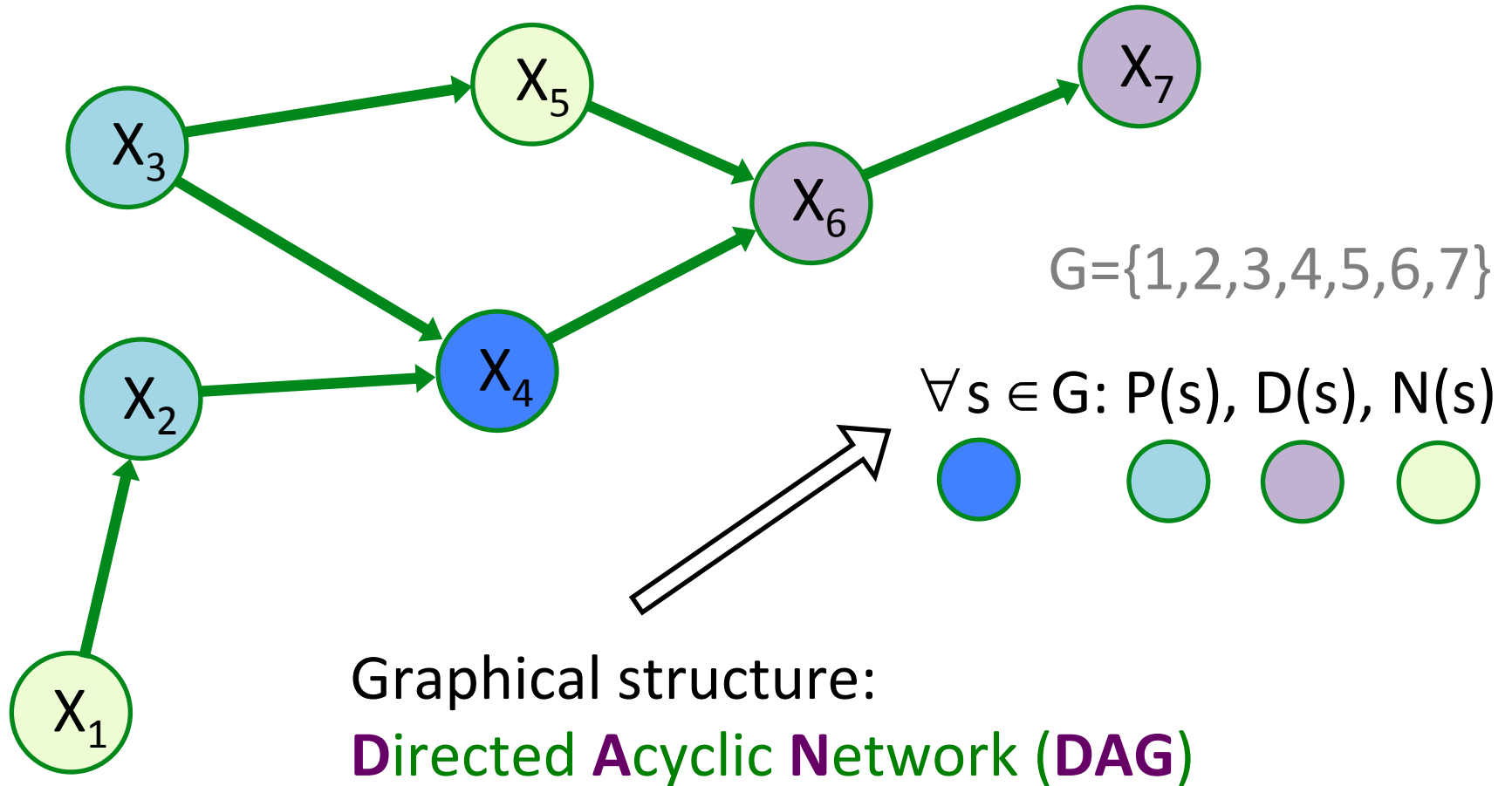
Credal networks: basic setup



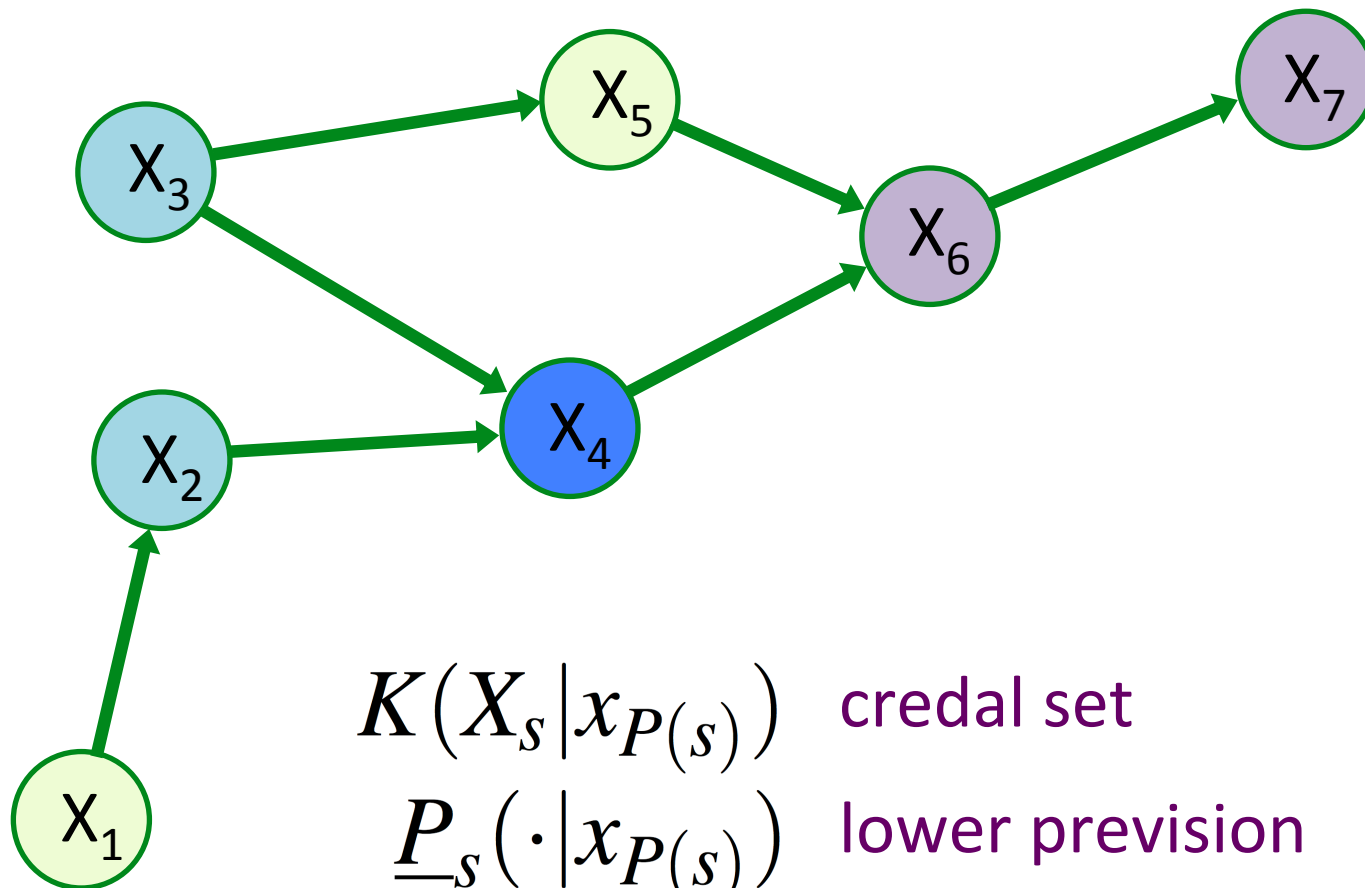
Credal networks: basic setup



Credal networks: basic setup



Credal networks: local uncertainty models



$$K(X_s | x_{P(s)})$$

credal set

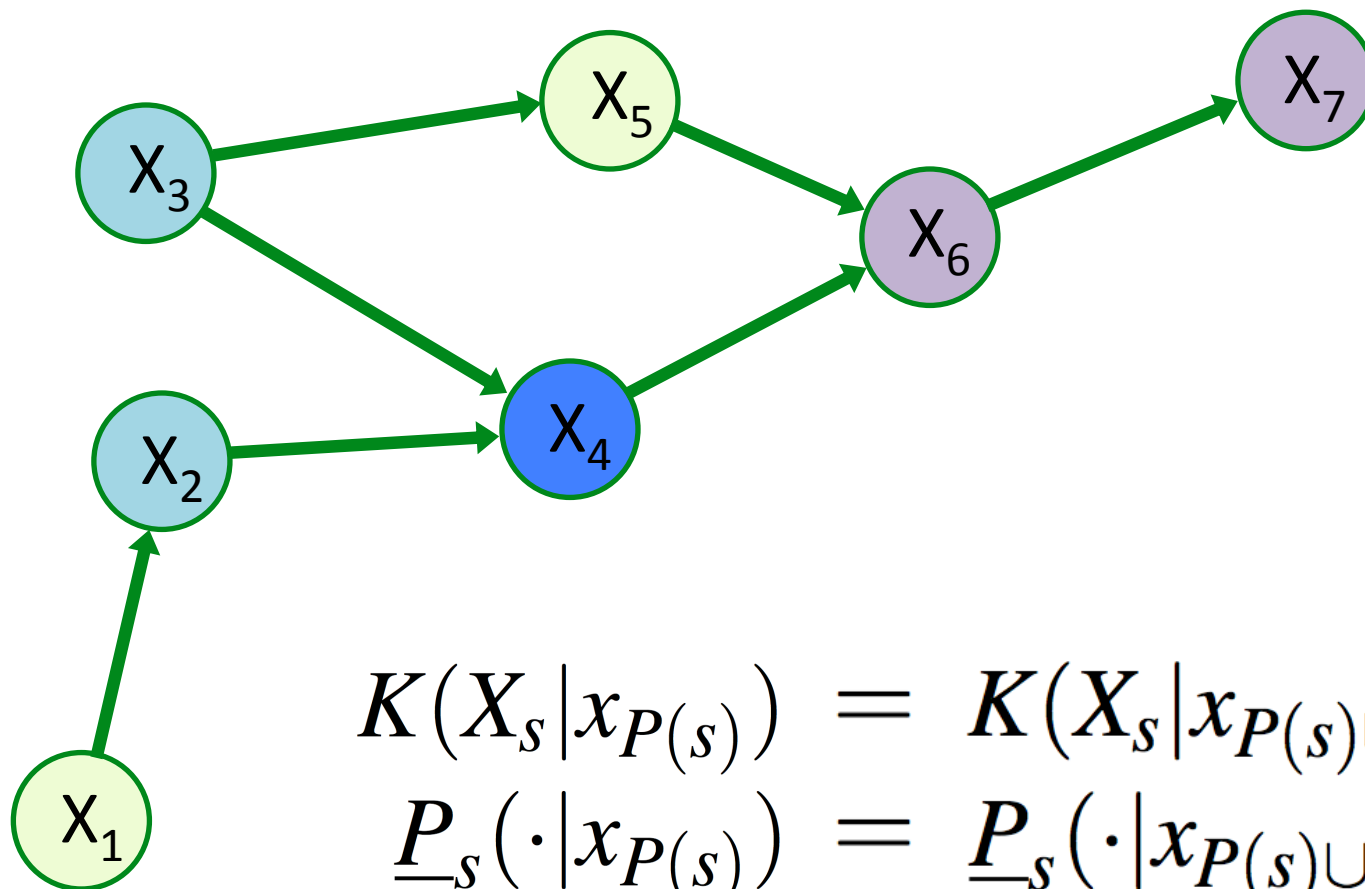
$$P_s(\cdot | x_{P(s)})$$

lower prevision

$$\mathcal{D}_s | x_{P(s)}$$

set of desirable gambles

Credal networks: epistemic irrelevance

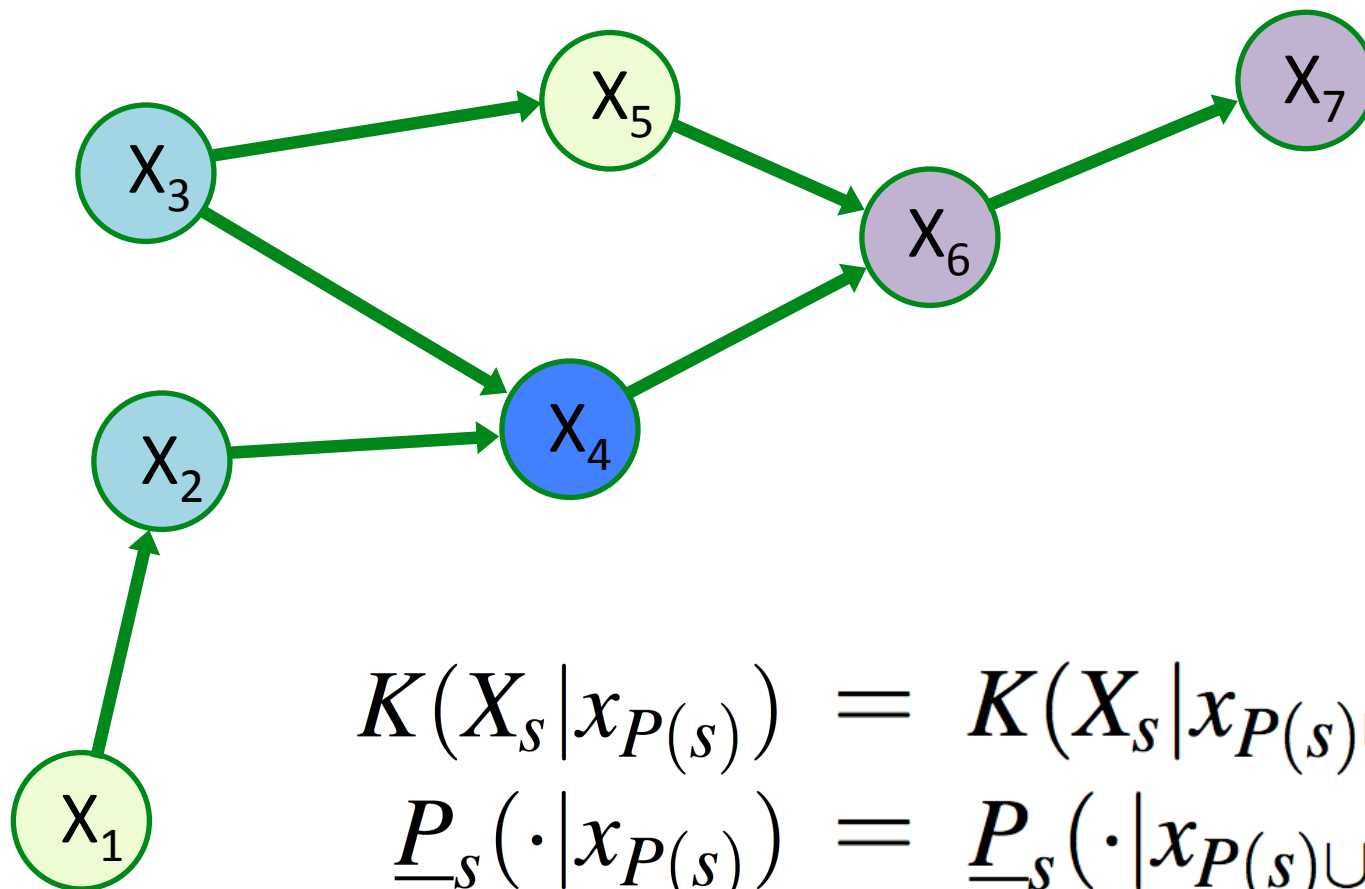


$$K(X_s | x_{P(s)}) = K(X_s | x_{P(s) \cup N(s)})$$

$$\underline{P}_s(\cdot | x_{P(s)}) = \underline{P}_s(\cdot | x_{P(s) \cup N(s)})$$

$$\mathcal{D}_s | x_{P(s)} = \mathcal{D}_s | x_{P(s) \cup N(s)}$$

Credal networks: a joint model



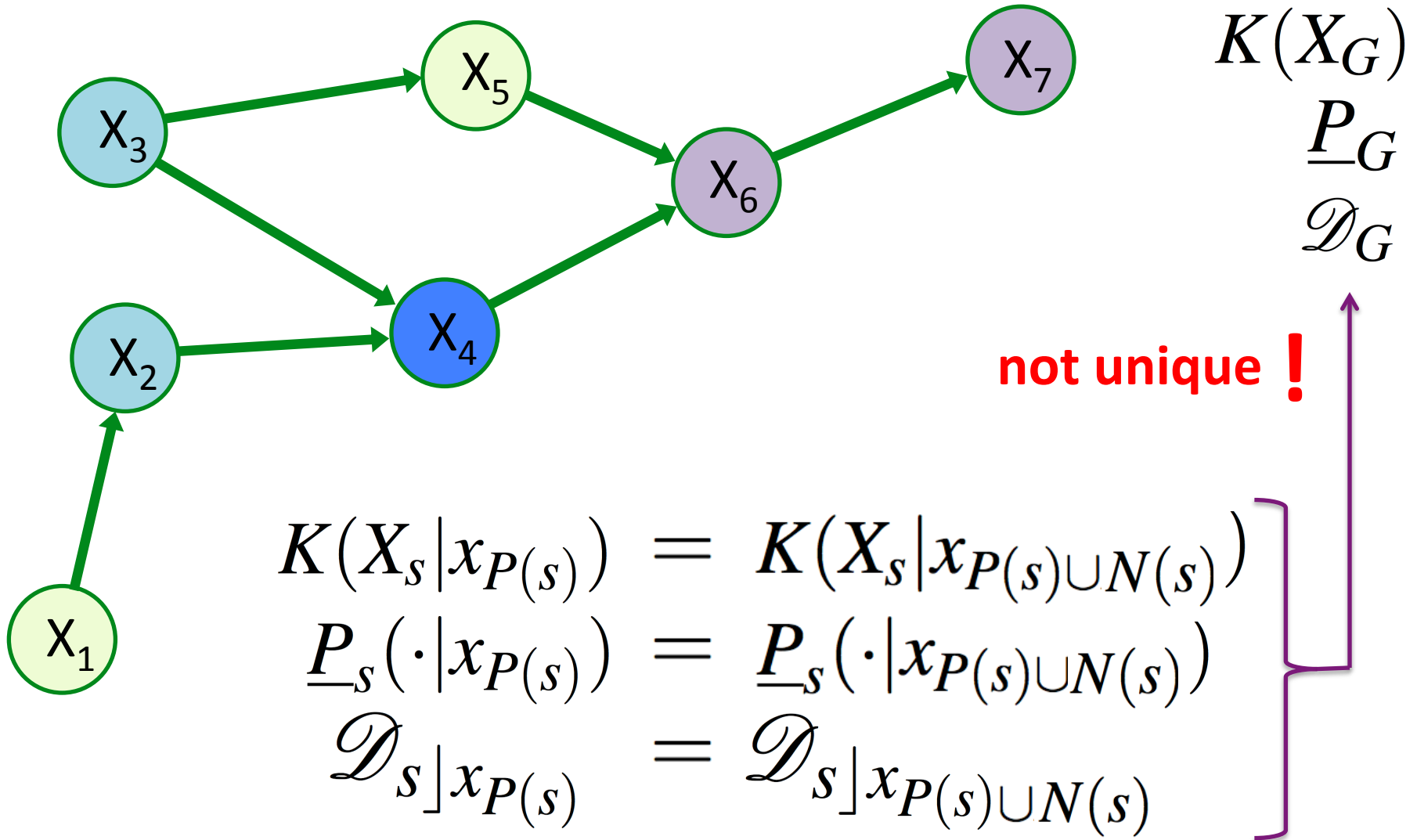
$K(X_G)$
 \underline{P}_G
 \mathcal{D}_G

?

