UNIVERSITEIT GENT

FACULTY OF ENGINEERING

The MePiCTIr message passing algorithm:

Expert system inference in credal trees under epistemic irrelevance

Gert de Cooman

Ghent University, SYSTeMS
gert.decooman@UGent.be

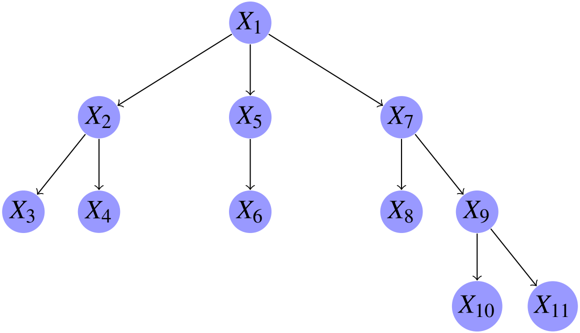
WPMSIIP2
Ludwig-Maximilians-Universität München
10 September 2009

Credal networks

The special case of a tree

Basic concept

Consider a directed tree, with a variable X_s attached to each node s.



Credal trees

Local uncertainty models

Local uncertainty model associated with each node s

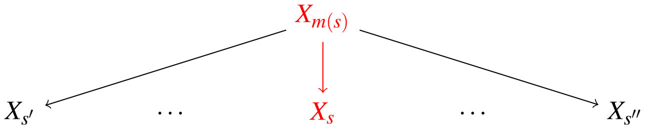
- the variable X_s may assume a value in the finite set $\mathscr{X}_s$;
- for each possible value $x_{m(s)} \in \mathscr{X}_{m(s)}$ of the mother variable $X_{m(s)}$, we have a conditional lower expectation

$$\underline{Q}_s(\cdot|x_{m(s)})\colon \mathscr{L}(\mathscr{X}_s) \to \mathbb{R}$$

where

$\underline{Q}_s(f|x_{m(s)}) =$ lower expectation of $f(X_s)$, given that $X_{m(s)} = x_{m(s)}$

- local model $\underline{Q}_s(\cdot|X_{m(s)})$ is a conditional lower expectation operator

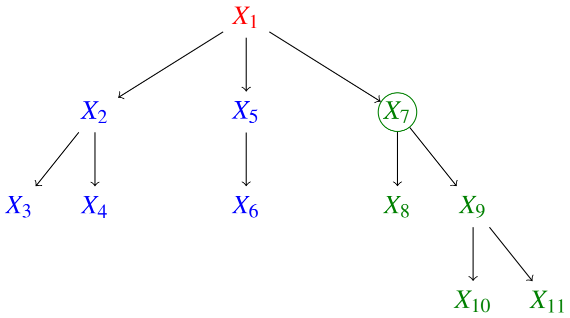


Credal trees under epistemic irrelevance

Definition

The graphical structure is interpreted as follows:

Conditional on the mother variable, the non-parent non-descendants of each node variable are epistemically irrelevant to it and its descendants.



Credal networks under epistemic irrelevance

As an expert system

When the credal network is a (Markov) tree we can find the joint model from the local models recursively, from leaves to root.

Exact message passing algorithm

- credal tree treated as an expert system
- linear complexity in the number of nodes

Python code

- written by Filip Hermans
- testing and connection with strong independence in cooperation with Marco Zaffalon and Alessandro Antonucci

Current (toy) applications in HMMs

- character recognition
- air traffic trajectory tracking and identification

Constructing global from local models

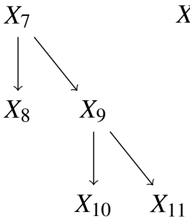
From local to global models

For each node s, we want to construct a global model

$$\underline{P}_s(\cdot|X_{m(s)})\colon \mathscr{L}(\mathscr{X}_{\downarrow s}) \to \mathbb{R}$$

representing the uncertainty about all variables $X_{\downarrow s}$ in the subtree below s:

$$\downarrow s = \{t\colon t \text{ lies below } s\}$$



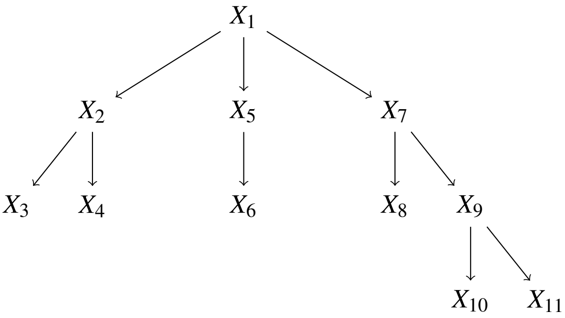
$$X_{\downarrow 7} = (X_7, X_8, X_9, X_{10}, X_{11})$$

Constructing global from local models

How to construct the global models

First crucial observation

We can build up any tree in a recursive fashion, repeating much simpler basic building blocks.