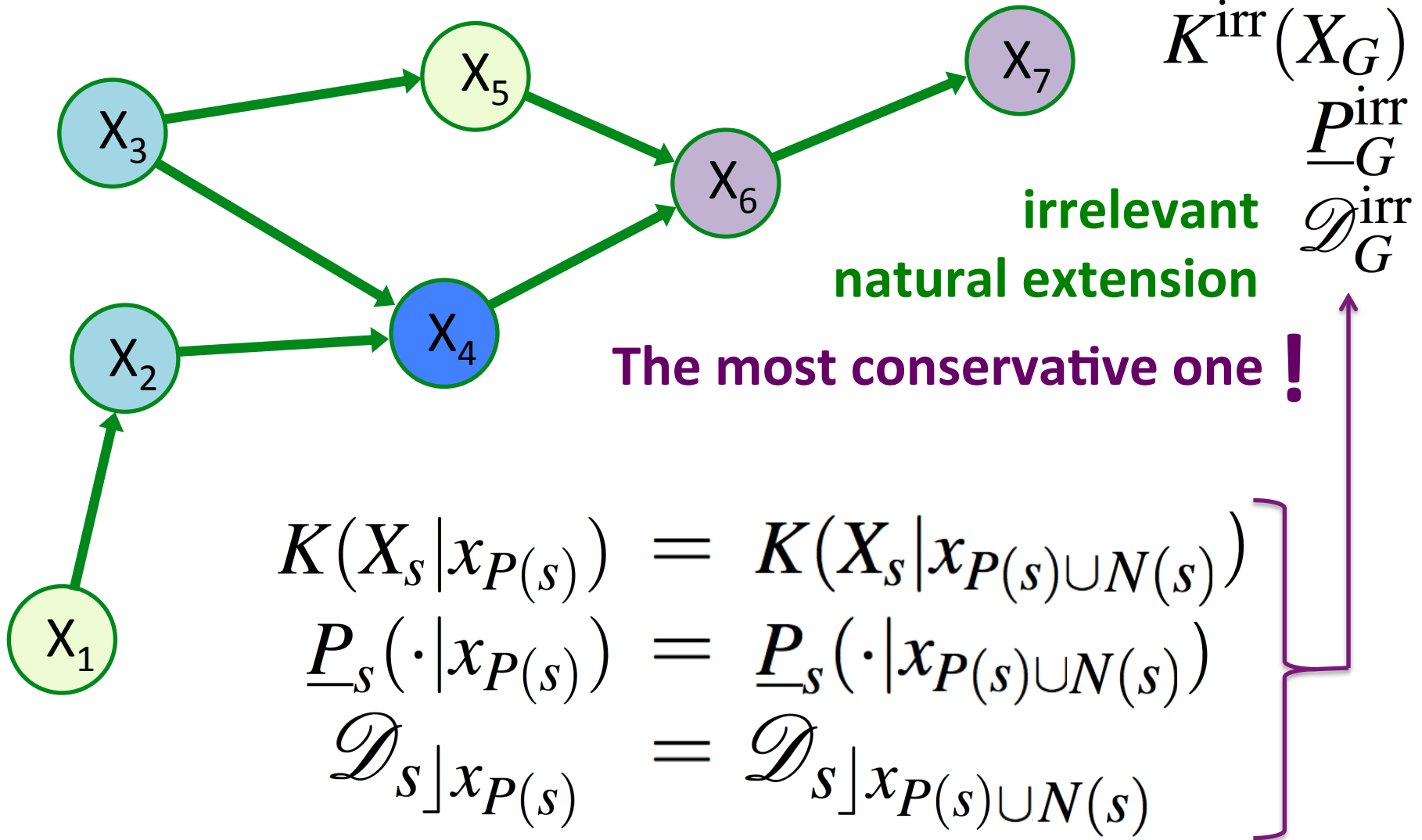
$$\begin{aligned}
 K(X_s | \mathbf{x}_{P(s)}) &= K(X_s | \mathbf{x}_{P(s) \cup N(s)}) \\
 \underline{P}_s(\cdot | \mathbf{x}_{P(s)}) &= \underline{P}_s(\cdot | \mathbf{x}_{P(s) \cup N(s)}) \\
 \mathcal{D}_s | \mathbf{x}_{P(s)} &= \mathcal{D}_s | \mathbf{x}_{P(s) \cup N(s)}
 \end{aligned}$$



Credal networks: a joint model

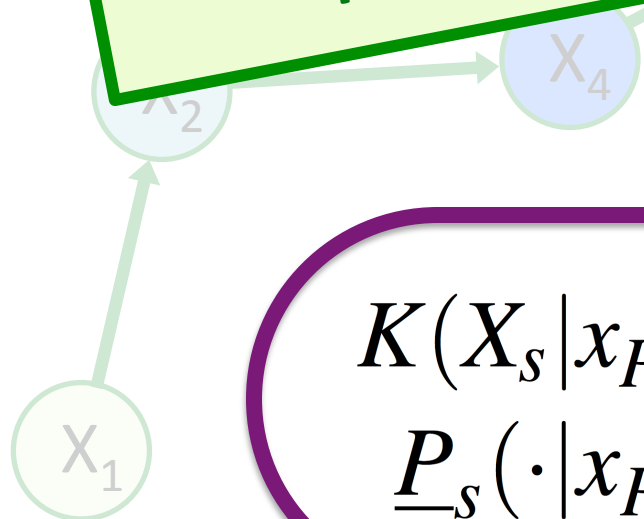


Credal networks: a joint model



Credal networks: a joint model

Allowing for probability zero in credal networks under epistemic irrelevance



$$K^{irr}(X_G)$$

$$\underline{P}_G^{irr}$$

The model is the one!

$$K(X_S | x_{P(S)}) = K(X_S | x_{P(S) \cup N(S)})$$

$$\underline{P}_S(\cdot | x_{P(S)}) = \underline{P}_S(\cdot | x_{P(S) \cup N(S)})$$

$\mathcal{D}_S | x_{P(S)} = \mathcal{D}_S | x_{P(S) \cup N(S)}$

Allowing for probability zero in credal networks under epistemic irrelevance

Jasper De Bock and Gert de Cooman
 (Jasper.DeBock, Gert.deCooman)@UGent.be
 SYSTEMS, Ghent University, Belgium

Abstract of the paper

We generalise Cooman's (2000) concept of a credal network under epistemic irrelevance to the case where lower (and upper) probabilities are allowed to be zero. Our main definition is expressed in terms of coherent lower previsions (BMT) and imposes epistemic irrelevance by means of strong coherence rather than element-wise Bayes's rule (L&R). We also present a number of alternative representations for the resulting joint model, both in terms of lower previsions and credal sets, amongst which an intuitive characterisation of the joint credal set by means of linear constraints (IV). We then apply our method to a simple case: the independent natural extension for two binary variables (V). This allows us to, for the first time, find analytical expressions for the extreme points of this special type of independent product.

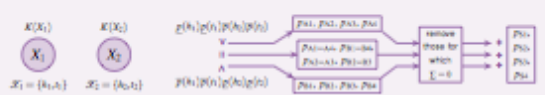
Basic modelling tools (BMT)

We will model a subject's beliefs about the value that a variable X assumes in some set \mathcal{X} by means of two different, although mathematically equivalent, imprecise-probabilistic methods. The approach that is perhaps best known is to use a credal set $K(X)$, defined as a closed convex subset of $\Sigma_{\mathcal{X}}$, which is the set containing all probability mass functions on \mathcal{X} . The second approach is to use the associated coherent lower prevision \underline{E} on $\mathcal{F}(\mathcal{X})$, where $\mathcal{F}(\mathcal{X})$ is the set of all gambles on \mathcal{X} . It is given by $\underline{E}(f) = \min\{P_j(f); P_j \in K(X)\}$ for all $f \in \mathcal{F}(\mathcal{X})$, where P_j is the expectation operator (prevision) for the probability mass function P_j . The credal set of such a coherent lower prevision is given by $K(X) = \{P \in \Sigma_{\mathcal{X}}; (\forall f \in \mathcal{F}(\mathcal{X})) P_j(f) \geq \underline{E}(f)\}$, thereby establishing the mathematical equivalence.

Linear constraints (IV)

It is well known that each local credal set $K(X_i)_{i \in G}$ (1), $i \in G$ and $X_{i \setminus G} \in \mathcal{X}_{i \setminus G}$, is the solution set to a local unitary constraint and a set of linear homogeneous inequalities of the form $\sum_{i \in G} P_i(X_{i \setminus G}) \geq 0$, where P_i takes values in some (possibly infinite, but often finite) set $\Gamma_i(X_{i \setminus G}) \subseteq \mathcal{F}(\mathcal{X}_i)$. We show that, even without the positivity assumption (II), these linear constraints can be used to derive an intuitive characterisation of the irrelevant natural extension $K^{\text{in}}(X_G)$ (18) in terms of linear constraints. $K^{\text{in}}(X_G)$ is the solution set to the global unitary constraint and, for all $i \in G$, $X_{i \setminus G} \in \mathcal{X}_{i \setminus G}$ and $\gamma \in \Gamma_i(X_{i \setminus G})$, a linear homogeneous inequality $\sum_{i \in G} P_i(X_{i \setminus G}) \gamma \geq 0$.

Independent natural extension for two binary variables (V)



	$P(A A)$	$P(A \bar{A})$	$P(\bar{A} A)$	$P(\bar{A} \bar{A})$	\underline{E}
$P(A A)$	$\underline{g}(A A)$	$\underline{g}(A \bar{A})$	$\underline{g}(\bar{A} A)$	$\underline{g}(\bar{A} \bar{A})$	1
$P(A \bar{A})$	$\underline{g}(A A)$	$\underline{g}(A \bar{A})$	$\underline{g}(\bar{A} A)$	$\underline{g}(\bar{A} \bar{A})$	1
$P(\bar{A} A)$	$\underline{g}(A A)$	$\underline{g}(A \bar{A})$	$\underline{g}(\bar{A} A)$	$\underline{g}(\bar{A} \bar{A})$	1
$P(\bar{A} \bar{A})$	$\underline{g}(A A)$	$\underline{g}(A \bar{A})$	$\underline{g}(\bar{A} A)$	$\underline{g}(\bar{A} \bar{A})$	1
$P(A A, A)$	$\underline{g}(A A, A)$	$\underline{g}(A \bar{A}, A)$	$\underline{g}(\bar{A} A, A)$	$\underline{g}(\bar{A} \bar{A}, A)$	$\underline{g}(A A) + \underline{g}(A \bar{A})$
$P(A A, \bar{A})$	$\underline{g}(A A, \bar{A})$	$\underline{g}(A \bar{A}, \bar{A})$	$\underline{g}(\bar{A} A, \bar{A})$	$\underline{g}(\bar{A} \bar{A}, \bar{A})$	$\underline{g}(A A) + \underline{g}(A \bar{A})$
$P(A \bar{A}, A)$	$\underline{g}(A A, A)$	$\underline{g}(A \bar{A}, A)$	$\underline{g}(\bar{A} A, A)$	$\underline{g}(\bar{A} \bar{A}, A)$	$\underline{g}(A A) + \underline{g}(A \bar{A})$
$P(A \bar{A}, \bar{A})$	$\underline{g}(A A, \bar{A})$	$\underline{g}(A \bar{A}, \bar{A})$	$\underline{g}(\bar{A} A, \bar{A})$	$\underline{g}(\bar{A} \bar{A}, \bar{A})$	$\underline{g}(A A) + \underline{g}(A \bar{A})$
$P(\bar{A} A, A)$	$\underline{g}(A A, A)$	$\underline{g}(A \bar{A}, A)$	$\underline{g}(\bar{A} A, A)$	$\underline{g}(\bar{A} \bar{A}, A)$	$\underline{g}(A A) + \underline{g}(\bar{A} \bar{A})$
$P(\bar{A} A, \bar{A})$	$\underline{g}(A A, \bar{A})$	$\underline{g}(A \bar{A}, \bar{A})$	$\underline{g}(\bar{A} A, \bar{A})$	$\underline{g}(\bar{A} \bar{A}, \bar{A})$	$\underline{g}(A A) + \underline{g}(\bar{A} \bar{A})$
$P(\bar{A} \bar{A}, A)$	$\underline{g}(A A, A)$	$\underline{g}(A \bar{A}, A)$	$\underline{g}(\bar{A} A, A)$	$\underline{g}(\bar{A} \bar{A}, A)$	$\underline{g}(A A) + \underline{g}(\bar{A} \bar{A})$
$P(\bar{A} \bar{A}, \bar{A})$	$\underline{g}(A A, \bar{A})$	$\underline{g}(A \bar{A}, \bar{A})$	$\underline{g}(\bar{A} A, \bar{A})$	$\underline{g}(\bar{A} \bar{A}, \bar{A})$	$\underline{g}(A A) + \underline{g}(\bar{A} \bar{A})$

Local uncertainty models (I)

With every node X_i of a finite directed acyclic graph (DAG), we associate a variable X_i , taking values in some finite set \mathcal{X}_i . The set of all nodes is denoted by G . For every subset $S \subseteq G$, the joint variable X_S takes values in \mathcal{X}_S . For every $i \in G$, we denote by $\text{pa}(i)$ the set consisting of the parent nodes of X_i . Similar to what is done in classical Bayesian networks, we attach local uncertainty models to the nodes of the network, conditional on the values of their parents. For all $i \in G$ and every $X_{\text{pa}(i)} \in \mathcal{X}_{\text{pa}(i)}$, we require a credal set $K(X_i|X_{\text{pa}(i)})$ on \mathcal{X}_i , which is a coherent lower prevision $\underline{E}_i(X_i|X_{\text{pa}(i)})$ (BMT).

Imposing epistemic irrelevance (II)

We provide the graphical structure of the network with the following interpretation: for any node $i \in G$, its non-parent non-descendant variables $X_{i \setminus G}$ are epistemically irrelevant to X_i , conditional on $X_{\text{pa}(i)}$. (In our paper, we also require this for subsets of $X_{i \setminus G}$.) We do not impose these additional assumptions on this poster because we have recently discovered that, at least for the unconditional joint model, they are redundant. Put more mathematically, and using $\text{pa}(i)$ as a shorthand notation for $P(\text{pa}(i)|X_{i \setminus G})$, we require that $K(X_i|X_{\text{pa}(i)}) = K(X_i|X_{\text{pa}(i)})$ for all $i \in G$ and $X_{i \setminus G} \in \mathcal{X}_{i \setminus G}$. (1) In order to translate this into a property of a joint model $K(X_G)$, it is often assumed that for every $P \in K(X_G)$, all events have strictly positive probability (Cooman 2000). Under this assumption, $K(X_G)$ can be conditioned by means of element-wise Bayes's rule (applying Bayes's rule to every $P(X_i) \in K(X_i)$), thereby making it possible to impose Eq. (1). We drop this positivity assumption by using an approach based on lower previsions, replacing Eq. (1) by the equivalent (BMT) statement that $\underline{E}_i(X_i|X_{\text{pa}(i)}) = \underline{E}_i(X_i|X_{\text{pa}(i)})$ for all $i \in G$ and $X_{i \setminus G} \in \mathcal{X}_{i \setminus G}$. (2) where, again, the right hand side is provided by the local models (1). Since, without the positivity assumption, conditioning is not uniquely defined, we use a different method for making the conditional models in Eq. (2) consistent with the joint model \underline{E}_G : we require them to be (strongly) coherent. We prove that, in our particular case, this is equivalent to requiring that $\underline{E}_G(X_{i \setminus G} | X_{\text{pa}(i)}) - \underline{E}_G(X_{i \setminus G} | X_{\text{pa}(i)}) = 0$ for all $i \in G$, $X_{i \setminus G} \in \mathcal{X}_{i \setminus G}$ and $g \in \mathcal{F}(\mathcal{X}_i)$. This formula is known as Generalised Bayes's Rule (GBR) and is equivalent to element-wise Bayes's rule if the positivity assumption is satisfied. It should therefore be clear that our approach is an extension of the one by Cooman (2000), coinciding with it under the positivity assumption.

■ Definition by means of strong coherence

The properties that we impose on our network (I & II) can be satisfied by multiple coherent lower previsions \underline{E}_G on $\mathcal{F}(X_G)$. However, amongst them, there is a unique most conservative (pointwise smallest) one. We call it the irrelevant natural extension of the network and denote it by $\underline{E}_G^{\text{in}}$. We show that $\underline{E}_G^{\text{in}}$ is the pointwise smallest coherent lower prevision on $\mathcal{F}(X_G)$ such that for all $i \in G$, $X_{i \setminus G} \in \mathcal{X}_{i \setminus G}$ and $g \in \mathcal{F}(\mathcal{X}_i)$ $\underline{E}_G^{\text{in}}(X_{i \setminus G} | X_{\text{pa}(i)}) - \underline{E}_G^{\text{in}}(X_{i \setminus G} | X_{\text{pa}(i)}) = 0$. We also prove the following simple characterisation of the corresponding credal set $K^{\text{in}}(X_G)$: it consists of all probability mass function P_j on \mathcal{X}_G for which for all $i \in G$ and $X_{i \setminus G} \in \mathcal{X}_{i \setminus G}$ there are a real number $\lambda \geq 0$ and a $P_i(X_i|X_{\text{pa}(i)}) \in K(X_i|X_{\text{pa}(i)})$ such that $\sum_{i \in G} P_j(X_{i \setminus G} | X_{\text{pa}(i)}) = \lambda P_i(X_i|X_{\text{pa}(i)})$, where we use P_j to denote the set consisting of the descendants of the node i . We believe that most of the marginalisation and graphical properties that are presented on our other poster can be translated to the current framework. Combined with linear programming (IV), this might allow for efficient inference algorithms.

Allowing for probability zero in credal networks under epistemic irrelevance

Jasper De Bock and Gert de Cooman
 (Jasper.DeBock, Gert.deCooman)@UGent.be
 SYSTEMS, Ghent University, Belgium

Abstract of the paper

We generalise Cooman's (2000) concept of a credal network under epistemic irrelevance to the case where lower (and upper) probabilities are allowed to be zero. Our main definition is expressed in terms of coherent lower previsions (BMT) and imposes epistemic irrelevance by means of strong coherence rather than element-wise Bayes's rule (L&B II). We also present a number of alternative representations for the resulting joint model, both in terms of lower previsions and credal sets, amongst which an intuitive characterisation of the joint credal set by means of linear constraints (IV). We then apply our method to a simple case: the independent natural extension for two binary variables (V). This allows us to, for the first time, find analytical expressions for the extreme points of this special type of independent product.

Basic modelling tools (BMT)

We will model a subject's beliefs about the value that a variable X assumes in some set \mathcal{X} by means of two different, although mathematically equivalent, imprecise-probabilistic methods. The approach that is perhaps best known is to use a credal set $K(X)$, defined as a closed convex subset of $\Sigma_{\mathcal{X}}$, which is the set containing all probability mass functions on \mathcal{X} . The second approach is to use coherent lower previsions $\mathcal{L}(X)$ on \mathcal{X} , where \mathcal{X} is the set of all gambles $f: \mathcal{X} \rightarrow \mathbb{R}$ and $\mathcal{L}(f) = \inf_{P \in \mathcal{L}(X)} \int_{\mathcal{X}} f(x) P(dx)$, where $\int_{\mathcal{X}}$ is the expectation of the probability mass function. The credal set is given by $K(X) = \{P \in \Sigma_{\mathcal{X}} : \forall f \in \mathcal{L}(X), \int_{\mathcal{X}} f(x) P(dx) \geq \mathcal{L}(f)\}$, thereby establishing the mathematical equivalence.

Linear constraints (IV)

It is well known that each local credal set $K(X_i)_{i \in G}$ (I), $x \in G$ and $x_{(i)} \in \mathcal{X}_{x_{(i)}}$ is the solution set to a local unitary constraint and a set of linear homogeneous inequalities of the form $\sum_{i \in G} P(A_i) \mathbb{1}_{A_i}(x) \geq 0$, where $\mathbb{1}_{A_i}$ takes values in some possibly infinite, but often finite set $\Gamma(A_i, x_{(i)}) \subseteq \mathcal{W}(X_i)$. We show that, even without the positivity assumption (II), these linear constraints can be used to derive an intuitive characterisation of the irrelevant natural extension $K^{\text{in}}(X_G)$ (III) in terms of linear constraints. $K^{\text{in}}(X_G)$ is the solution set to the global unitary constraint and, for all $x \in G$, $x_{(i)} \in \mathcal{X}_{x_{(i)}}$ and $\gamma \in \Gamma(A_i, x_{(i)})$, a linear homogeneous inequality $\sum_{i \in G} P(A_i) \mathbb{1}_{A_i}(x) \geq 0$.

Independent natural extension for two binary variables (V)



For the simple credal network above, $K^{\text{in}}(X_{12})$ is the so-called independent natural extension of $K(X_1)$ and $K(X_2)$. Every $K(X_i)$, with $i \in \{1, 2\}$, is fully determined by the lower probability $\beta(A_i)$ and upper probability $\beta(A_i)$ of heads. The probability of tails is bounded by $\beta(A_i) = 1 - \beta(A_i)$ and $\beta(A_i) = 1 - \beta(A_i)$. Using the linear constraints in (IV), we have derived analytical expressions for the extreme points of $K^{\text{in}}(X_{12})$. They can be found using the table and diagram to the right; see our paper for more details.

$\beta(A_1)$	$\beta(A_2)$	$\beta(A_1, A_2)$	$\beta(A_1, \bar{A}_2)$	$\beta(\bar{A}_1, A_2)$	$\beta(\bar{A}_1, \bar{A}_2)$
0	0	0	0	0	0
0	1	0	0	0	0
1	0	0	0	0	0
1	1	0	0	0	0
0	0	0	0	0	0
0	1	0	0	0	0
1	0	0	0	0	0
1	1	0	0	0	0

Local uncertainty models (I)

With every node x of a finite directed acyclic graph (DAG), we associate a variable X_x taking values in some finite, non-empty set \mathcal{X}_x . The set of all nodes is denoted by G . For every subset $S \subseteq G$, the joint variable X_S takes values in $\mathcal{X}_S := \prod_{x \in S} \mathcal{X}_x$. For every $x \in G$, we denote by $\mathcal{P}(x)$ the set consisting of the parent nodes of x . Similar to what is done in classical Bayesian networks, we attach local uncertainty models to the nodes of the network, conditional on the value of their parents. For all $x \in G$ and every instantiation $x_{\mathcal{P}(x)} \in \mathcal{X}_{\mathcal{P}(x)}$, we require a credal set $K(X_x | x_{\mathcal{P}(x)})$ or, equivalently, a coherent lower prevision $\mathcal{L}_x(X_x | x_{\mathcal{P}(x)})$ on $\mathcal{W}(X_x)$ (BMT).

Imposing epistemic irrelevance (II)

We provide the graphical structure of the network with the following interpretation: for any node $x \in G$, its non-parent non-descendant variables $X_{x_{(i)}}$ are epistemically irrelevant to X_x , conditional on $x_{\mathcal{P}(x)}$. (In our paper, we also require this for subsets of $X_{x_{(i)}}$). We do not impose these additional assu-

ptions. In the unconditional joint model, they are redundant. Put more mathematically, and using $\mathbb{P}(X_x)$ as a shorthand notation for $\mathbb{P}(X_x | U_{x_{(i)}})$, we require that $K(X_x | x_{\mathcal{P}(x)}) = K(X_x | x_{\mathcal{P}(x)}, x_{x_{(i)}})$ for all $x \in G$ and $x_{\mathcal{P}(x)} \in \mathcal{X}_{\mathcal{P}(x)}$. (I)

In order to translate this into a property of a joint model $K(X_G)$, it is often assumed that for every $P(X_G) \in K(X_G)$, all events have strictly positive probability (Cooman 2000). Under this assumption, $K(X_G)$ can be conditioned by means of element-wise Bayes's rule (applying Bayes's rule to every $P(X_G) \in K(X_G)$), thereby making it possible to impose Eq. (I).

We drop this positivity assumption by using an approach based on lower previsions, replacing Eq. (I) by the equivalent (BMT) statement that

$$\mathcal{L}_x(X_x | x_{\mathcal{P}(x)}) = \mathcal{L}_x(X_x | x_{\mathcal{P}(x)}, x_{x_{(i)}}) \quad (2)$$

where, again, the right hand side is provided by the local models (I). Since, without the positivity assumption, conditioning is not uniquely defined, we use a different method for making the conditional models in Eq. (2) consistent with the joint model \mathcal{L}_x ; we require them to be (strongly) coherent. We prove that, in our particular case, this is equivalent to requiring that

$$\mathcal{L}_x(X_x | x_{\mathcal{P}(x)}, x_{x_{(i)}}) - \mathcal{L}_x(X_x | x_{\mathcal{P}(x)}) = 0$$

for all $x \in G$, $x_{\mathcal{P}(x)} \in \mathcal{X}_{\mathcal{P}(x)}$ and $x_{x_{(i)}} \in \mathcal{X}_{x_{(i)}}$. This formula is known as Generalised Bayes's Rule (GBR) and is equivalent to element-wise Bayes's rule if the positivity assumption is satisfied. It should therefore be clear that our approach is an extension of the one by Cooman (2000), coinciding with it under the positivity assumption.

Irrelevant natural extension (III)

The properties that we impose on our network (I & II) can be satisfied by multiple coherent lower previsions \mathcal{L}_x on $\mathcal{W}(X_x)$. However, amongst them, there is a unique most conservative (pointwise smallest) one. We call it the irrelevant natural extension of the network and denote it by \mathcal{L}^{in} . We show that \mathcal{L}^{in} is the pointwise smallest coherent lower prevision on $\mathcal{W}(X_G)$ such that for all $x \in G$, $x_{\mathcal{P}(x)} \in \mathcal{X}_{\mathcal{P}(x)}$ and $g \in \mathcal{W}(X_x)$

$$\mathcal{L}^{\text{in}}(X_x | x_{\mathcal{P}(x)}, g) - \mathcal{L}^{\text{in}}(X_x | x_{\mathcal{P}(x)}, g) = 0$$

We also prove the following simple characterisation of the corresponding credal set $K^{\text{in}}(X_G)$: it consists of all probability mass functions $P(X_G)$ on \mathcal{X}_G for which for all $x \in G$ and $x_{\mathcal{P}(x)} \in \mathcal{X}_{\mathcal{P}(x)}$ there are a real number $\lambda \geq 0$ and a $P(X_x | x_{\mathcal{P}(x)}) \in K(X_x | x_{\mathcal{P}(x)})$ such that

$$\sum_{x \in G} P(X_G) \mathbb{1}_{A_x}(x) = \lambda P(X_x | x_{\mathcal{P}(x)})$$

where we use $\mathcal{D}(x)$ to denote the set consisting of the descendants of the node x .

We believe that most of the marginalisation and graphical properties that are presented on our other poster can be translated to the current framework. Combined with linear programming (IV), this might allow for efficient inference algorithms.

- Definition by means of strong coherence
- Description in terms of linear constraints

Allowing for probability zero in credal networks under epistemic irrelevance

Jasper De Bock and Gert de Cooman
 (Jasper.DeBock, Gert.deCooman)@UGent.be
 SYStEMS, Ghent University, Belgium

Abstract of the paper

We generalise Cooman's (2000) concept of a credal network under epistemic irrelevance to the case where lower (and upper) probabilities are allowed to be zero. Our main definition is expressed in terms of coherent lower previsions (BMT) and imposes epistemic irrelevance by means of strong coherence rather than element-wise Bayes's rule (L&B II). We also present a number of alternative representations for the resulting joint model, both in terms of lower previsions and credal sets, amongst which an intuitive characterisation of the joint credal set by means of linear constraints (IV). We then apply our method to a simple case: the independent natural extension for two binary variables (V). This allows us to, for the first time, find analytical expressions for the extreme points of this special type of independent product.

Basic modelling tools (BMT)

We will model a subject's beliefs about the value that a variable X assumes in some set \mathcal{X} by means of two different, although mathematically equivalent, imprecise-probabilistic methods. The approach that is perhaps best known is to use a credal set $K(X)$, defined as a closed convex subset of $\Sigma_{\mathcal{X}}$, which is the set containing all probability mass functions on \mathcal{X} . The second approach is to use the associated coherent lower prevision \underline{E} on $\mathcal{W}(\mathcal{X})$, where $\mathcal{W}(\mathcal{X})$ is the set of all gambles on \mathcal{X} . It is given by $\underline{E}(f) = \min\{E(f); p(X) \in K(X)\}$ for all $f \in \mathcal{W}(\mathcal{X})$, where E is the expectation operator (prevision) for the probability mass function $p(X)$. The credal set of such a coherent lower prevision is given by $K(X) = \{p \in \Sigma_{\mathcal{X}}; \forall f \in \mathcal{W}(\mathcal{X}) [p(f) \geq \underline{E}(f)]\}$, thereby establishing the mathematical equivalence.

Linear constraints (IV)

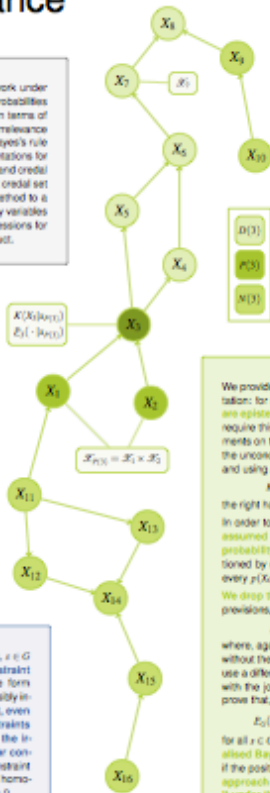
It is well known that each local credal set $K(X_i|_{X_{\setminus i}})$ (I), $x \in G$ and $x_{\setminus i} \in \mathcal{X}_{\setminus i}$, is the solution set to a local unitary constraint and a set of linear homogeneous inequalities of the form $\sum_{j \in \mathcal{X}_i} p_j(x_{\setminus i}, \gamma_j) \geq 0$, where γ takes values in some possibly infinite, but often finite, set $\Gamma(x_{\setminus i}, \gamma) \subseteq \mathcal{W}(\mathcal{X}_i)$. We show that, even without the positivity assumption (II), these linear constraints can be used to derive an intuitive characterisation of the irrelevant natural extension $K^{\text{irr}}(X_G)$ (III) in terms of linear constraints. $K^{\text{irr}}(X_G)$ is the solution set to the global unitary constraint and, for all $x \in G$, $x_{\setminus i} \in \mathcal{X}_{\setminus i}$ and $\gamma \in \Gamma(x_{\setminus i}, \gamma)$, a linear homogeneous inequality $\sum_{j \in \mathcal{X}_i} p_j(x_{\setminus i}, \gamma_j) \geq 0$.

Independent natural extension for two binary variables (V)



For the simple credal network above, $K^{\text{irr}}(X_{12})$ is the so-called independent natural extension of $K(X_1)$ and $K(X_2)$. Every $K(X_i)$, with $i \in \{1, 2\}$, is fully determined by the lower probability $p(A)$ and upper probability $\beta(A)$ of heads. The probability of tails is bounded by $g(A) = 1 - p(A)$ and $\beta(A) = 1 - g(A)$. Using the linear constraints in (IV), we have derived analytical expressions for the extreme points of $K^{\text{irr}}(X_{12})$. They can be found using the table and diagram to the right; see our paper for more details.

$p(A), \beta(A)$	$p(A), \beta(A)$	$p(A), \beta(A)$	$p(A), \beta(A)$	β
$g(A) \geq \beta(A)$	$g(A) \geq \beta(A)$	$g(A) \geq \beta(A)$	$g(A) \geq \beta(A)$	1
$g(A) \geq \beta(A)$	$g(A) \geq \beta(A)$	$g(A) \geq \beta(A)$	$g(A) \geq \beta(A)$	1
$g(A) \geq \beta(A)$	$g(A) \geq \beta(A)$	$g(A) \geq \beta(A)$	$g(A) \geq \beta(A)$	1
$g(A) \geq \beta(A)$	$g(A) \geq \beta(A)$	$g(A) \geq \beta(A)$	$g(A) \geq \beta(A)$	1
$g(A) \geq \beta(A)$	$g(A) \geq \beta(A)$	$g(A) \geq \beta(A)$	$g(A) \geq \beta(A)$	1
$g(A) \geq \beta(A)$	$g(A) \geq \beta(A)$	$g(A) \geq \beta(A)$	$g(A) \geq \beta(A)$	1
$g(A) \geq \beta(A)$	$g(A) \geq \beta(A)$	$g(A) \geq \beta(A)$	$g(A) \geq \beta(A)$	1
$g(A) \geq \beta(A)$	$g(A) \geq \beta(A)$	$g(A) \geq \beta(A)$	$g(A) \geq \beta(A)$	1
$g(A) \geq \beta(A)$	$g(A) \geq \beta(A)$	$g(A) \geq \beta(A)$	$g(A) \geq \beta(A)$	1
$g(A) \geq \beta(A)$	$g(A) \geq \beta(A)$	$g(A) \geq \beta(A)$	$g(A) \geq \beta(A)$	1
$g(A) \geq \beta(A)$	$g(A) \geq \beta(A)$	$g(A) \geq \beta(A)$	$g(A) \geq \beta(A)$	1
$g(A) \geq \beta(A)$	$g(A) \geq \beta(A)$	$g(A) \geq \beta(A)$	$g(A) \geq \beta(A)$	1
$g(A) \geq \beta(A)$	$g(A) \geq \beta(A)$	$g(A) \geq \beta(A)$	$g(A) \geq \beta(A)$	1
$g(A) \geq \beta(A)$	$g(A) \geq \beta(A)$	$g(A) \geq \beta(A)$	$g(A) \geq \beta(A)$	1
$g(A) \geq \beta(A)$	$g(A) \geq \beta(A)$	$g(A) \geq \beta(A)$	$g(A) \geq \beta(A)$	1
$g(A) \geq \beta(A)$	$g(A) \geq \beta(A)$	$g(A) \geq \beta(A)$	$g(A) \geq \beta(A)$	1
$g(A) \geq \beta(A)$	$g(A) \geq \beta(A)$	$g(A) \geq \beta(A)$	$g(A) \geq \beta(A)$	1
$g(A) \geq \beta(A)$	$g(A) \geq \beta(A)$	$g(A) \geq \beta(A)$	$g(A) \geq \beta(A)$	1
$g(A) \geq \beta(A)$	$g(A) \geq \beta(A)$	$g(A) \geq \beta(A)$	$g(A) \geq \beta(A)$	1
$g(A) \geq \beta(A)$	$g(A) \geq \beta(A)$	$g(A) \geq \beta(A)$	$g(A) \geq \beta(A)$	1



Local uncertainty models (I)

With every node x of a finite directed acyclic graph (DAG), we associate a variable X_x . The set of all nodes is denoted by G . For every subset $S \subseteq G$, the joint variable X_S takes values in $\mathcal{X}_S = \prod_{x \in S} \mathcal{X}_x$. For every $x \in G$, we denote by $\text{Pa}(x)$ the set consisting of the parent nodes of x . Similar to what is done in classical Bayesian networks, we attach local uncertainty models to the nodes of the network, conditional on the value of their parents. For all $x \in G$ and every instantiation $x_{\text{Pa}(x)} \in \mathcal{X}_{\text{Pa}(x)}$, we require a credal set $K(X_x|x_{\text{Pa}(x)})$ or, equivalently, a coherent lower prevision $\underline{E}_x(\cdot|x_{\text{Pa}(x)})$ on $\mathcal{W}(\mathcal{X}_x)$ (BMT).

Imposing epistemic irrelevance (II)

We provide the graphical structure of the network with the following interpretation: for any node $x \in G$, its non-parent non-descendant variables $X_{\setminus x}$ are epistemically irrelevant to X_x , conditional on $X_{\text{Pa}(x)}$. (In our paper, we also require this for subsets of $X_{\setminus x}$.) We do not impose these additional assumptions on this poster because we have recently discovered that, at least for the unconditional joint model, they are redundant. But more importantly, and using $\text{Pa}(x)$ as a shorthand notation for $\text{Pa}(G \setminus \{x\})$, we require that $K(X_x|x_{\text{Pa}(x)}) = K(X_x|x_{\text{Pa}(x)})$ for all $x \in G$ and $x_{\text{Pa}(x)} \in \mathcal{X}_{\text{Pa}(x)}$. (I) In order to translate this into a property of a joint model $K(X_G)$, it is often assumed that for every $p(X_G) \in K(X_G)$, all events have strictly positive probability (Cooman 2000). Under this assumption, $K(X_G)$ can be conditioned by means of element-wise Bayes's rule (applying Bayes's rule to every $p(X_G) \in K(X_G)$), thereby making it possible to impose Eq. (I).

We drop this positivity assumption by using an approach based on lower previsions, replacing Eq. (I) by the equivalent (BMT) statement that $\underline{E}_x(\cdot|x_{\text{Pa}(x)}) = \underline{E}_x(\cdot|x_{\text{Pa}(x)})$ for all $x \in G$ and $x_{\text{Pa}(x)} \in \mathcal{X}_{\text{Pa}(x)}$. (II) Since, without the positivity assumption, conditioning is not uniquely defined, we use a different method for making the conditional models in Eq. (I) consistent with the joint model \underline{E}_G : we require them to be (strongly) coherent. We prove that, in our particular case, this is equivalent to requiring that $\underline{E}_G(K_{x_{\text{Pa}(x)}}[x] - \underline{E}_x(\cdot|x_{\text{Pa}(x)})) = 0$ and $\underline{E}_G(K_{x_{\text{Pa}(x)}}[x] - \underline{E}_x(\cdot|x_{\text{Pa}(x)})) = 0$ for all $x \in G$, $x_{\text{Pa}(x)} \in \mathcal{X}_{\text{Pa}(x)}$ and $g \in \mathcal{W}(\mathcal{X}_x)$. This formula is known as Generalised Bayes's Rule (GBR) and is equivalent to element-wise Bayes's rule if the positivity assumption is satisfied. It should therefore be clear that our approach is an extension of the one by Cooman (2000), coinciding with it under the positivity assumption.

Irrelevant natural extension (III)

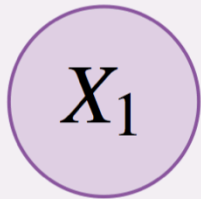
The properties that we impose on our model can be satisfied by multiple coherent \underline{E}_G on $\mathcal{W}(X_G)$. However, amongst the unique most conservative (pointwise) one. We call it the irrelevant natural extension and denote it by $\underline{E}_G^{\text{irr}}$. We show that $\underline{E}_G^{\text{irr}}$ is the pointwise smallest coherent lower prevision such that for all $x \in G$, $x_{\text{Pa}(x)} \in \mathcal{X}_{\text{Pa}(x)}$ and $g \in \mathcal{W}(\mathcal{X}_x)$, $\underline{E}_G^{\text{irr}}(K_{x_{\text{Pa}(x)}}[x] - \underline{E}_x(\cdot|x_{\text{Pa}(x)})) = 0$. We also prove the following simple characterisation: it consists of all probability mass function $p(X_G)$ on \mathcal{X}_G for which for all $x \in G$ and $x_{\text{Pa}(x)} \in \mathcal{X}_{\text{Pa}(x)}$, there are a real number $\lambda \geq 0$ and a $p(X_G|_{x_{\text{Pa}(x)}}) \in K(X_G|_{x_{\text{Pa}(x)}})$ such that $\sum_{x_{\setminus x} \in \mathcal{X}_{\setminus x}} p(X_G|_{x_{\setminus x}, x_{\text{Pa}(x)}}) = \lambda p(X_G|_{x_{\text{Pa}(x)}})$, where we use $D(x)$ to denote the set consisting of the descendants of the node x . We believe that most of the marginalisation and graphical properties that are presented on our other poster can be translated to the current framework. Combined with linear programming (IV), this might allow for efficient inference algorithms.

- Definition by means of strong coherence
- Description in terms of linear constraints
- Independent natural extension for two binary variables

Independent natural extension for two binary variables

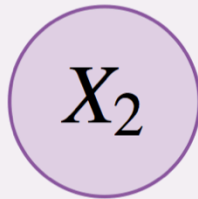
$$K^{\text{irr}}(X_G) = K(X_1) \otimes K(X_2)$$

$K(X_1)$



$$\mathcal{X}_1 = \{h_1, t_1\}$$

$K(X_2)$



$$\mathcal{X}_2 = \{h_2, t_2\}$$

$$\bar{p}(t_1)p(h_1, h_2) - \underline{p}(h_1)p(t_1, h_2) \geq 0$$

$$-\underline{p}(t_1)p(h_1, h_2) + \bar{p}(h_1)p(t_1, h_2) \geq 0$$

$$\bar{p}(t_1)p(h_1, t_2) - \underline{p}(h_1)p(t_1, t_2) \geq 0$$

$$-\underline{p}(t_1)p(h_1, t_2) + \bar{p}(h_1)p(t_1, t_2) \geq 0$$

$$\bar{p}(t_2)p(h_1, h_2) - \underline{p}(h_2)p(h_1, t_2) \geq 0$$

$$-\underline{p}(t_2)p(h_1, h_2) + \bar{p}(h_2)p(h_1, t_2) \geq 0$$

$$\bar{p}(t_2)p(t_1, h_2) - \underline{p}(h_2)p(t_1, t_2) \geq 0$$

$$-\underline{p}(t_2)p(t_1, h_2) + \bar{p}(h_2)p(t_1, t_2) \geq 0$$

$$K(X_i) = \{p \in \Sigma_{\mathcal{X}_i} : p(h_i) \in [\underline{p}(h_i), \bar{p}(h_i)]\}$$

$$\bar{p}(t_i) := 1 - \underline{p}(h_i)$$

$$\underline{p}(t_i) := 1 - \bar{p}(h_i)$$

Independent natural extension for two binary variables

Analytical expressions for the extreme points

	$p(h_1, h_2) \Sigma$	$p(h_1, t_2) \Sigma$	$p(t_1, h_2) \Sigma$	$p(t_1, t_2) \Sigma$	Σ
PS_1	$\underline{p}(h_1)\underline{p}(h_2)$	$\underline{p}(h_1)\bar{p}(t_2)$	$\bar{p}(t_1)\underline{p}(h_2)$	$\bar{p}(t_1)\bar{p}(t_2)$	1
PS_2	$\underline{p}(h_1)\bar{p}(h_2)$	$\underline{p}(h_1)\underline{p}(t_2)$	$\bar{p}(t_1)\bar{p}(h_2)$	$\bar{p}(t_1)\underline{p}(t_2)$	1
PS_3	$\bar{p}(h_1)\underline{p}(h_2)$	$\bar{p}(h_1)\bar{p}(t_2)$	$\underline{p}(t_1)\underline{p}(h_2)$	$\underline{p}(t_1)\bar{p}(t_2)$	1
PS_4	$\bar{p}(h_1)\bar{p}(h_2)$	$\bar{p}(h_1)\underline{p}(t_2)$	$\underline{p}(t_1)\bar{p}(h_2)$	$\underline{p}(t_1)\underline{p}(t_2)$	1
PA_1	$\underline{p}(h_1)\bar{p}(h_1)\underline{p}(h_2)$	$\underline{p}(h_1)\bar{p}(h_1)\bar{p}(t_2)$	$\bar{p}(t_1)\bar{p}(h_1)\underline{p}(h_2)$	$\underline{p}(h_1)\underline{p}(t_1)\bar{p}(t_2)$	$\underline{p}(h_1)\bar{p}(t_2) + \bar{p}(h_1)\underline{p}(h_2)$
PA_2	$\underline{p}(h_1)\bar{p}(h_1)\bar{p}(h_2)$	$\underline{p}(h_1)\bar{p}(h_1)\underline{p}(t_2)$	$\underline{p}(h_1)\underline{p}(t_1)\bar{p}(h_2)$	$\bar{p}(t_1)\bar{p}(h_1)\underline{p}(t_2)$	$\underline{p}(h_1)\bar{p}(h_2) + \bar{p}(h_1)\underline{p}(t_2)$
PA_3	$\bar{p}(h_1)\bar{p}(t_1)\underline{p}(h_2)$	$\underline{p}(t_1)\underline{p}(h_1)\bar{p}(t_2)$	$\underline{p}(t_1)\bar{p}(t_1)\underline{p}(h_2)$	$\underline{p}(t_1)\bar{p}(t_1)\bar{p}(t_2)$	$\underline{p}(t_1)\bar{p}(t_2) + \bar{p}(t_1)\underline{p}(h_2)$
PA_4	$\underline{p}(t_1)\underline{p}(h_1)\bar{p}(h_2)$	$\bar{p}(h_1)\bar{p}(t_1)\underline{p}(t_2)$	$\underline{p}(t_1)\bar{p}(t_1)\bar{p}(h_2)$	$\underline{p}(t_1)\bar{p}(t_1)\underline{p}(t_2)$	$\underline{p}(t_1)\bar{p}(h_2) + \bar{p}(t_1)\underline{p}(t_2)$
PB_1	$\underline{p}(h_2)\bar{p}(h_2)\underline{p}(h_1)$	$\bar{p}(t_2)\bar{p}(h_2)\underline{p}(h_1)$	$\underline{p}(h_2)\bar{p}(h_2)\bar{p}(t_1)$	$\underline{p}(h_2)\underline{p}(t_2)\bar{p}(t_1)$	$\underline{p}(h_2)\bar{p}(t_1) + \bar{p}(h_2)\underline{p}(h_1)$
PB_2	$\bar{p}(h_2)\bar{p}(t_2)\underline{p}(h_1)$	$\underline{p}(t_2)\bar{p}(t_2)\underline{p}(h_1)$	$\underline{p}(t_2)\underline{p}(h_2)\bar{p}(t_1)$	$\underline{p}(t_2)\bar{p}(t_2)\bar{p}(t_1)$	$\underline{p}(t_2)\bar{p}(t_1) + \bar{p}(t_2)\underline{p}(h_1)$
PB_3	$\underline{p}(h_2)\bar{p}(h_2)\bar{p}(h_1)$	$\underline{p}(h_2)\underline{p}(t_2)\bar{p}(h_1)$	$\underline{p}(h_2)\bar{p}(h_2)\underline{p}(t_1)$	$\bar{p}(t_2)\bar{p}(h_2)\underline{p}(t_1)$	$\underline{p}(h_2)\bar{p}(h_1) + \bar{p}(h_2)\underline{p}(t_1)$
PB_4	$\underline{p}(t_2)\underline{p}(h_2)\bar{p}(h_1)$	$\underline{p}(t_2)\bar{p}(t_2)\bar{p}(h_1)$	$\bar{p}(h_2)\bar{p}(t_2)\underline{p}(t_1)$	$\underline{p}(t_2)\bar{p}(t_2)\underline{p}(t_1)$	$\underline{p}(t_2)\bar{p}(h_1) + \bar{p}(t_2)\underline{p}(t_1)$