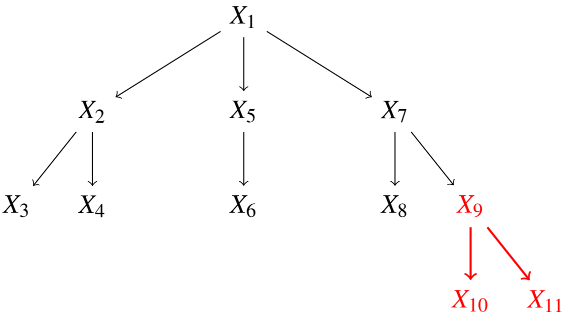


Constructing global from local models

How to construct the global models

First crucial observation

We can build up any tree in a recursive fashion, repeating much simpler basic building blocks.

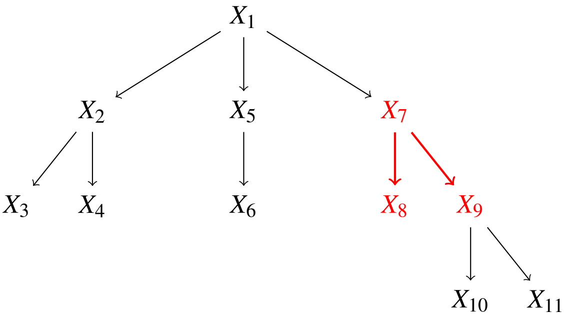


Constructing global from local models

How to construct the global models

First crucial observation

We can build up any tree in a recursive fashion, repeating much simpler basic building blocks.

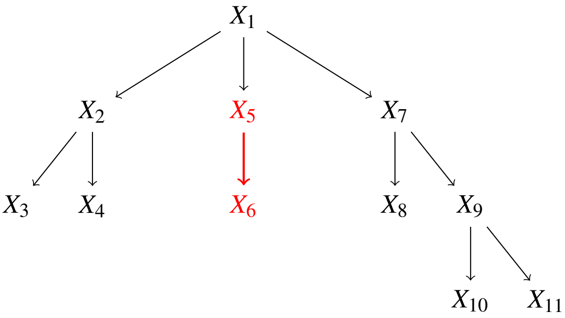


Constructing global from local models

How to construct the global models

First crucial observation

We can build up any tree in a recursive fashion, repeating much simpler basic building blocks.

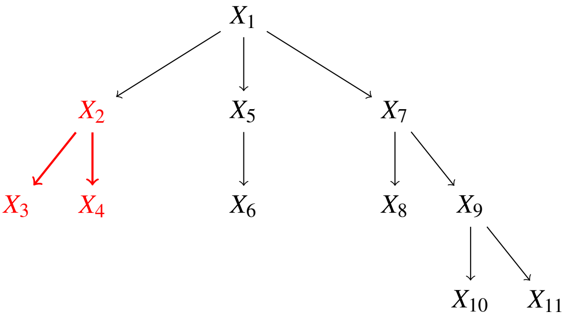


Constructing global from local models

How to construct the global models

First crucial observation

We can build up any tree in a recursive fashion, repeating much simpler basic building blocks.

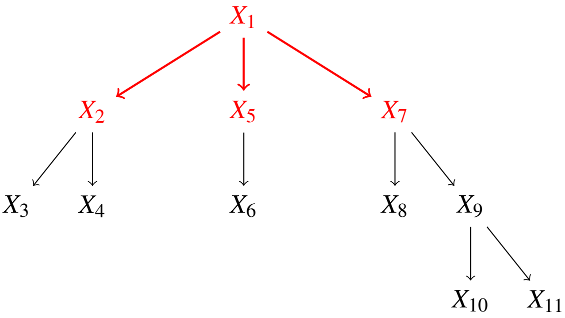


Constructing global from local models

How to construct the global models

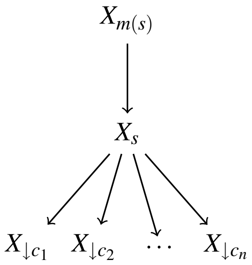
First crucial observation

We can build up any tree in a recursive fashion, repeating much simpler basic building blocks.