Independent natural extension for two binary variables

Analytical expressions for the extreme points

	$p(h_1, h_2) \Sigma$	$p(h_1, t_2) \Sigma$	$p(t_1, h_2) \Sigma$	$p(t_1, t_2) \Sigma$	Σ
PS1	$\underline{p}(h_1)\underline{p}(h_2)$	$\underline{p}(h_1)\bar{p}(t_2)$	$\bar{p}(t_1)\underline{p}(h_2)$	$\bar{p}(t_1)\bar{p}(t_2)$	$\underline{p}(h_1)\underline{p}(h_2) + \underline{p}(h_1)\bar{p}(t_2) + \bar{p}(t_1)\underline{p}(h_2) + \bar{p}(t_1)\bar{p}(t_2)$
PS2	$\underline{p}(h_1)\bar{p}(h_2)$	$\underline{p}(h_1)\underline{p}(t_2)$	$\underline{p}(t_1)\underline{p}(h_2)$	$\underline{p}(t_1)\underline{p}(t_2)$	$\underline{p}(h_1)\bar{p}(h_2) + \underline{p}(h_1)\underline{p}(t_2) + \underline{p}(t_1)\underline{p}(h_2) + \underline{p}(t_1)\underline{p}(t_2)$
PS3	$\bar{p}(h_1)\underline{p}(h_2)$	$\bar{p}(h_1)\bar{p}(t_2)$	$\bar{p}(t_1)\bar{p}(h_2)$	$\bar{p}(t_1)\bar{p}(t_2)$	$\bar{p}(h_1)\underline{p}(h_2) + \bar{p}(h_1)\bar{p}(t_2) + \bar{p}(t_1)\bar{p}(h_2) + \bar{p}(t_1)\bar{p}(t_2)$
PS4	$\bar{p}(h_1)\bar{p}(h_2)$	$\bar{p}(h_1)\underline{p}(t_2)$	$\bar{p}(t_1)\underline{p}(h_2)$	$\bar{p}(t_1)\underline{p}(t_2)$	$\bar{p}(h_1)\bar{p}(h_2) + \bar{p}(h_1)\underline{p}(t_2) + \bar{p}(t_1)\underline{p}(h_2) + \bar{p}(t_1)\underline{p}(t_2)$
PA1	$\underline{p}(h_1)\underline{p}(t_1)\bar{p}(h_2)\bar{p}(t_2)$				
PA2	$\underline{p}(h_1)\bar{p}(h_2)\underline{p}(t_1)\bar{p}(t_2)$				
PA3	$\bar{p}(h_1)\underline{p}(h_2)\bar{p}(t_1)\underline{p}(t_2)$				
PA4	$\underline{p}(t_1)\underline{p}(h_2)\bar{p}(h_1)\bar{p}(t_2)$				
PB1	$\underline{p}(h_2)\bar{p}(h_2)\underline{p}(h_1)$	$\bar{p}(t_2)\bar{p}(h_2)\underline{p}(h_1)$	$\underline{p}(h_2)\bar{p}(h_2)\bar{p}(t_1)$	$\underline{p}(h_2)\underline{p}(t_2)\bar{p}(t_1)$	$\underline{p}(h_2)\bar{p}(t_1) + \bar{p}(h_2)\underline{p}(h_1)$
PB2	$\bar{p}(h_2)\bar{p}(t_2)\underline{p}(h_1)$	$\underline{p}(t_2)\bar{p}(t_2)\underline{p}(h_1)$	$\underline{p}(t_2)\underline{p}(h_2)\bar{p}(t_1)$	$\underline{p}(t_2)\bar{p}(t_2)\bar{p}(t_1)$	$\underline{p}(t_2)\bar{p}(t_1) + \bar{p}(t_2)\underline{p}(h_1)$
PB3	$\underline{p}(h_2)\bar{p}(h_2)\bar{p}(h_1)$	$\underline{p}(h_2)\underline{p}(t_2)\bar{p}(h_1)$	$\underline{p}(h_2)\bar{p}(h_2)\underline{p}(t_1)$	$\bar{p}(t_2)\bar{p}(h_2)\underline{p}(t_1)$	$\underline{p}(h_2)\bar{p}(h_1) + \bar{p}(h_2)\underline{p}(t_1)$
PB4	$\underline{p}(t_2)\underline{p}(h_2)\bar{p}(h_1)$	$\underline{p}(t_2)\bar{p}(t_2)\bar{p}(h_1)$	$\bar{p}(h_2)\bar{p}(t_2)\underline{p}(t_1)$	$\underline{p}(t_2)\bar{p}(t_2)\underline{p}(t_1)$	$\underline{p}(t_2)\bar{p}(h_1) + \bar{p}(t_2)\underline{p}(t_1)$

PA1, PA2, PA3, PA4

PA1=A4, PB1=B4, PA2=A3, PB2=B3

PB1, PB2, PB3, PB4

remove those for which $\Sigma = 0$

PS1, PS2, PS3, PS4

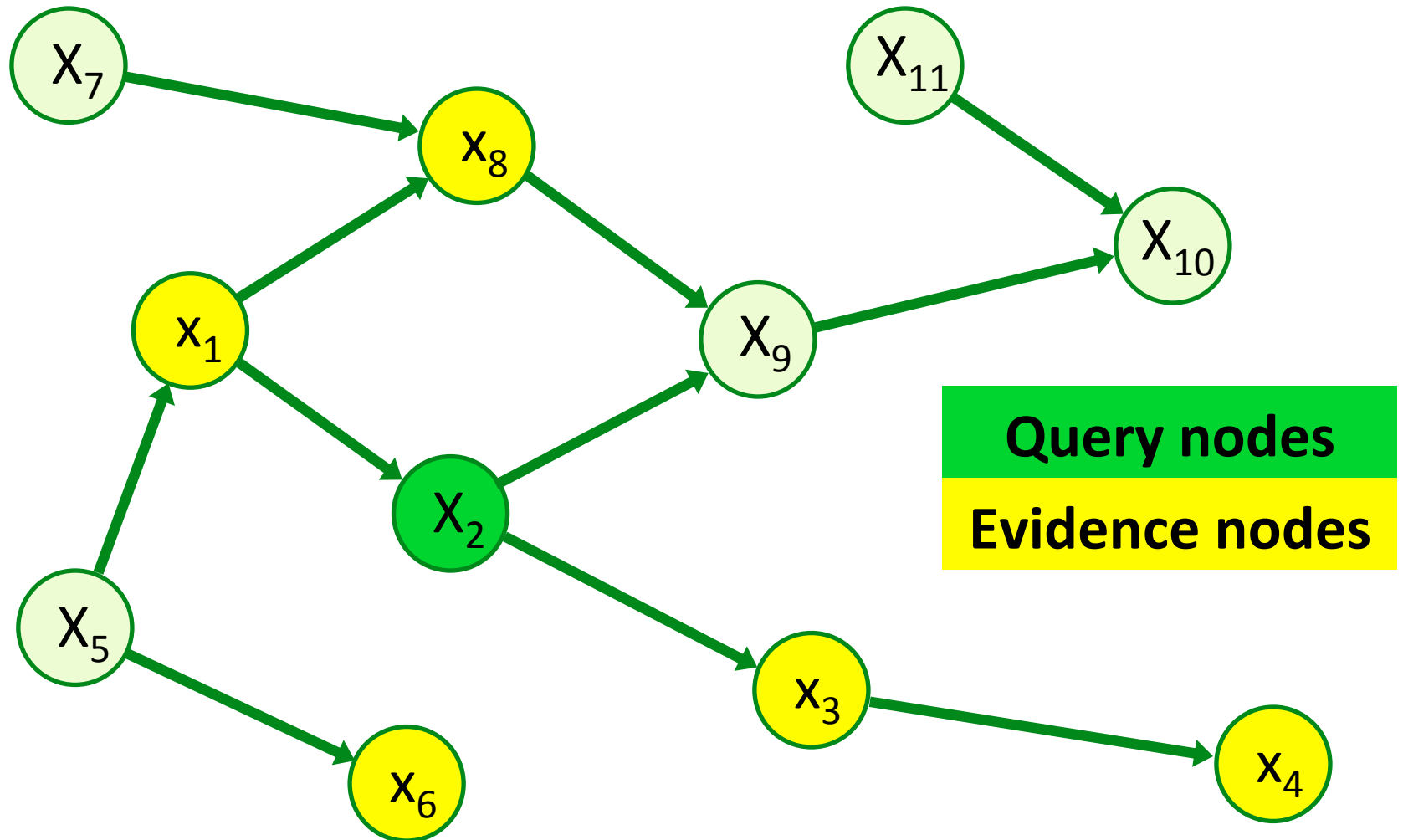
Independent natural extension for two binary variables

Analytical expressions for the ϵ

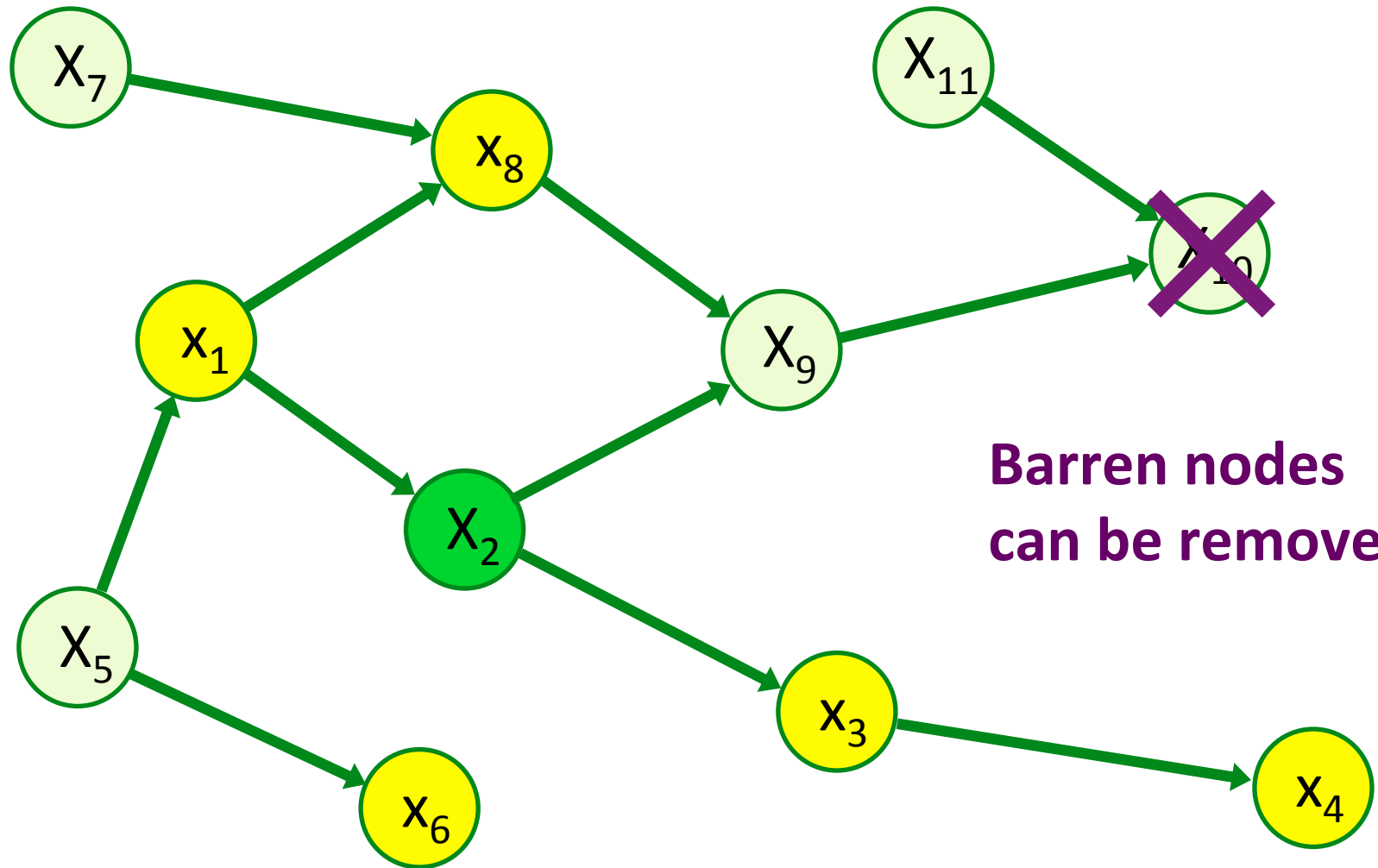
	$p(h_1, h_2) \Sigma$	$p(h_1, t_2) \Sigma$	$p(t_1, h_2) \Sigma$
PS1	$\underline{p}(h_1)\underline{p}(h_2)$	$\underline{p}(h_1)\bar{p}(t_2)$	$\bar{p}(t_1)\underline{p}(h_2)$
PS2	$\underline{p}(h_1)\bar{p}(h_2)$	$\underline{p}(h_1)\underline{p}(t_2)$	$\underline{p}(h_1)\underline{p}(h_2)$
PS3	$\bar{p}(h_1)\underline{p}(h_2)$	$\bar{p}(h_1)\bar{p}(t_2)$	$\bar{p}(h_1)\underline{p}(h_2)$
PS4	$\bar{p}(h_1)\bar{p}(h_2)$	$\bar{p}(h_1)\underline{p}(t_2)$	$\bar{p}(h_1)\underline{p}(h_2)$
PA1	$\underline{p}(h_1)\underline{p}(t_1)\bar{p}(h_2)\bar{p}(t_2)$	$\underline{p}(h_1)\underline{p}(t_1)\bar{p}(h_2)\bar{p}(t_2)$	$\underline{p}(h_1)\underline{p}(t_1)\bar{p}(h_2)\bar{p}(t_2)$
PA2	$\underline{p}(h_1)\underline{p}(t_1)\bar{p}(h_2)\underline{p}(t_2)$	$\underline{p}(h_1)\underline{p}(t_1)\bar{p}(h_2)\underline{p}(t_2)$	$\underline{p}(h_1)\underline{p}(t_1)\bar{p}(h_2)\underline{p}(t_2)$
PA3	$\bar{p}(h_1)\bar{p}(t_1)\underline{p}(h_2)\underline{p}(t_2)$	$\bar{p}(h_1)\bar{p}(t_1)\underline{p}(h_2)\underline{p}(t_2)$	$\bar{p}(h_1)\bar{p}(t_1)\underline{p}(h_2)\underline{p}(t_2)$
PA4	$\underline{p}(t_1)\underline{p}(h_2)\bar{p}(h_1)\bar{p}(t_2)$	$\bar{p}(h_1)\bar{p}(t_1)\underline{p}(t_2)$	$\underline{p}(t_1)\underline{p}(h_2)\bar{p}(h_1)\bar{p}(t_2)$
PB1	$\underline{p}(h_2)\bar{p}(h_2)\underline{p}(h_1)$	$\bar{p}(t_2)\bar{p}(h_2)\underline{p}(h_1)$	$\underline{p}(h_2)\bar{p}(h_2)\underline{p}(h_1)$
PB2	$\bar{p}(h_2)\bar{p}(t_2)\underline{p}(h_1)$	$\underline{p}(t_2)\bar{p}(t_2)\underline{p}(h_1)$	$\underline{p}(t_2)\underline{p}(h_2)\bar{p}(t_1)$
PB3	$\underline{p}(h_2)\bar{p}(h_2)\bar{p}(h_1)$	$\underline{p}(h_2)\underline{p}(t_2)\bar{p}(h_1)$	$\underline{p}(h_2)\bar{p}(h_2)\underline{p}(t_1)$
PB4	$\underline{p}(t_2)\underline{p}(h_2)\bar{p}(h_1)$	$\underline{p}(t_2)\bar{p}(t_2)\bar{p}(h_1)$	$\bar{p}(h_2)\bar{p}(t_2)\underline{p}(t_1)$

$\underline{p}(h_1)\underline{p}(t_1)\bar{p}(h_2)\bar{p}(t_2)$ → PA1, PA2, PA3, PA4
 $\underline{p}(h_1)\underline{p}(t_1)\bar{p}(h_2)\underline{p}(t_2)$ → PA1=A4, PA2=B4
 $\bar{p}(h_1)\bar{p}(t_1)\underline{p}(h_2)\underline{p}(t_2)$ → PB1, PB2, PB3, PB4

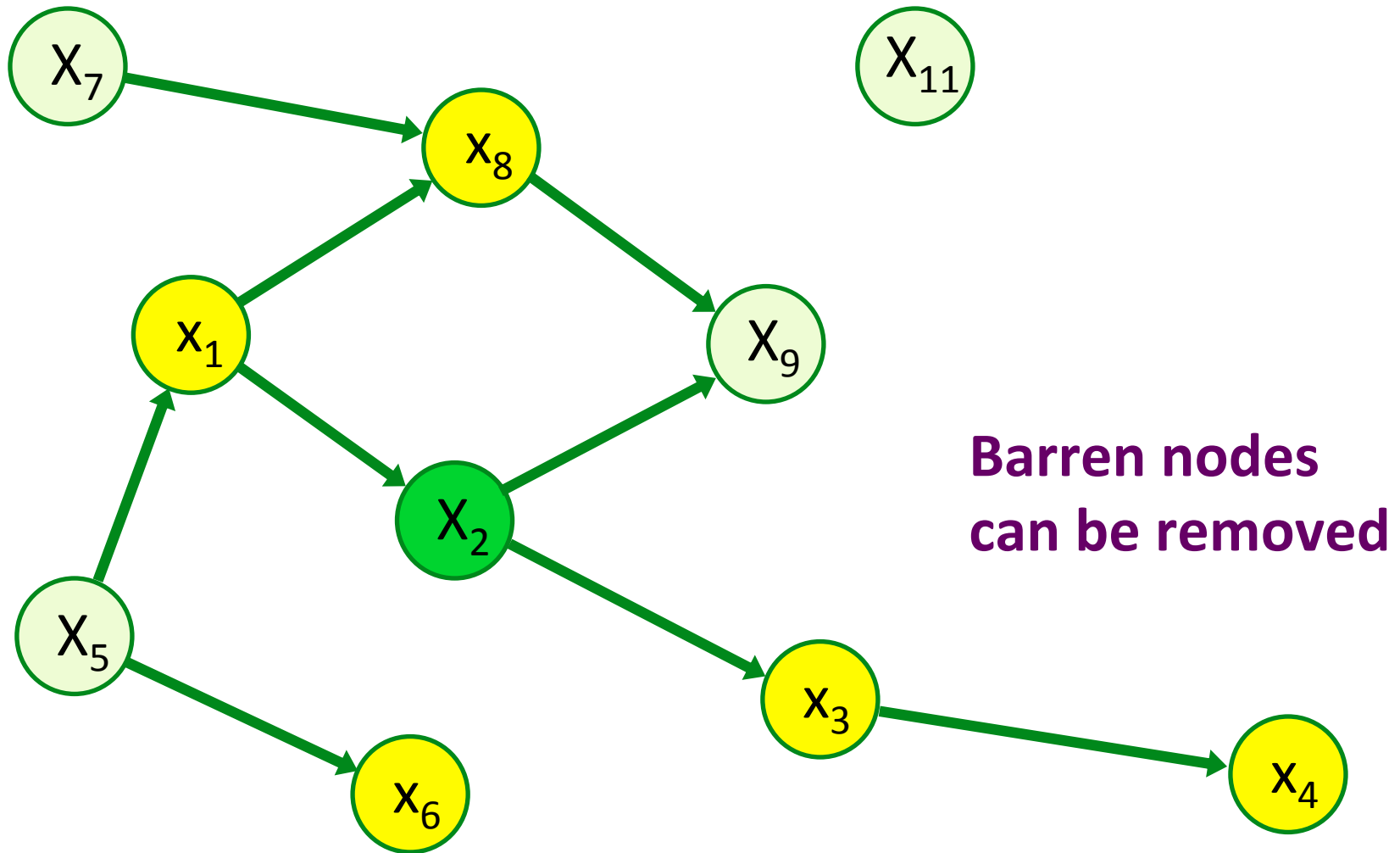
Bayesian networks: useful properties



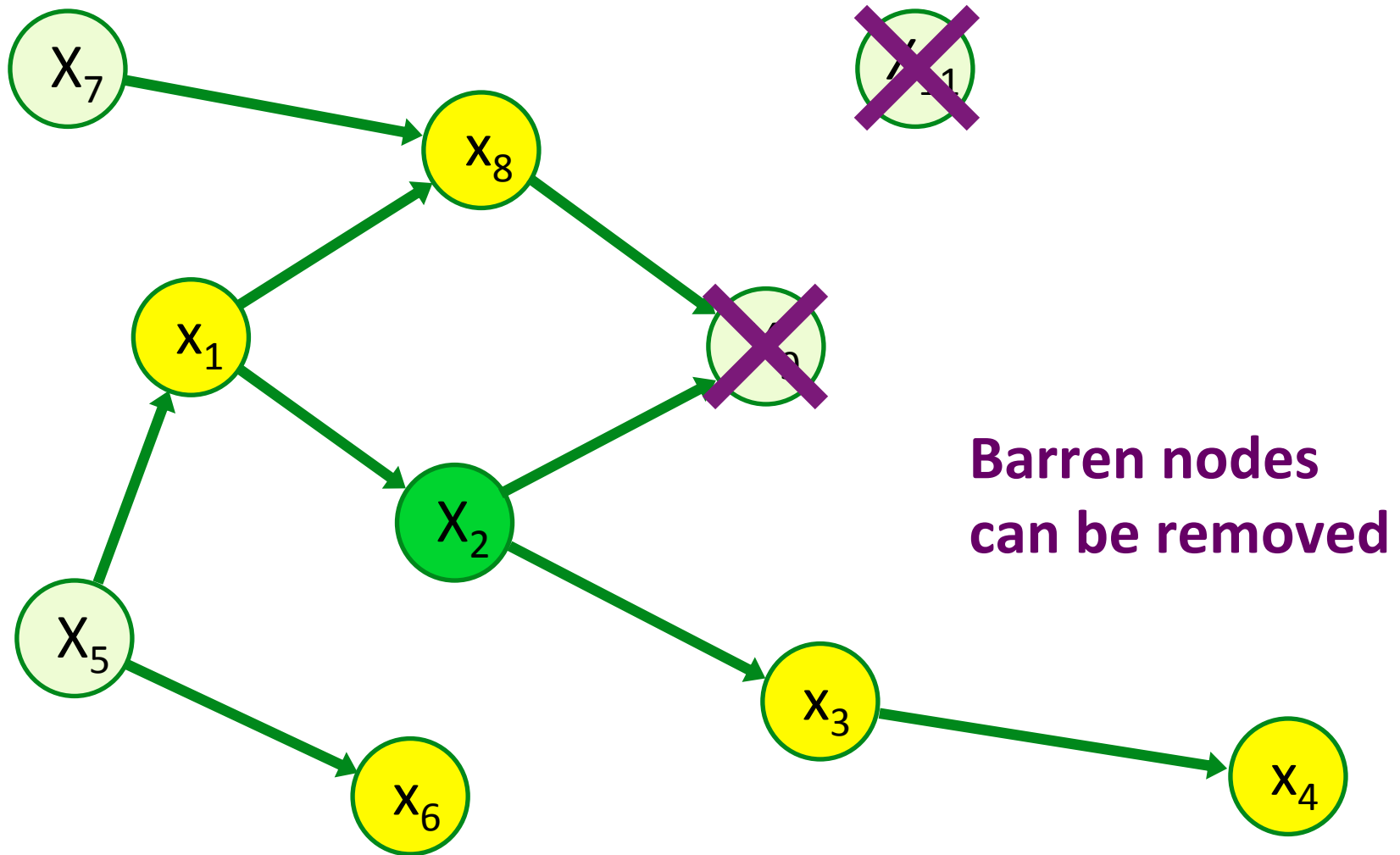
Bayesian networks: useful properties



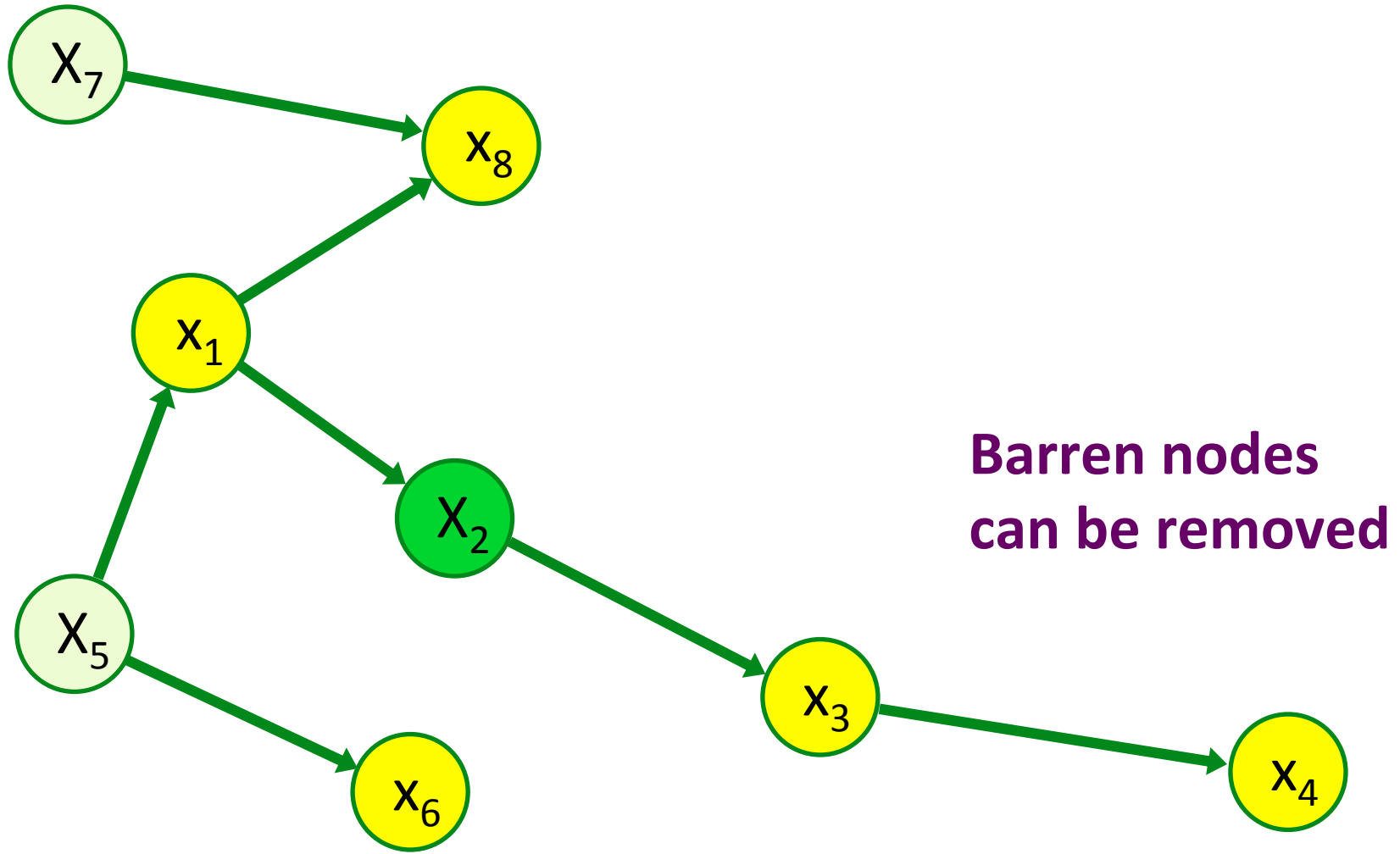
Bayesian networks: useful properties



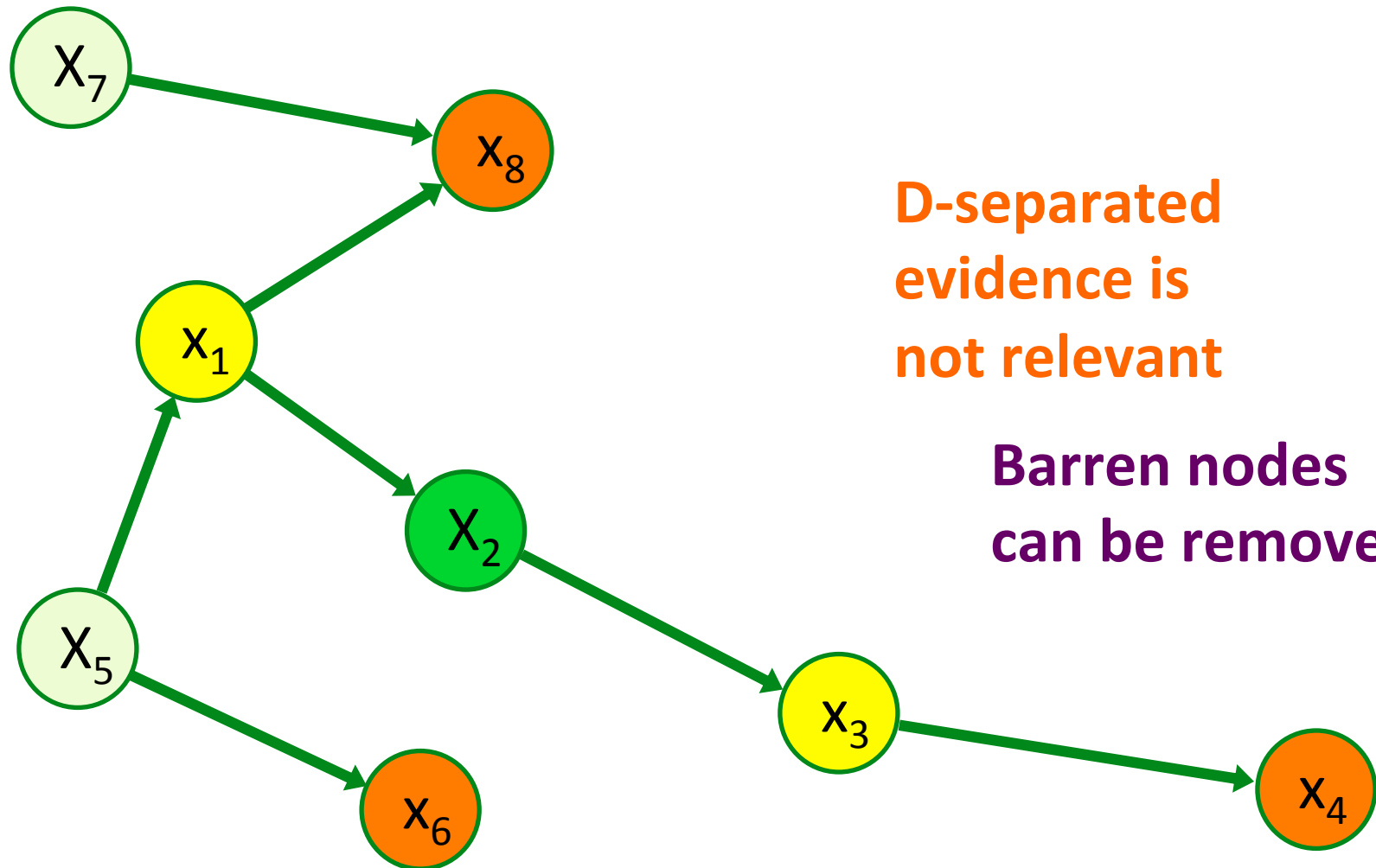
Bayesian networks: useful properties



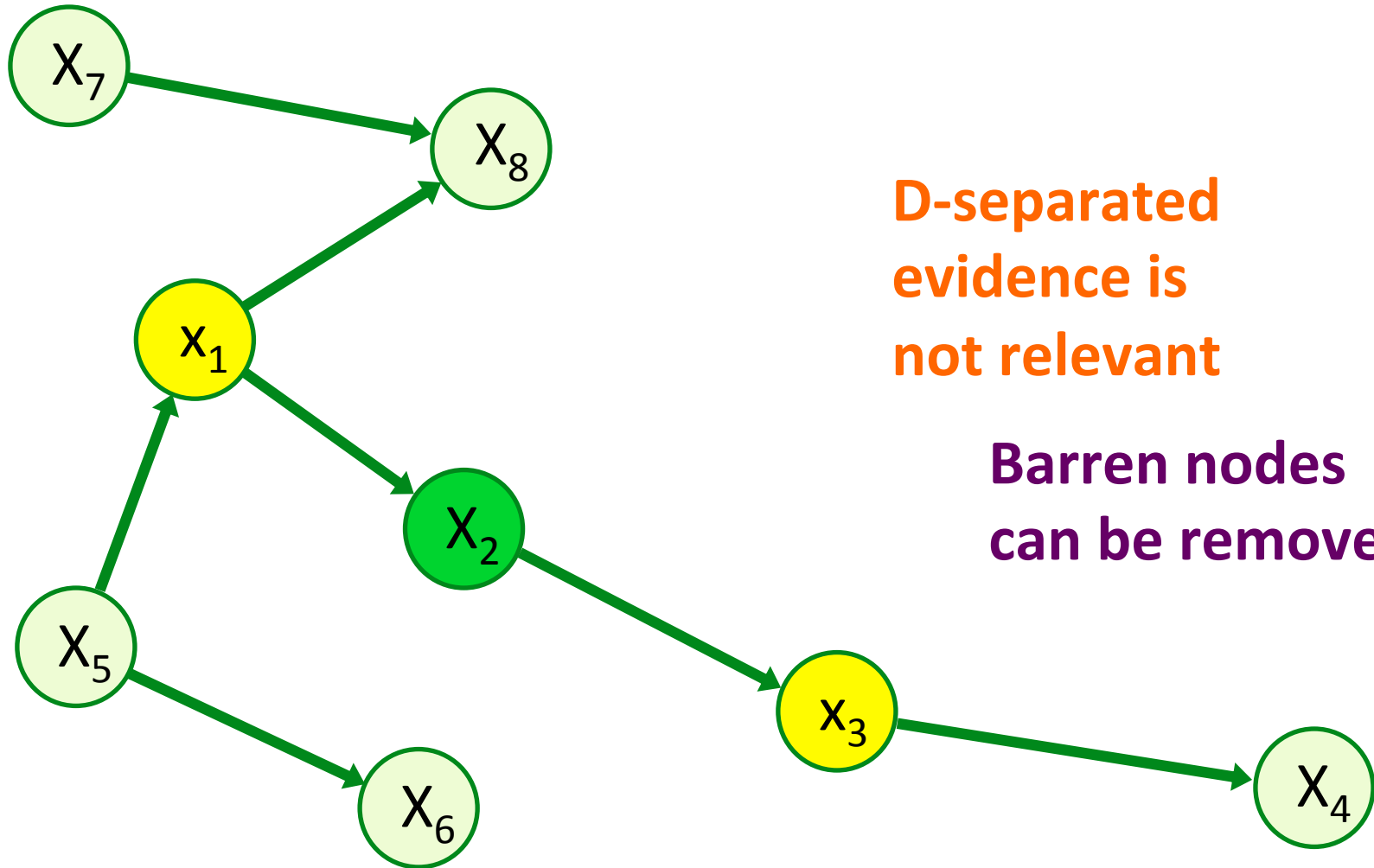
Bayesian networks: useful properties



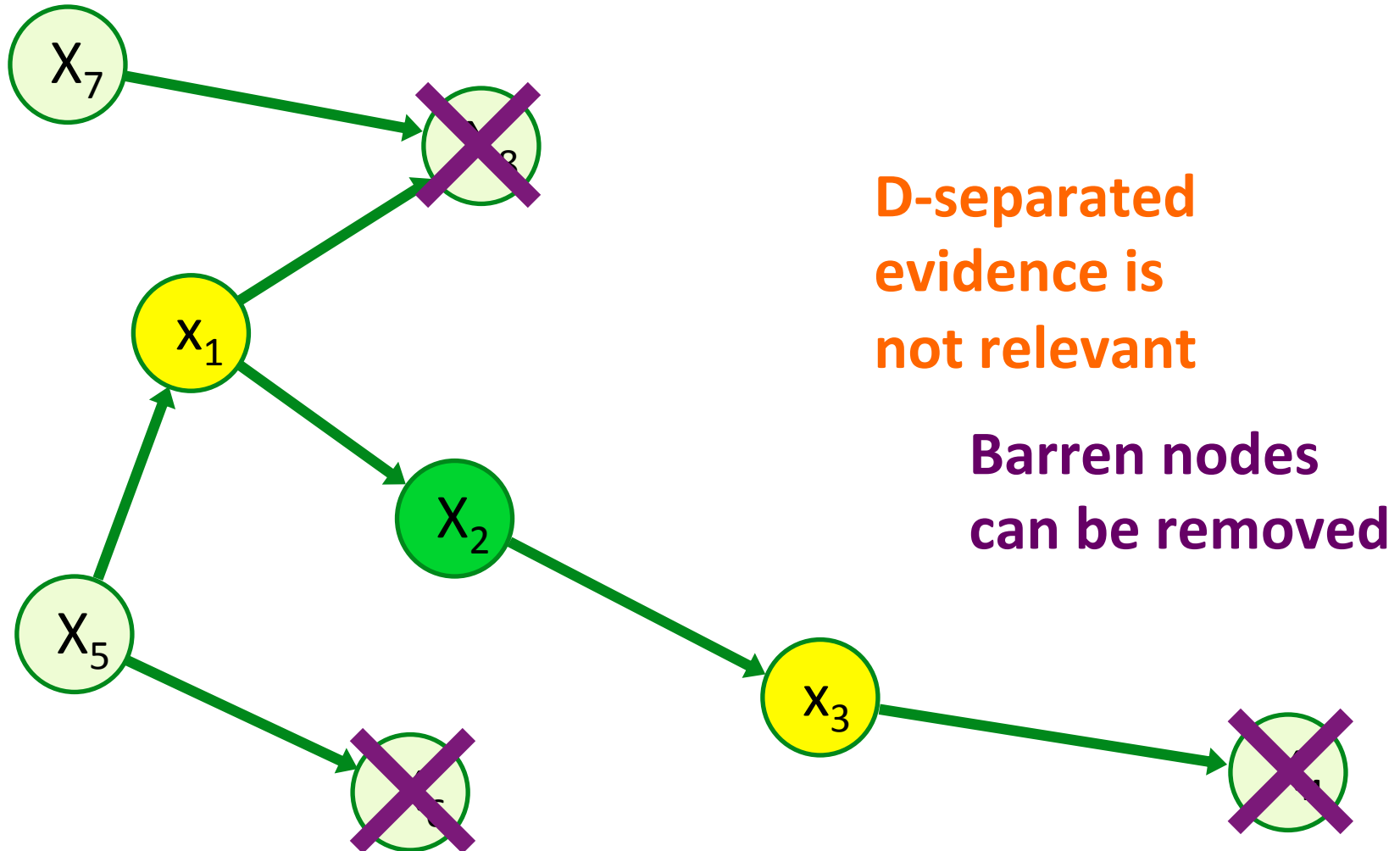
Bayesian networks: useful properties



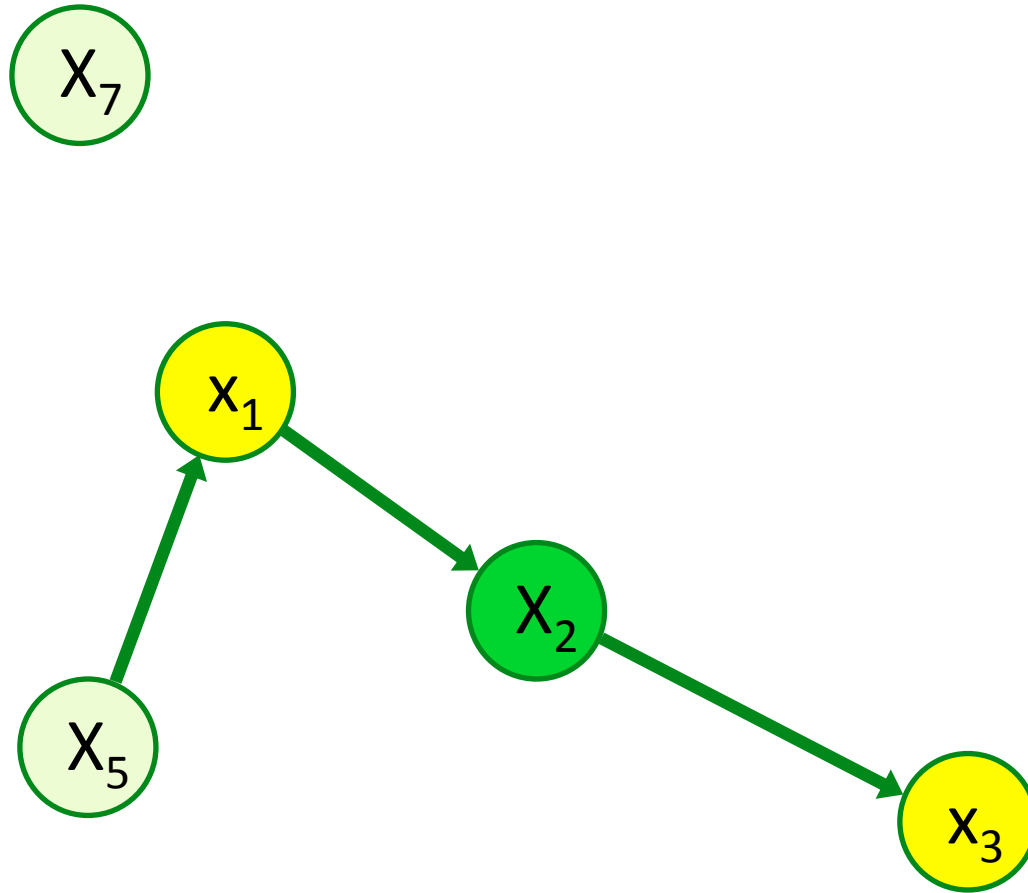
Bayesian networks: useful properties



Bayesian networks: useful properties



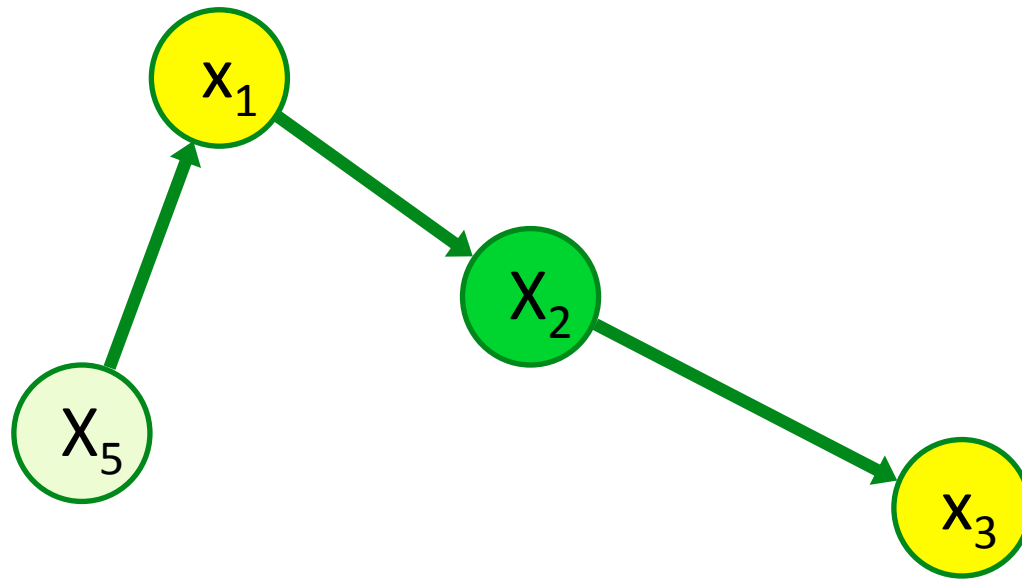
Bayesian networks: useful properties



**D-separated
evidence is
not relevant**

**Barren nodes
can be removed**

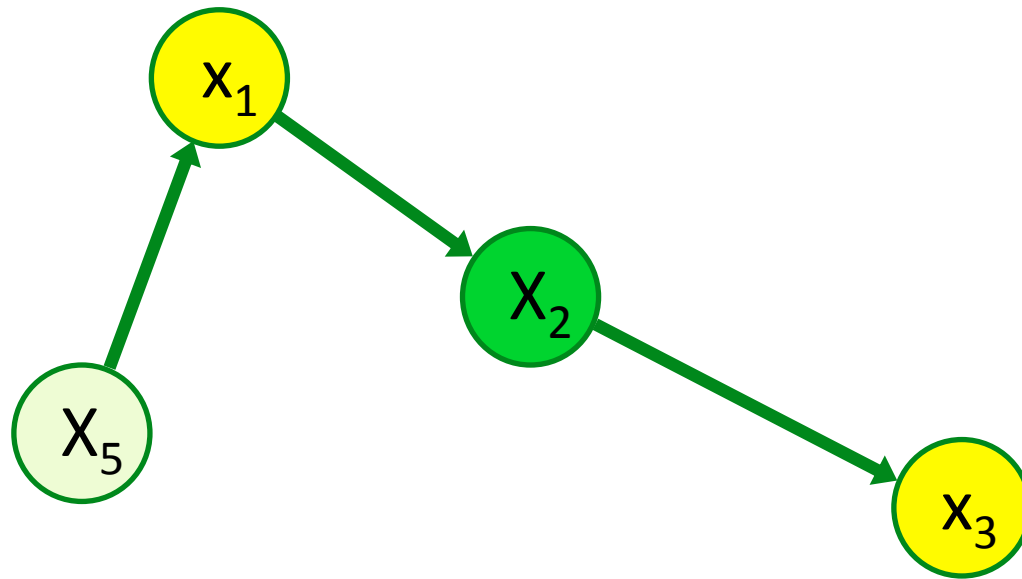
Bayesian networks: useful properties



**D-separated
evidence is
not relevant**

**Barren nodes
can be removed**

Bayesian networks: useful properties



**D-separated
evidence is
not relevant**

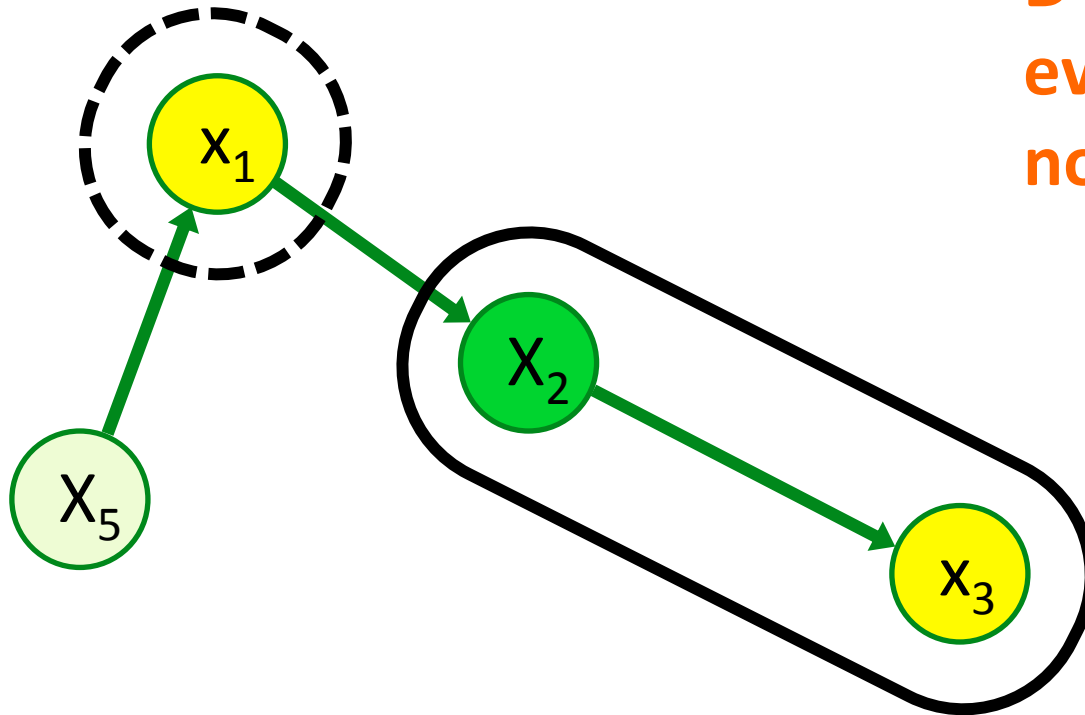
**Barren nodes
can be removed**

Bayesian networks: useful properties

Conditional
marginalisation
properties

D-separated
evidence is
not relevant

Barren nodes
can be removed

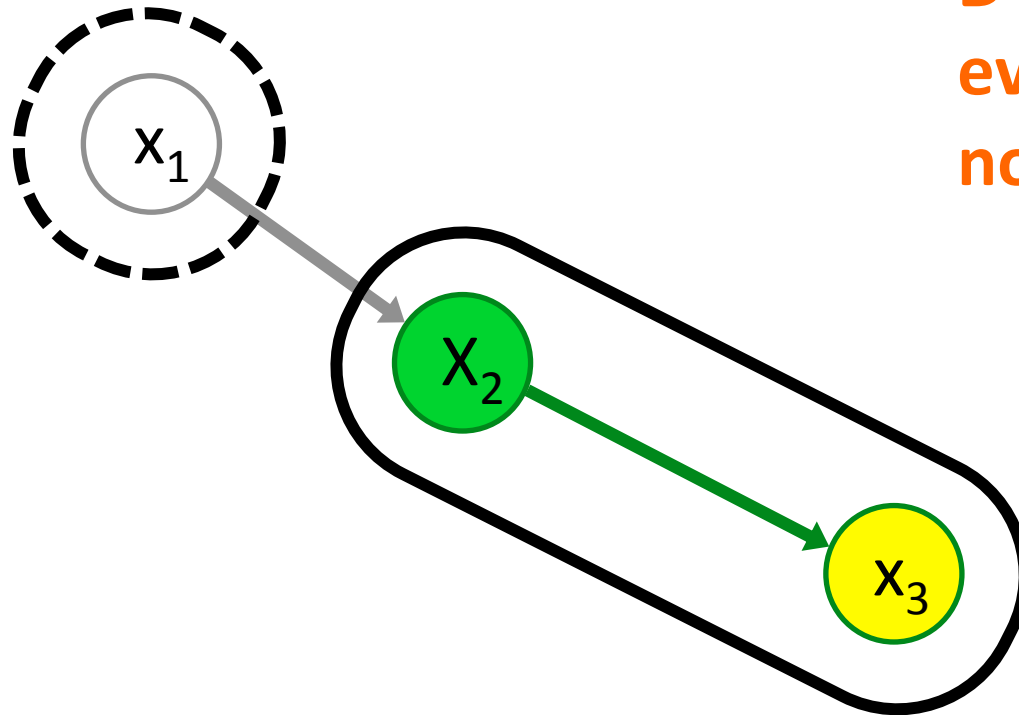


Bayesian networks: useful properties

Conditional
marginalisation
properties

D-separated
evidence is
not relevant

Barren nodes
can be removed

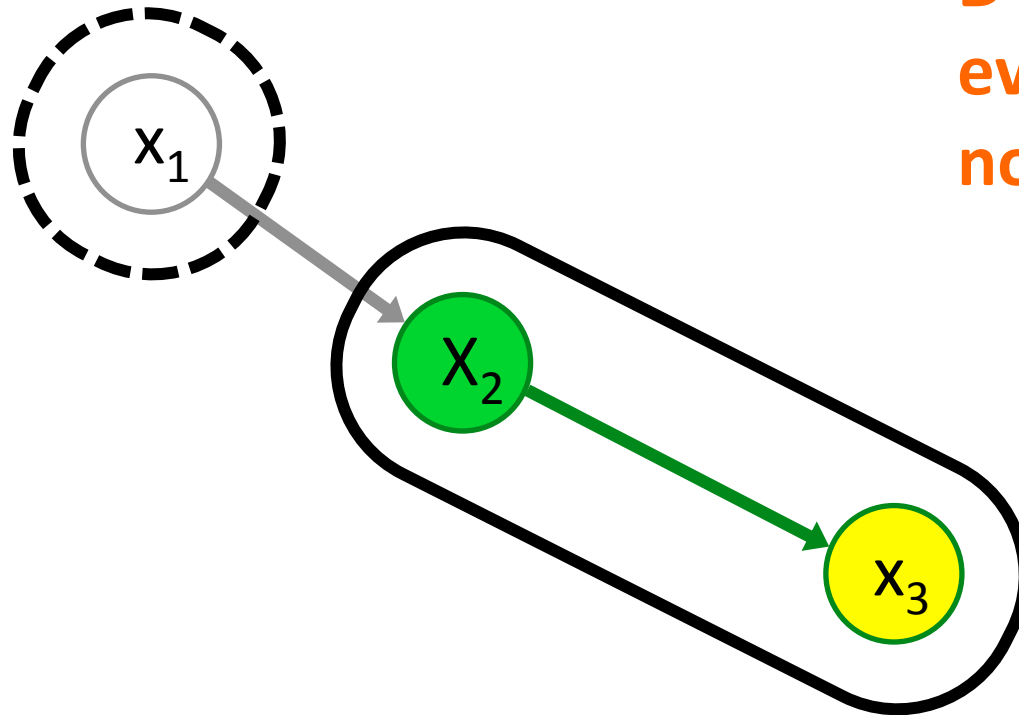


Credal networks: useful properties ?

Conditional
marginalisation
properties

D-separated
evidence is
not relevant

Barren nodes
can be removed



Credal networks: useful properties ?



**Conditional
marginalisation
properties**



**D-separated
evidence is
not relevant**

**Barren nodes
can be removed**



**Credal networks under
epistemic irrelevance**

Credal networks: useful properties ?

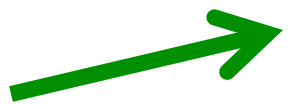
? Conditional marginalisation properties

D-separated evidence is not relevant

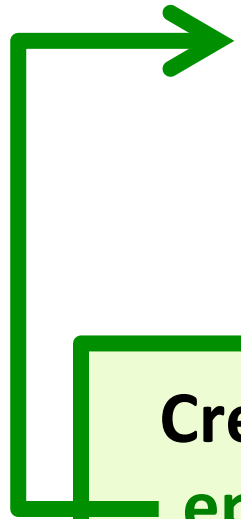
Barren nodes can be removed



~~graphoid axioms~~



Credal networks under epistemic irrelevance



Credal networks: a joint model

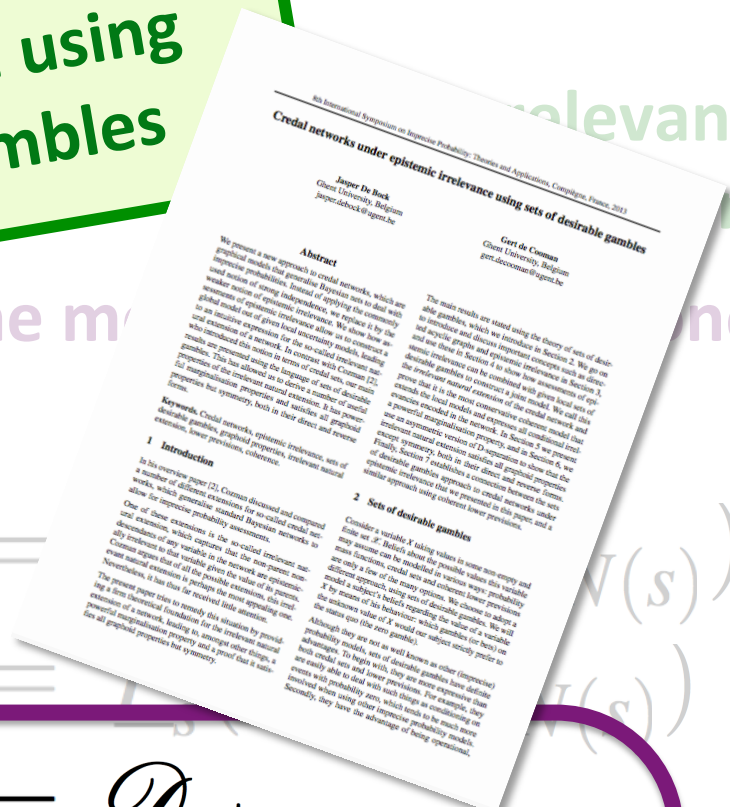
Credal networks under epistemic irrelevance using sets of desirable gambles



$$K^{irr}(X_G)$$

$$P_G^{irr}$$

$$\mathcal{D}_G^{irr}$$



The m

ne!

$$K(X_S | x_{P(S)}) = \dots$$

$$P_S(\cdot | x_{P(S)}) = \dots$$

$$\mathcal{D}_S | x_{P(S)} = \mathcal{D}_S | x_{P(S) \cup N(S)}$$



Credal networks under epistemic irrelevance using sets of desirable gambles

Jasper De Bock and Gert de Cooman
 {Jasper.DeBock, Gert.deCooman}@UGent.be
 SYSTeMS, Ghent University, Belgium

Abstract of the paper

We present a new approach to credal networks, which are graphical models that generalise Bayesian nets to deal with imprecise probabilities. Instead of applying the commonly used notion of strong independence, we replace it by the weaker notion of epistemic irrelevance (II). We show how assessments of epistemic irrelevance allow us to construct a global model out of given local uncertainty models (I), leading to an intuitive expression for the so-called irrelevant natural extension (III) of a network. In contrast with Cozman (2000) who introduced this notion in terms of credal sets, our main results are presented using the language of sets of desirable gambles (SDG). This has allowed us to derive a number of useful properties of the irrelevant natural extension. It has powerful marginalisation properties (IV) and satisfies all graphoid properties but symmetry, both in their direct and reverse forms (V & VI).

Sets of desirable gambles (SDG)

We will model a subject's beliefs about the value that a variable X assumes in some set \mathcal{X} , by means of his behaviour, which gambles (real-valued maps) f on \mathcal{X} does he strictly prefer to the status quo. This results in a set of desirable gambles $\mathcal{D} \subseteq \mathcal{F}(\mathcal{X})$, where $\mathcal{F}(\mathcal{X})$ is the set of all gambles on \mathcal{X} . \mathcal{D} is called coherent if it satisfies the rationality requirements D1–D4 for all $f, f_1, f_2 \in \mathcal{F}(\mathcal{X})$ and all real $\lambda > 0$.

- D1 $f \leq 0 \Rightarrow f \notin \mathcal{D}$
- D2 $f > 0 \Rightarrow f \in \mathcal{D}$
- D3 $f \in \mathcal{D} \Rightarrow \lambda f \in \mathcal{D}$
- D4 $f_1, f_2 \in \mathcal{D} \Rightarrow f_1 + f_2 \in \mathcal{D}$

Although they are not as well known as other (imprecise) probability models, sets of desirable gambles have definite advantages. To give a few examples: they are operational, are easily able to deal with conditioning on events with probability zero, allow for intuitive geometrically flavoured proofs and are more expressive than both credal sets and lower previsions (see our papers or the second poster for credal networks under epistemic irrelevance that use these alternative models).

Local uncertainty models (I)

With every node x of a finite directed acyclic graph (DAG), we associate a variable X , taking values in some finite, non-empty set \mathcal{X} . The set of all nodes is denoted by G . For every subset $S \subseteq G$, the joint variable X_S takes values in $\mathcal{X}_S := \times_{x \in S} \mathcal{X}_x$. For every $x \in G$, we denote by $P(x)$ the set consisting of the parent nodes of x . Similar to what is done in classical Bayesian networks, we attach local uncertainty models to the nodes of the network, conditional on the value of their parents. For all $x \in G$ and every instantiation $x_{P(x)} \in \mathcal{X}_{P(x)}$, we require a coherent set $\mathcal{D}_{x|P(x)}$ of desirable gambles (SDG) on \mathcal{X}_x .

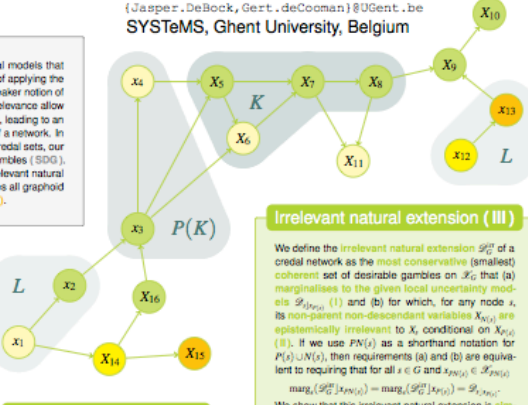
AD-separation (V)

Consider any path x_1, \dots, x_n in G , with $n \geq 1$. We say that this path is blocked by a set of nodes $C \subseteq G$ whenever at least one of the following four conditions holds:

- B1 $x_1 \in C$;
- B2 there is some $1 < i < n$ such that $x_i \rightarrow x_{i+1}$ and $x_i \in C$;
- B3 there is some $1 < i < n$ such that $x_{i-1} \rightarrow x_i \leftarrow x_{i+1}$, $x_i \in C$ and $D(x_i) \cap C = \emptyset$;
- B4 $x_n \in C$.

Now consider (not necessarily disjoint) subsets I, O and C of G . We say that O is AD-separated from I by C , denoted as $AD(I, O | C)$, if every path $I \rightarrow x_1, \dots, x_n \rightarrow O$, $n \geq 1$, from a node $I \in I$ to a node $O \in O$, is blocked by C .

This asymmetric version of D-separation is similar to, yet different from both Moral's (2005) version of AD-separation and the notion of L-separation, as introduced by Vantaggi (2002). Our reason for not using one of these existing concepts is that our version of AD-separation has stronger properties: it satisfies all graphoid properties except symmetry: it satisfies redundancy, decomposition, weak union, contraction and intersection both in their direct and reverse forms.



Irrelevant natural extension (III)

We define the irrelevant natural extension $\mathcal{D}_G^{\text{irr}}$ of a credal network as the most conservative (smallest) coherent set of desirable gambles on \mathcal{X}_G that (a) marginalises to the given local uncertainty models $\mathcal{D}_{x|P(x)}$ (I) and (b) for which, for any node x , its non-parent non-descendant variables $X_{N(x)}$ are epistemically irrelevant to X_x conditional on $X_{P(x)}$ (II). If we use $P_N(x)$ as a shorthand notation for $P(x) \cup N(x)$, then requirements (a) and (b) are equivalent to requiring that for all $s \in G$ and $x_{P_N(s)} \in \mathcal{X}_{P_N(s)}$

$$\text{marg}_G(\mathcal{D}_{x|P_N(s)}^{\text{irr}}) = \text{marg}_G(\mathcal{D}_{x|P_N(s)}^{\text{irr}}) = \mathcal{D}_{x|P_N(s)}$$

We show that this irrelevant natural extension is sim-

operator generates the set of all finite positive linear combinations of elements in its argument set

$$\mathcal{D}_G^{\text{irr}} := \{ \sum_{i=1}^n \lambda_i f_i : f_i \in \mathcal{D}_{x|P_N(s)}^{\text{irr}}, \lambda_i \geq 0 \}$$

Consider a global set of desirable gambles \mathcal{D}_G (SDG) on \mathcal{X}_G and disjoint subsets S and K of G . Then the marginal model for X_S conditional on the information that X_K assumes a value $x_K \in \mathcal{X}_K$, is given by $\text{marg}_K(\mathcal{D}_G | x_K) = \{ f \in \mathcal{F}(\mathcal{X}_S) : \exists_{x_K} f \in \mathcal{D}_G \}$. Consider now three subsets $C, I, O \subseteq G$, with $I \cap C$ and $O \cap C$ disjoint. We say that X_I is epistemically irrelevant to X_O conditional on X_C , denoted as $IR(I, O | C)$, if and only if for all $x_{C \cup I} \in \mathcal{X}_{C \cup I}$ we have

$$\text{marg}_{O \cup C}(\mathcal{D}_G | x_{C \cup I}) = \text{marg}_{O \cup C}(\mathcal{D}_G | x_C)$$

Our paper also considers epistemic subset-irrelevance, which although interesting, is not discussed on this poster.

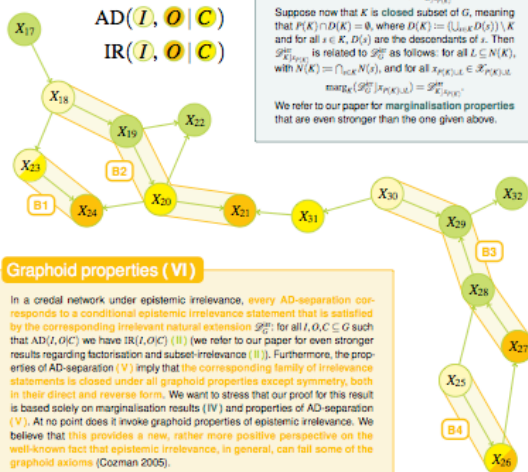
Marginalisation properties (IV)

For any $K \subseteq G$, we construct a sub-DAG of the original DAG by eliminating the nodes $s \in G \setminus K$ and their associated edges. The parents of a node $x \in K$, with respect to this sub-DAG, are denoted by $P_K(x) := P(x) \cap K$. We derive local models $\mathcal{D}_{x|P_K(x)}$ for this sub-DAG from the original local models $\mathcal{D}_{x|P(x)}$ by fixing $x_{P(x) \setminus K}$. We do this consistently for all $x \in K$ at once by fixing $x_{P(x) \setminus K}$ where $P(K) := (\cup_{x \in K} P(x)) \setminus K$. For any $x_{P(K)} \in \mathcal{X}_{P(K)}$, we use the resulting local models to construct an irrelevant natural extension of the sub-DAG and denote it by $\mathcal{D}_{x|P_K(x)}^{\text{irr}}$.

Suppose now that K is closed subset of G , meaning that $P(K) \cap D(K) = \emptyset$, where $D(K) := (\cup_{x \in K} D(x)) \setminus K$ and for all $x \in K$, $D(x)$ are the descendants of x . Then $\mathcal{D}_{x|P_K(x)}^{\text{irr}}$ is related to $\mathcal{D}_G^{\text{irr}}$ as follows: for all $I \subseteq N(K)$, with $N(K) = \cap_{x \in K} N(x)$, and for all $x_{P(K) \cup I} \in \mathcal{X}_{P(K) \cup I}$

$$\text{marg}_I(\mathcal{D}_{x|P_K(x)}^{\text{irr}} | x_{P(K) \cup I}) = \mathcal{D}_{x|P_K(x)}^{\text{irr}}$$

We refer to our paper for marginalisation properties that are even stronger than the one given above.



Graphoid properties (VI)

In a credal network under epistemic irrelevance, every AD-separation corresponds to a conditional epistemic irrelevance statement that is satisfied by the corresponding irrelevant natural extension $\mathcal{D}_G^{\text{irr}}$: for all $I, O, C \subseteq G$ such that $AD(I, O | C)$ we have $IR(I, O | C)$ (II) (we refer to our paper for even stronger results regarding factorisation and subset-irrelevance (II)). Furthermore, the properties of AD-separation (V) imply that the corresponding family of irrelevance statements is closed under all graphoid properties except symmetry, both in their direct and reverse form. We want to stress that our proof for this result is based solely on marginalisation results (IV) and properties of AD-separation (V). At no point does it invoke graphoid properties of epistemic irrelevance. We believe that this provides a new, rather more positive perspective on the well-known fact that epistemic irrelevance, in general, can fall some of the graphoid axioms (Cozman 2005).

Advantages of sets of desirable gambles

Credal networks under epistemic irrelevance using sets of desirable gambles

Jasper De Bock and Gert de Cooman
{Jasper.DeBock, Gert.deCooman}@UGent.be
SYSTeMS, Ghent University, Belgium

Abstract of the paper

We present a new approach to credal networks, which are graphical models that generalise Bayesian nets to deal with imprecise probabilities. Instead of applying the commonly used notion of strong independence, we replace it by the weaker notion of epistemic irrelevance (II). We show how assessments of epistemic irrelevance allow us to construct a global model out of given local uncertainty models (I), leading to an intuitive expression for the so-called irrelevant natural extension (III) of a network. In contrast with Cozman (2000) who introduced this notion in terms of credal sets, our main results are presented using the language of sets of desirable gambles (SDG). This has allowed us to derive a number of useful properties of the irrelevant natural extension. It has powerful marginalisation properties (IV) and satisfies all graphoid properties but symmetry, both in their direct and reverse forms (V & VI).

Sets of desirable gambles (SDG)

We will model a subject's beliefs about the value that a variable X , assumes in some set \mathcal{X} , by means of his behaviour, which gambles (real-valued maps) f on \mathcal{X} does he strictly prefer to the status quo. This results in a set of desirable gambles $\mathcal{D} \subseteq \mathcal{F}(\mathcal{X})$, where $\mathcal{F}(\mathcal{X})$ is the set of all gambles on \mathcal{X} . \mathcal{D} is called coherent if it satisfies the rationality requirements D1–D4 for all $f, f_1, f_2 \in \mathcal{F}(\mathcal{X})$ and all real $\lambda > 0$.

- D1 $f \leq 0 \Rightarrow f \in \mathcal{D}$
- D2 $f > 0 \Rightarrow f \in \mathcal{D}$
- D3 $f \in \mathcal{D} \Rightarrow \lambda f \in \mathcal{D}$
- D4 $f_1, f_2 \in \mathcal{D} \Rightarrow f_1 + f_2 \in \mathcal{D}$

Although they are not as well known as other (imprecise) probability models, sets of desirable gambles have definite advantages. To give a few examples: they are operational, are easily able to deal with conditioning on events with probability zero, allow for intuitive geometrically framed proofs and are more expressive than both credal sets and lower previsions (see our papers or the second poster for credal networks under epistemic irrelevance that use these alternative models).

Local uncertainty models (I)

With every node x of a finite directed acyclic graph (DAG), we associate a variable X_x taking values in some finite, non-empty set \mathcal{X}_x . The set of all nodes is denoted by G . For every subset $S \subseteq G$, the joint variable X_S takes values in $\mathcal{X}_S := \times_{x \in S} \mathcal{X}_x$. For every $x \in G$, we denote by $P(x)$ the set consisting of the parent nodes of x . Similar to what is done in classical Bayesian networks, we attach local uncertainty models to the nodes of the network, conditional on the value of their parents. For all $x \in G$ and every instantiation $x_{P(x)} \in \mathcal{X}_{P(x)}$, we require a coherent set $\mathcal{D}_{x|x_{P(x)}}$ of desirable gambles (SDG) on \mathcal{X}_x .

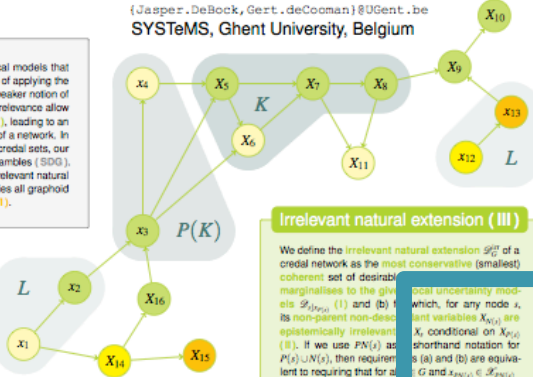
AD-separation (V)

Consider any path x_1, \dots, x_n in G , with $n \geq 1$. We say that this path is blocked by a set of nodes $C \subseteq G$ whenever at least one of the following four conditions holds:

- B1 $x_1 \in C$;
- B2 there is some $1 < i < n$ such that $x_i \rightarrow x_{i+1}$ and $x_i \in C$;
- B3 there is some $1 < i < n$ such that $x_{i-1} \rightarrow x_i \leftarrow x_{i+1}$, $x_i \in C$ and $D(x_i) \cap C = \emptyset$;
- B4 $x_n \in C$.

Now consider (not necessarily disjoint) subsets I, O and C of G . We say that O is AD-separated from I by C , denoted as $AD(I, O | C)$, if every path $I \rightarrow x_1, \dots, x_n \rightarrow O$, $n \geq 1$, from a node $I \in I$ to a node $O \in O$, is blocked by C .

This asymmetrical version of D-separation is similar to, yet different from both Moral's (2005) version of AD-separation and the notion of L-separation, as introduced by Vantaggi (2002). Our reason for not using one of these existing concepts is that our version of AD-separation has stronger properties: it satisfies all graphoid properties except symmetry: it satisfies redundancy, decomposition, weak union, contraction and intersection both in their direct and reverse forms.

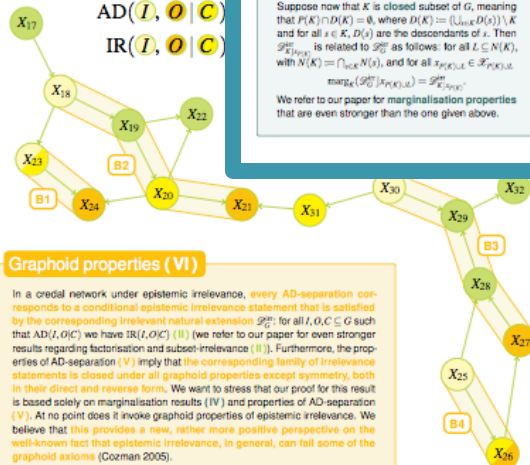


Epistemic irrelevance (II)

Consider a global set of desirable gambles \mathcal{D}_G (SDG) on \mathcal{X}_G and disjoint subsets S and K of G . Then the marginal model for X_S , conditional on the information that X_K assumes a value $x_K \in \mathcal{X}_K$, is given by $\text{marg}_K(\mathcal{D}_G | x_K) = \{f \in \mathcal{F}(\mathcal{X}_S) : \exists_{x_K} f \in \mathcal{D}_G\}$. Consider now three subsets $C, I, O \subseteq G$, with $I \cap C$ and $O \cap C$ disjoint. We say that X_C is epistemically irrelevant to X_O conditional on X_I , denoted as $IR(I, O | C)$, if and only if for all $x_{C \cup I} \in \mathcal{X}_{C \cup I}$ we have $\text{marg}_{O \cap C}(\mathcal{D}_G | x_{C \cup I}) = \text{marg}_{O \cap C}(\mathcal{D}_G | x_I)$.

Our paper also considers epistemically AD-separated subsets, which, although interesting, is not discussed on this poster.

$AD(I, O | C)$
 $IR(I, O | C)$



Graphoid properties (VI)

In a credal network under epistemic irrelevance, every AD-separation corresponds to a conditional epistemic irrelevance statement that is satisfied by the corresponding irrelevant natural extension \mathcal{D}_G^{irr} . For all $I, O, C \subseteq G$ such that $AD(I, O | C)$ we have $IR(I, O | C)$ (we refer to our paper for even stronger results regarding factorisation and subset irrelevance (II)). Furthermore, the properties of AD-separation (V) imply that the corresponding family of irrelevance statements is closed under all graphoid properties except symmetry, both in their direct and reverse form. We want to stress that our proof for this result is based solely on marginalisation results (IV) and properties of AD-separation (V). At no point does it invoke graphoid properties of epistemic irrelevance. We believe that this provides a new, rather more positive perspective on the well-known fact that epistemic irrelevance, in general, can fall some of the graphoid axioms (Cozman 2005).

Irrelevant natural extension (III)

We define the irrelevant natural extension \mathcal{D}_G^{irr} of a credal network as the most conservative (smallest) coherent set of desirable gambles to the given local uncertainty models (I) and (II) which, for any node x , its non-parent non-descendants $X_{N(x)}$ are epistemically irrelevant (II). If we use $FN(x)$ as shorthand notation for $s(x) \cup N(x)$, then requirement (II) and (b) are equivalent to requiring that for all $x \in G$ and $x_{FN(x)} \in \mathcal{X}_{FN(x)}$, $\text{marg}_x(\mathcal{D}_G^{irr} | x_{FN(x)}) = \text{marg}_x(\mathcal{D}_{x|x_{FN(x)}})$. We show that this irrelevant natural extension is simple to construct: $\mathcal{D}_G^{irr} := \text{operator generates the combinations of elements } \mathcal{D}_G^{irr} := \{ \sum_{i \in I} \lambda_i f_i : f_i \in \mathcal{D}_{x_i | x_{N(x_i)}}, \lambda_i > 0 \}$.

Marginalisation properties (IV)

For any $K \subseteq G$, we construct a sub-DAG of the original DAG by eliminating the nodes $s \in G \setminus K$ and their associated edges. The parents of a node $x \in K$, with respect to this sub-DAG, are denoted by $P_K(x) := P(x) \cap K$. We derive local models $\mathcal{D}_{x|x_{P_K(x)}}$ for this sub-DAG from the original local models $\mathcal{D}_{x|x_{P(x)}}$ by fixing $x_{P(x) \setminus K}$. We do this consistently for all $x \in K$ at once by fixing $x_{P(x) \setminus K}$ where $P(x) := (\cup_{x \in K} P(x)) \cap K$. For any $x \in K$, we use the resulting local models to construct an irrelevant natural extension of the sub-DAG and denote it by $\mathcal{D}_{x|x_{P_K(x)}}$. Suppose now that K is closed subset of G , meaning that $P(K) \cap D(K) = \emptyset$, where $D(K) := (\cup_{x \in K} D(x)) \cap K$ and for all $x \in K$, $D(x)$ are the descendants of x . Then $\mathcal{D}_{x|x_{P_K(x)}}$ is related to \mathcal{D}_G^{irr} as follows: for all $I \subseteq N(K)$, with $N(K) := \cap_{x \in K} N(x)$, and for all $x_{P_K(I, K)} \in \mathcal{X}_{P_K(I, K)}$, $\text{marg}_I(\mathcal{D}_G^{irr} | x_{P_K(I, K)}) = \mathcal{D}_{x|x_{P_K(I, K)}}$. We refer to our paper for marginalisation properties that are even stronger than the one given above.

- Advantages of sets of desirable gambles
- Marginalisation properties

Credal networks under epistemic irrelevance using sets of desirable gambles

Jasper De Bock and Gert de Cooman
 {Jasper.DeBock, Gert.deCooman}@UGent.be
 SYSTeMS, Ghent University, Belgium

Abstract of the paper

We present a new approach to credal networks, which are graphical models that generalise Bayesian nets to deal with imprecise probabilities. Instead of applying the commonly used notion of strong independence, we replace it by the weaker notion of epistemic irrelevance (II). We show how assessments of epistemic irrelevance allow us to construct a global model out of given local uncertainty models (I), leading to an intuitive expression for the so-called irrelevant natural extension (III) of a network. In contrast with Cozman (2000) who introduced this notion in terms of credal sets, our main results are presented using the language of sets of desirable gambles (SDG). This has allowed us to derive a number of useful properties of the irrelevant natural extension. It has powerful marginalisation properties (IV) and satisfies all graphoid properties but symmetry, both in their direct and reverse forms (V & VI).

Sets of desirable gambles (SDG)

We will model a subject's beliefs about the value that a variable X assumes in some set \mathcal{X} , by means of his behaviour, which gambles (real-valued maps) f on \mathcal{X} does he strictly prefer to the status quo. This results in a set of desirable gambles $\mathcal{D} \subseteq \mathcal{F}(\mathcal{X})$, where $\mathcal{F}(\mathcal{X})$ is the set of all gambles on \mathcal{X} . \mathcal{D} is called coherent if it satisfies the rationality requirements D1–D4 for all $f, f_1, f_2 \in \mathcal{F}(\mathcal{X})$ and all real $\lambda > 0$.

- D1 $f \leq 0 \Rightarrow f \notin \mathcal{D}$
- D2 $f > 0 \Rightarrow f \in \mathcal{D}$
- D3 $f \in \mathcal{D} \Rightarrow \lambda f \in \mathcal{D}$
- D4 $f_1, f_2 \in \mathcal{D} \Rightarrow f_1 + f_2 \in \mathcal{D}$

Although they are not as well known as other (imprecise) probability models, sets of desirable gambles have definite advantages. To give a few examples: they are operational, are easily able to deal with conditioning on events with probability zero, allow for intuitive geometrically flavoured proofs and are more expressive than both credal sets and lower previsions (see our papers or the second poster for credal networks under epistemic irrelevance that use these alternative models).

Local uncertainty models (I)

With every node x of a finite directed acyclic graph (DAG), we associate a variable X_x taking values in some finite, non-empty set \mathcal{X}_x . The set of all nodes is denoted by G . For every subset $S \subseteq G$, the joint variable X_S takes values in $\mathcal{X}_S := \times_{x \in S} \mathcal{X}_x$. For every $x \in G$, we denote by $P(x)$ the set consisting of the parent nodes of x . Similar to what is done in classical Bayesian networks, we attach local uncertainty models to the nodes of the network, conditional on the value of their parents. For all $x \in G$ and every instantiation $x_{P(x)} \in \mathcal{X}_{P(x)}$, we require a coherent set $\mathcal{D}_{x|P(x)}$ of desirable gambles (SDG) on \mathcal{X}_x .

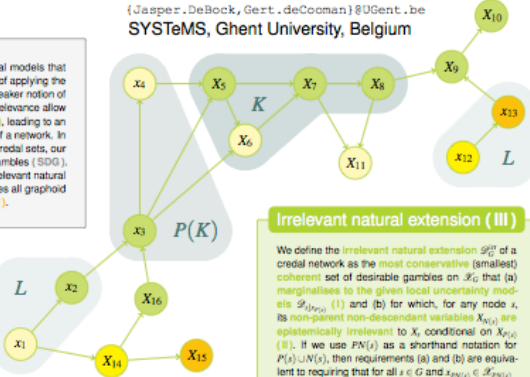
AD-separation (V)

Consider any path x_1, \dots, x_n in G , with $n \geq 1$. We say that this path is blocked by a set of nodes $C \subseteq G$ whenever at least one of the following four conditions holds:

- B1 $x_1 \in C$;
- B2 there is some $1 < i < n$ such that $x_i \rightarrow x_{i+1}$ and $x_i \in C$;
- B3 there is some $1 < i < n$ such that $x_{i-1} \rightarrow x_i \leftarrow x_{i+1}$, $x_i \in C$ and $D(x_i) \cap C = \emptyset$;
- B4 $x_n \in C$.

Now consider (not necessarily disjoint) subsets I, O and C of G . We say that O is AD-separated from I by C , denoted as $AD(I, O|C)$, if every path $I = x_1, \dots, x_n = o$, $n \geq 1$, from a node $i \in I$ to a node $o \in O$, is blocked by C .

This asymmetrical version of D-separation is similar to, yet different from both Moral's (2005) version of AD-separation and the notion of L-separation, as introduced by Vantaggi (2002). Our reason for not using one of these existing concepts is that our version of AD-separation has stronger properties: it satisfies all graphoid properties except symmetry: it satisfies redundancy, decomposition, weak union, contraction and intersection both in their direct and reverse forms.

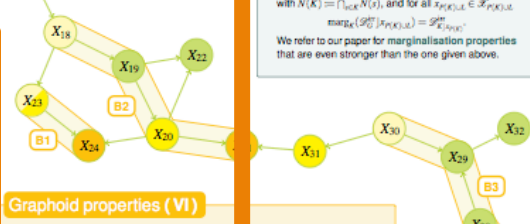


Epistemic irrelevance (II)

Consider a global set of desirable gambles \mathcal{D} (SDG) on \mathcal{X}_G (1) and disjoint subsets S and K of G . Then the marginal model for X_S , conditional on the information that X_K assumes a value $x_K \in \mathcal{X}_K$, is given by $\text{marg}_K(\mathcal{D}_{x_K}) = \{f \in \mathcal{F}(\mathcal{X}_S) : 1_{x_K} f \in \mathcal{D}\}$. Consider now three subsets $C, I, O \subseteq G$, with $I \cap C$ and $O \cap C$ disjoint. We say that X_I is epistemically irrelevant to X_O conditional on X_C , denoted as $IR(I, O|C)$, if and only if for all $x_{C \cup I} \in \mathcal{X}_{C \cup I}$ we have $\text{marg}_{O \cap C}(\mathcal{D}_{x_{C \cup I}}) = \text{marg}_{O \cap C}(\mathcal{D}_{x_{C \cup I}})$.

Our paper also considers epistemically sub-set-irrelevance, which although interesting, is not discussed on this poster.

AD(I, O|C) and IR(I, O|C)



Graphoid properties (VI)

Every AD-separation corresponds to a conditional epistemic irrelevance statement that is satisfied by the corresponding irrelevant natural extension $\mathcal{D}_{x_{P(x) \setminus C}}^{\text{irr}}$ for all $I, O, C \subseteq G$ such that $AD(I, O|C)$ (II) (we refer to our paper for even stronger results regarding factorisation and subset-irrelevance (II)). Furthermore, the properties of AD-separation (V) imply that the corresponding family of irrelevance statements is closed under all graphoid properties except symmetry, both in their direct and reverse form. We want to stress that our proof for this result is based solely on marginalisation results (IV) and properties of AD-separation (V). At no point does it invoke graphoid properties of epistemic irrelevance. We believe that this provides a new, rather more positive perspective on the well-known fact that epistemic irrelevance, in general, can fall some of the graphoid axioms (Cozman 2005).

Irrelevant natural extension (III)

We define the irrelevant natural extension $\mathcal{D}_{x_{P(x)}}^{\text{irr}}$ of a credal network as the most conservative (smallest) coherent set of desirable gambles on \mathcal{X}_G that (a) marginalises to the given local uncertainty models $\mathcal{D}_{x_{P(x)}}$ (1) and (b) for which, for any node x , its non-parent non-descendant variables $X_{N(x)}$ are epistemically irrelevant to X_x conditional on $X_{P(x)}$ (II). If we use $FN(x)$ as a shorthand notation for $P(x) \cup N(x)$, then requirements (a) and (b) are equivalent to requiring that for all $x \in G$ and $x_{P(x)} \in \mathcal{X}_{P(x)}$:

$\text{marg}_x(\mathcal{D}_{x_{P(x)}}^{\text{irr}}) = \text{marg}_x(\mathcal{D}_{x_{P(x)}})$ and $\text{marg}_x(\mathcal{D}_{x_{P(x)}}^{\text{irr}}) = \mathcal{D}_{x_{P(x)}}$. We show that this irrelevant natural extension is simple to construct: $\mathcal{D}_{x_{P(x)}}^{\text{irr}} := \text{pos}(a_{x_{P(x)}}^{\text{irr}})$, where the 'post'-operator generates the set of all finite positive linear combinations of elements in its argument set $a_{x_{P(x)}}^{\text{irr}} := \{1_{x_{P(x)}} f : f \in \mathcal{D}_{x_{P(x)}}, f \in \mathcal{D}_{x_{P(x) \setminus C}}\}$.

Marginalisation properties (IV)

For any $K \subseteq G$, we construct a sub-DAG of the original DAG by eliminating the nodes $s \in G \setminus K$ and their associated edges. The parents of a node $x \in K$, with respect to this sub-DAG, are denoted by $P_K(x) := P(x) \cap K$. We derive local models $\mathcal{D}_{x|P_K(x)}$ for this sub-DAG from the original local models $\mathcal{D}_{x|P(x)}$ by fixing $x_{P(x)}$. We do this consistently for all $x \in K$ at once by fixing $x_{P(K)}$ where $P(K) := (\cup_{x \in K} P(x)) \cap K$. For any $x_{P(K)} \in \mathcal{X}_{P(K)}$, we use the resulting local models to construct an irrelevant natural extension of the sub-DAG and denote it by $\mathcal{D}_{x_{P(K)}}^{\text{irr}}$.

Suppose now that K is closed subset of G , meaning that $P(K) \cap D(K) = \emptyset$, where $D(K) := (\cup_{x \in K} D(x)) \cap K$ and for all $x \in K$, $D(x)$ are the descendants of x . Then $\mathcal{D}_{x_{P(K)}}^{\text{irr}}$ is related to $\mathcal{D}_{x_{P(x)}}$ as follows: for all $x \in \mathcal{X}_{P(K)}$, with $N(x) = \cap_{x \in K} N(x)$, and for all $x_{P(x) \setminus K} \in \mathcal{X}_{P(x) \setminus K}$, $\text{marg}_{x_{P(x) \setminus K}}(\mathcal{D}_{x_{P(K)}}^{\text{irr}}) = \mathcal{D}_{x_{P(x) \setminus K}}^{\text{irr}}$.

We refer to our paper for marginalisation properties that are even stronger than the one given above.

- Advantages of sets of desirable gambles
- Marginalisation properties
- AD-separation satisfies every graphoid property except symmetry



Credal networks under epistemic irrelevance using sets of desirable gambles

Jasper De Bock and Gert de Cooman
{Jasper.DeBock, Gert.deCooman}@UGent.be
SYSTeMS, Ghent University, Belgium

Abstract of the paper

We present a new approach to credal networks, which are graphical models that generalise Bayesian nets to deal with imprecise probabilities. Instead of applying the commonly used notion of strong independence, we replace it by the weaker notion of epistemic irrelevance (II). We show how assessments of epistemic irrelevance allow us to construct a global model out of given local uncertainty models (I), leading to an intuitive expression for the so-called irrelevant natural extension (III) of a network. In contrast with Cozman (2000) who introduced this notion in terms of credal sets, our main results are presented using the language of sets of desirable gambles (SDG). This has allowed us to derive a number of useful properties of the irrelevant natural extension. It has powerful marginalisation properties (IV) and satisfies all graphoid properties but symmetry, both in their direct and reverse forms (V & VI).

Sets of desirable gambles (SDG)

We will model a subject's beliefs about the value that a variable X assumes in some set \mathcal{X} , by means of his behaviour, which gambles (real-valued maps) f on \mathcal{X} does he strictly prefer to the status quo. This results in a set of desirable gambles $\mathcal{D} \subseteq \mathcal{F}(\mathcal{X})$, where $\mathcal{F}(\mathcal{X})$ is the set of all gambles on \mathcal{X} . \mathcal{D} is called coherent if it satisfies the rationality requirements D1–D4 for all $f, f_1, f_2 \in \mathcal{F}(\mathcal{X})$ and all real $\lambda > 0$.

- D1 $f \leq 0 \Rightarrow f \notin \mathcal{D}$
- D2 $f > 0 \Rightarrow f \in \mathcal{D}$
- D3 $f \in \mathcal{D} \Rightarrow \lambda f \in \mathcal{D}$
- D4 $f_1, f_2 \in \mathcal{D} \Rightarrow f_1 + f_2 \in \mathcal{D}$

Although they are not as well known as other (imprecise) probability models, sets of desirable gambles have definite advantages. To give a few examples: they are operational, are easily able to deal with conditioning on events with probability zero, allow for intuitive geometrically flavoured proofs and are more expressive than both credal sets and lower previsions (see our papers or the second poster for credal networks under epistemic irrelevance that use these alternative models).

Local uncertainty models (I)

With every node x of a finite directed acyclic graph (DAG), we associate a variable X_x taking values in some finite, non-empty set \mathcal{X}_x . The set of all nodes is denoted by G . For every subset $S \subseteq G$, the joint variable X_S takes values in $\mathcal{X}_S := \times_{x \in S} \mathcal{X}_x$. For every $x \in G$, we denote by $P(x)$ the set consisting of the parent nodes of x . Similar to what is done in classical Bayesian networks, we attach local uncertainty models to the nodes of the network, conditional on the value of their parents. For all $x \in G$ and every instantiation $x_{P(x)} \in \mathcal{X}_{P(x)}$, we require a coherent set $\mathcal{D}_{x|P(x)}$ of desirable gambles (SDG) on \mathcal{X}_x .

AD-separation (V)

Consider any path x_1, \dots, x_n in G , with $n \geq 1$. We say that this path is blocked by a set of nodes $C \subseteq G$ whenever at least one of the following four conditions holds:

- B1 $x_1 \in C$;
- B2 there is some $1 < i < n$ such that $x_i \rightarrow x_{i+1}$ and $x_i \in C$;
- B3 there is some $1 < i < n$ such that $x_{i-1} \rightarrow x_i \leftarrow x_{i+1}$, $x_i \in C$ and $D(x_i) \cap C = \emptyset$;
- B4 $x_n \in C$.

Now consider (not necessarily disjoint) subsets I, O and C of G . We say that O is AD-separated from I by C , denoted as $AD(I, O|C)$, if every path $I \rightarrow x_1, \dots, x_n \leftarrow O$, $n \geq 1$, from a node $I \in I$ to a node $O \in O$, is blocked by C .

This asymmetrical version of D-separation is similar to, yet different from both Moral's (2005) version of AD-separation and the notion of L-separation, as introduced by Vantaggi (2002). Our reason for not using one of these existing concepts is that our version of AD-separation has stronger properties: it satisfies all graphoid properties except symmetry: it satisfies redundancy, decomposition, weak union, contraction and intersection both in their direct and reverse forms.

Epistemic irrelevance (II)

Consider a global set of desirable gambles \mathcal{D} (SDG) on \mathcal{X}_G and disjoint subsets S and K of G . Then the marginal model for X_S , conditional on the information that X_K assumes a value $x_K \in \mathcal{X}_K$, is given by $\text{marg}_K(\mathcal{D}_{S|X_S}) = \{f \in \mathcal{F}(\mathcal{X}_S) : \exists_{x_K} f \in \mathcal{D}\}$. Consider now three subsets $C, I, O \subseteq G$, with $I \cap C$ and $O \cap C$ disjoint. We say that X_I is epistemically irrelevant to X_O conditional on X_C , denoted as $IR(I, O|C)$, if and only if for all $x_{C \cup O} \in \mathcal{X}_{C \cup O}$ we have

$$\text{marg}_{O \cup C}(\mathcal{D}_{I|X_I, x_{C \cup O}}) = \text{marg}_{O \cup C}(\mathcal{D}_{I|X_I, C})$$

Our paper also considers epistemically sub-set-irrelevance, which although interesting, is not discussed in this poster.

$$AD(I, O|C) \\ IR(I, O|C)$$

Graphoid properties (VI)

In a credal network under epistemic irrelevance, every AD-separation corresponds to a conditional epistemic irrelevance statement that is satisfied by the corresponding irrelevant natural extension $\mathcal{D}_{G|G}^{\text{irr}}$, for all $I, O, C \subseteq G$ such that $AD(I, O|C)$ we have $IR(I, O|C)$ (II) (we refer to our paper for even stronger results regarding factorisation and subset-irrelevance (II)). Furthermore, the properties of AD-separation (V) imply that the corresponding family of irrelevance statements is closed under all graphoid properties except symmetry, both in their direct and reverse form. We want to stress that our proof for this result is based solely on marginalisation results (IV) and properties of AD-separation (V). At no point does it invoke graphoid properties of epistemic irrelevance. We believe that this provides a new, rather more positive perspective on the well-known fact that epistemic irrelevance, in general, can fail some of the graphoid axioms (Cozman 2005).

Irrelevant natural extension (III)

We define the irrelevant natural extension $\mathcal{D}_{G|G}^{\text{irr}}$ of a credal network as the most conservative (smallest) coherent set of desirable gambles on \mathcal{X}_G that (a) marginalises to the given local uncertainty models $\mathcal{D}_{x|P(x)}$ (I) and (b) for which, for any node x , its non-parent non-descendant variables $X_{N(x)}$ are epistemically irrelevant to X_x conditional on $X_{P(x)}$ (II). If we use $FN(x)$ as a shorthand notation for $P(x) \cup N(x)$, then requirements (a) and (b) are equivalent to requiring that for all $x \in G$ and $x_{P(x)} \in \mathcal{X}_{P(x)}$

$$\text{marg}_x(\mathcal{D}_{x|P(x)}^{\text{irr}}) = \text{marg}_x(\mathcal{D}_{x|P(x)}^{\text{irr}}) = \mathcal{D}_{x|P(x)}$$

We show that this irrelevant natural extension is simple to construct: $\mathcal{D}_{G|G}^{\text{irr}} := \text{pos}(\mathcal{A}_{G|G}^{\text{irr}})$, where the 'pos'-operator generates the set of all finite positive linear combinations of elements in its argument set

$$\mathcal{A}_{G|G}^{\text{irr}} := \{1_{x_{P(x)}}, f : x \in G, x_{P(x)} \in \mathcal{X}_{P(x)}, f \in \mathcal{D}_{x|P(x)}\}$$

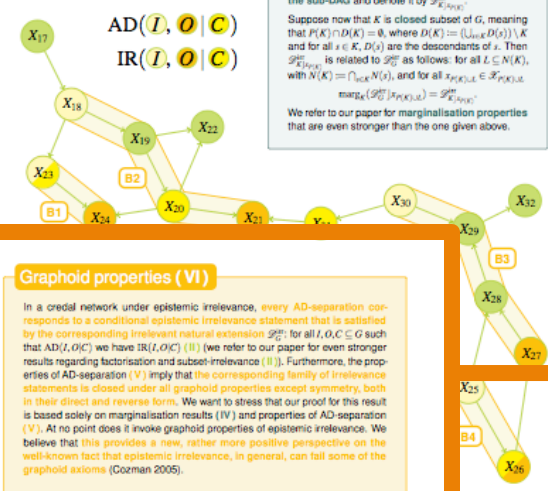
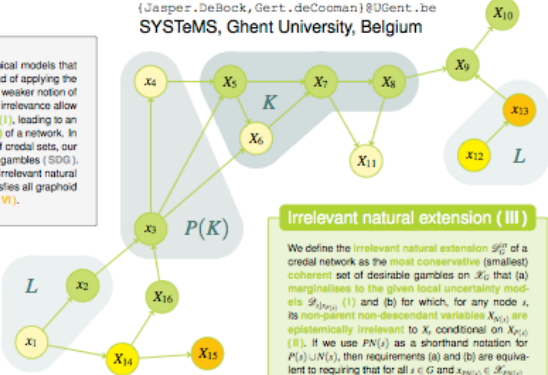
Marginalisation properties (IV)

For any $K \subseteq G$, we construct a sub-DAG of the original DAG by eliminating the nodes $s \in G \setminus K$ and their associated edges. The parents of a node $x \in K$, with respect to this sub-DAG, are denoted by $P_K(x) := P(x) \cap K$. We derive local models $\mathcal{D}_{x|P_K(x)}$ for this sub-DAG from the original local models $\mathcal{D}_{x|P(x)}$ by fixing $x_{P(x)}$. We do this consistently for all $x \in K$ at once by fixing $x_{P(K)}$ where $P(K) := (\cup_{x \in K} P(x)) \cap K$. For any $x_{P(K)} \in \mathcal{X}_{P(K)}$, we use the resulting local models to construct an irrelevant natural extension of the sub-DAG and denote it by $\mathcal{D}_{G|G}^{\text{irr}}$.

Suppose now that K is closed subset of G , meaning that $P(K) \cap D(K) = \emptyset$, where $D(K) := (\cup_{x \in K} D(x)) \cap K$ and for all $x \in K$, $D(x)$ are the descendants of x . Then $\mathcal{D}_{G|G}^{\text{irr}}$ is related to $\mathcal{D}_{G|G}^{\text{irr}}$ as follows: for all $I \subseteq N(K)$, with $N(K) := \cap_{x \in K} N(x)$, and for all $x_{P(K) \cup I} \in \mathcal{X}_{P(K) \cup I}$

$$\text{marg}_I(\mathcal{D}_{G|G}^{\text{irr}}) = \mathcal{D}_{G|G}^{\text{irr}}$$

We refer to our paper for marginalisation properties that are even stronger than the one given above.



- Advantages of sets of desirable gambles
- Marginalisation properties
- AD-separation satisfies every graphoid property except symmetry
- AD-separation implies epistemic irrelevance

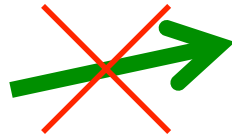
Credal networks: useful properties ✓



Conditional
marginalisation
properties

AD-separated
evidence is
not relevant

~~graphoid
axioms~~



Barren nodes
can be removed



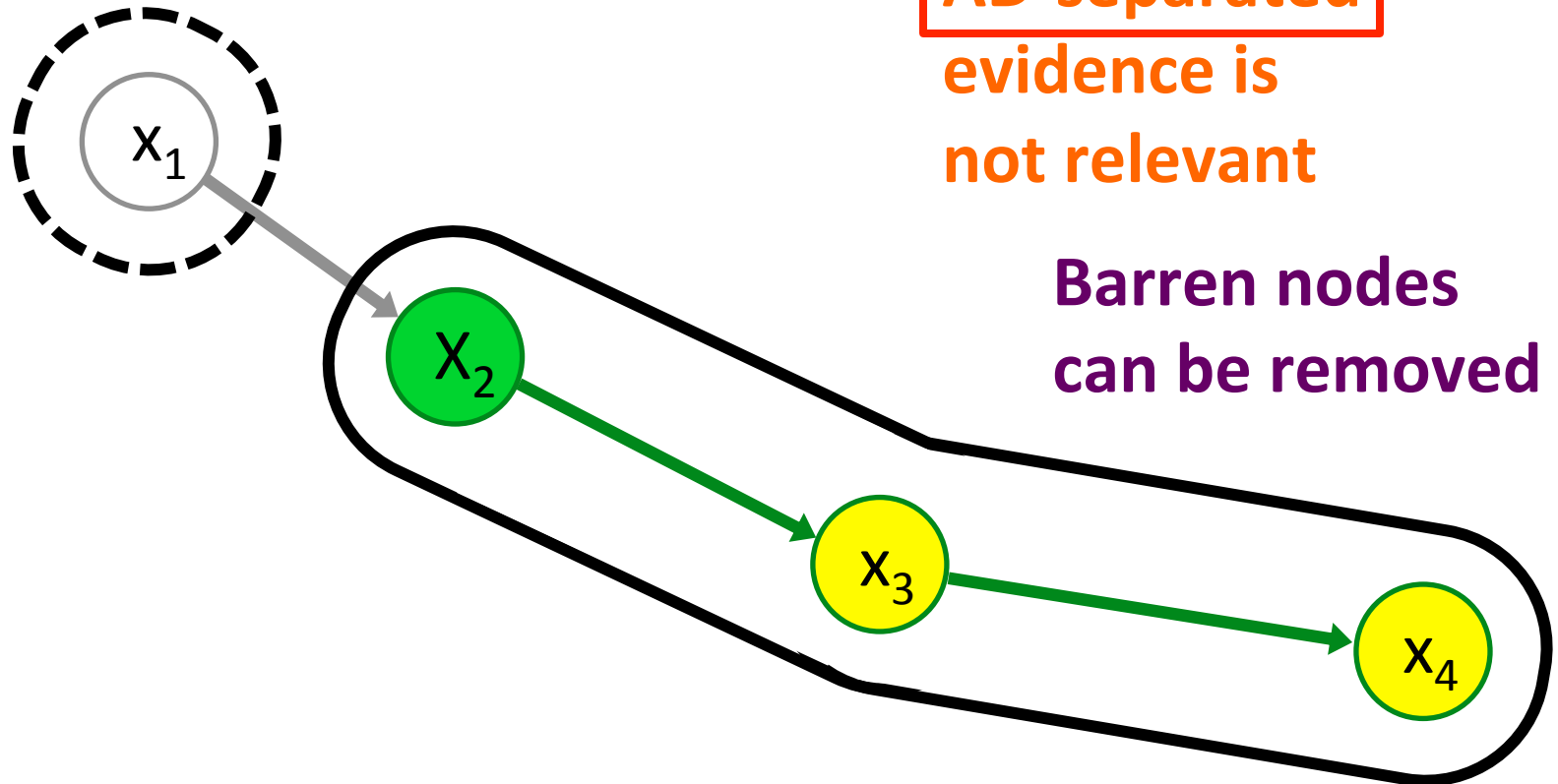
Credal networks under
epistemic irrelevance

Credal networks: useful properties ✓

Conditional
marginalisation
properties

AD-separated
evidence is
not relevant

Barren nodes
can be removed



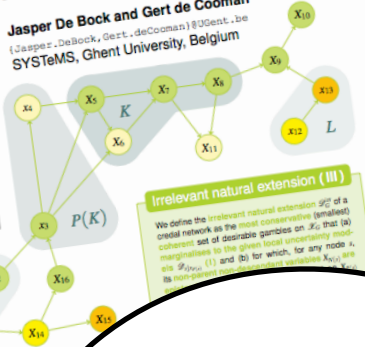
Hope to see you at the poster session!

Credal networks under epistemic irrelevance using sets of desirable gambles

Jasper De Bock and Gert de Cooman
 (Jasper_DeBock, Gert_DeCooman)@UGent.be
 SYSTeMS, Ghent University, Belgium

Abstract of the paper

We present a new approach to credal networks, which are graphical models that generalize Bayesian nets to deal with imprecise probabilities. Instead of applying the generative Bayesian nets to deal with imprecise probabilities, we replace it by the weaker notion of commonly used notions of (local) epistemic irrelevance, leading to a credal network. We show how assessments of epistemic irrelevance allow us to construct a global model out of given local uncertainty models (LUM), leading to an intuitive expression for the so-called irrelevant natural extension (INE) of a network, in contrast with Cozman (2005) who introduced this notion in terms of credal sets. Our main results are presented using the language of sets of desirable gambles (SDG), which has allowed us to derive a number of useful properties of the irrelevant natural extension. It has powerful marginalization properties (IV) and satisfies all graphoid properties but symmetry, both in their direct and reverse forms (V & VI).



Irrelevant natural extension (INE)

We define the irrelevant natural extension \mathcal{G}^I of a credal network as the most conservative (smallest) coherent set of desirable gambles on X_0 that marginalizes to the given local uncertainty models \mathcal{G}_i for $i \in I$ and for which, for any node x_i , its non-parent nodes $X_{pa(x_i)}$ are its non-parent nodes in \mathcal{G}_i .

Sets of desirable gambles (SDG)

We will model a subject's beliefs about the value that a variable X assumes in some set X , by means of its behaviour, which gambles (real-valued maps) f on X does he strictly prefer to the status quo. This results in a set of desirable gambles $\mathcal{D} \subseteq \mathcal{F}(X)$, where $\mathcal{F}(X)$ is the set of all gambles on X . \mathcal{D} is called coherent if it satisfies the rationality requirements D1–D4 for all $f, g, h \in \mathcal{F}(X)$ and all real $\lambda > 0$.

Although they are not as well known as other (imprecise) probability models, sets of desirable gambles have definite advantages. To give a few examples: they are operationally able to deal with conditioning on events with probability zero, allow for intuitive graphical representations with probability zero, are more expressive than rationally favoured points and lower previsions (see our papers both on credal sets and lower previsions under epistemic irrelevance that use these alternative models).

Local uncertainty models (LUM)

With every node x_i of a finite directed acyclic graph (DAG), we associate a variable X_i taking values in some finite, non-empty set \mathcal{X}_i . The set of all nodes is denoted by \mathcal{X} . For every subset $S \subseteq \mathcal{X}$, the joint variable X_S takes values in $\mathcal{X}_S := \prod_{x_i \in S} \mathcal{X}_i$. For every $x_i \in \mathcal{X}$, we denote by $P(x_i)$ the set consisting of the parent nodes of x_i . Similar to what is done in classical Bayesian networks, we attach local uncertainty models to the nodes of the network, conditional on the value of their parents. For all $x_i \in \mathcal{X}$ and every instantiation $x_{pa(x_i)} \in \mathcal{X}_{pa(x_i)}$ (SDG) on X_i , we require a coherent set $\mathcal{D}_{x_i|x_{pa(x_i)}}$ of desirable gambles (SDG) on X_i .

AD-separation (V)

Consider any path x_1, \dots, x_n in G , with $n \geq 2$. We say that this path is *broken* by a set of nodes $C \subseteq G$ whenever at least one of the following four conditions holds:
 B1 $x_1 \in C$;
 B2 $x_n \in C$;
 B3 $x_i \in C$ and $x_{i+1} \in C$ such that $x_i \rightarrow x_{i+1}$ and $x_{i+1} \rightarrow x_i$;
 B4 $x_i \in C$ and $x_{i+1} \in C$ derived from $x_i \rightarrow x_{i+1}$ and $x_{i+1} \rightarrow x_i$.

Graphoid properties (VI)

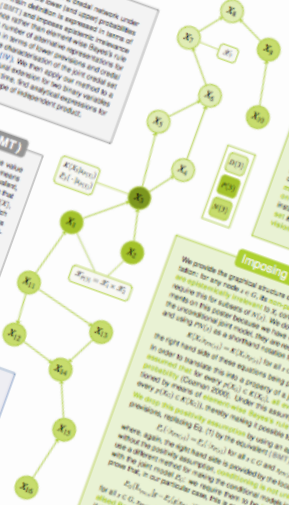
In a credal network under epistemic irrelevance, every AD-separation corresponds to a conditional epistemic irrelevance statement that is satisfied by the corresponding irrelevant natural extension \mathcal{G}^I : for all $A, B, C \subseteq G$ such that $(A, B | C)$ is an AD-separation (V) we have (IRLOCAL) (I) (we refer to our paper for even stronger results that $(A, B | C)$ imply that the corresponding family of irrelevance results regarding factorization and subadditivity properties, both in their direct and reverse form. We want to stress that our proof for this result is based solely on marginalization results (IV) and properties of AD-separation (V). At no point does it involve graphoid properties of epistemic irrelevance. We believe that this provides a new, rather more intuitive perspective on the well-known fact that epistemic irrelevance, in general, can fall some of the graphoid axioms (Cozman 2005).

Allowing for probability zero in credal networks under epistemic irrelevance

Jasper De Bock and Gert de Cooman
 (Jasper_DeBock, Gert_DeCooman)@UGent.be
 SYSTeMS, Ghent University, Belgium

Abstract of the paper

We generalize Cozman (2005) concept of a credal network under epistemic irrelevance to the case where local uncertainty models are allowed to be zero. Our main definition is based on local uncertainty models (LUM) that consist of a set of desirable gambles (SDG) on the node X_i , which may be empty. This allows for the possibility of having probability zero in the network, which is not possible in the original framework of Cozman (2005). We show how this is achieved by means of a new characterization of the irrelevant natural extension (INE) of a network. We also show how this is achieved by means of a new characterization of the irrelevant natural extension (INE) of a network. We also show how this is achieved by means of a new characterization of the irrelevant natural extension (INE) of a network.



Local uncertainty models (LUM)

We have introduced a finite directed acyclic graph (DAG), we associate a variable X_i taking values in some finite, non-empty set \mathcal{X}_i . The set of all nodes is denoted by \mathcal{X} . For every subset $S \subseteq \mathcal{X}$, the joint variable X_S takes values in $\mathcal{X}_S := \prod_{x_i \in S} \mathcal{X}_i$. For every $x_i \in \mathcal{X}$, we denote by $P(x_i)$ the set consisting of the parent nodes of x_i . Similar to what is done in classical Bayesian networks, we attach local uncertainty models to the nodes of the network, conditional on the value of their parents. For all $x_i \in \mathcal{X}$ and every instantiation $x_{pa(x_i)} \in \mathcal{X}_{pa(x_i)}$ (LUM) on X_i , we require a coherent set $\mathcal{D}_{x_i|x_{pa(x_i)}}$ of desirable gambles (SDG) on X_i .

Imposing epistemic irrelevance (II)

We provide the precise structure of the network with the following interpretation: for any node $x_i \in G$, the local uncertainty model $\mathcal{D}_{x_i|x_{pa(x_i)}}$ is a coherent set of desirable gambles (SDG) on X_i , which may be empty. This allows for the possibility of having probability zero in the network, which is not possible in the original framework of Cozman (2005). We show how this is achieved by means of a new characterization of the irrelevant natural extension (INE) of a network. We also show how this is achieved by means of a new characterization of the irrelevant natural extension (INE) of a network.

AD-separation for two binary variables (V)

X_1	X_2	X_3	X_4	X_5
0	0	0	0	0
0	0	0	1	0
0	0	1	0	0
0	0	1	1	0
0	1	0	0	0
0	1	0	1	0
0	1	1	0	0
0	1	1	1	0
1	0	0	0	0
1	0	0	1	0
1	0	1	0	0
1	0	1	1	0
1	1	0	0	0
1	1	0	1	0
1	1	1	0	0
1	1	1	1	0

Irrelevant natural extension (INE)

We define the irrelevant natural extension \mathcal{G}^I of a credal network as the most conservative (smallest) coherent set of desirable gambles on X_0 that marginalizes to the given local uncertainty models \mathcal{G}_i for $i \in I$ and for which, for any node x_i , its non-parent nodes $X_{pa(x_i)}$ are its non-parent nodes in \mathcal{G}_i .