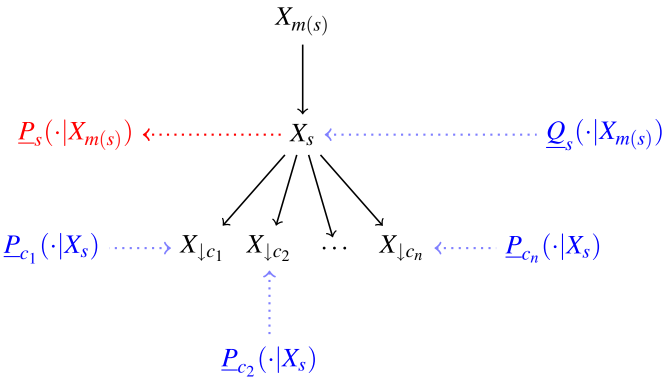
Constructing global from local models

Recursive construction of the joint: graphical representation



Constructing global from local models

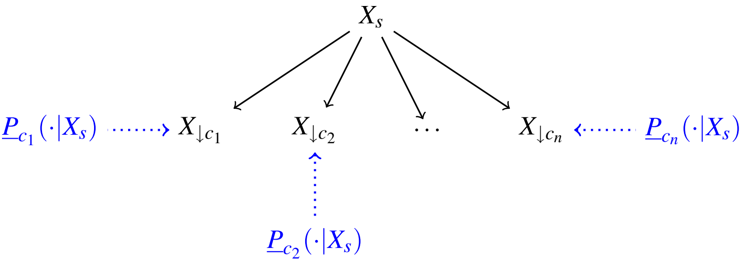
Recursive construction of the joint: graphical representation



Constructing global from local models

Recursive construction of the joint: epistemically independent product

Conditional on X_s, the child variables X_c, $c \in C(s)$ are epistemically independent.

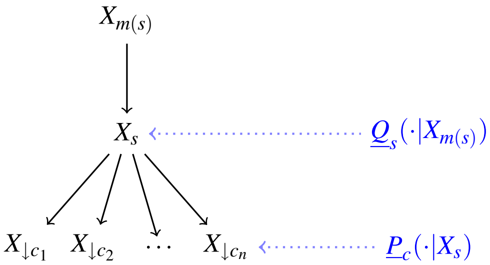


Epistemically independent product on $\mathscr{L}(X_{\downarrow s \setminus \{s\}})$:

$$\underline{E}_s(\cdot|X_s) = \otimes_{c \in C(s)} \underline{P}_c(\cdot|X_s)$$

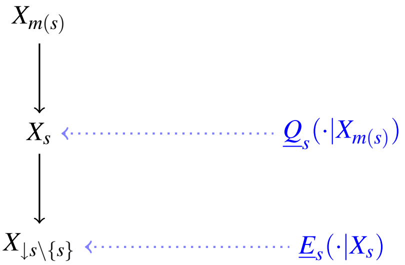
Constructing global from local models

Recursive construction of the joint: marginal extension



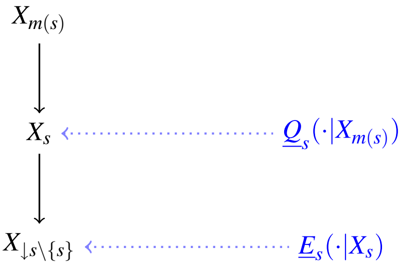
Constructing global from local models

Recursive construction of the joint: marginal extension



Constructing global from local models

Recursive construction of the joint: marginal extension



Marginal extension or Law of iterated expectation:

$$\underline{P}_s(\cdot|X_{m(s)}) = \underline{Q}_s(\underline{E}_s(\cdot|X_s)|X_{m(s)}) = \underline{Q}_s(\otimes_{c\in C(s)}\underline{P}_c(\cdot|X_s)|X_{m(s)})$$

Constructing global from local models

Recursive construction of the joint: marginal extension

Marginal extension or Law of iterated expectation:

$$\underline{P}_s(\cdot|X_{m(s)}) = \underline{Q}_s(\underline{E}_s(\cdot|X_s)|X_{m(s)}) = \underline{Q}_s(\otimes_{c\in C(s)}\underline{P}_c(\cdot|X_s)|X_{m(s)})$$

Constructing global from local models

Fundamental result

We start at the leaves with:

$$\underline{P}_t(\cdot|X_{m(t)}) = \underline{Q}_t(\cdot|X_{m(t)}) \text{ for all leaves } t$$

and move recursively upwards to the root of the tree.

Theorem

In this recursive fashion, we construct the point-wise smallest (most conservative) collection of global models $\underline{P}_s(\cdot|X_{m(s)})$ that

(i) are coherent;

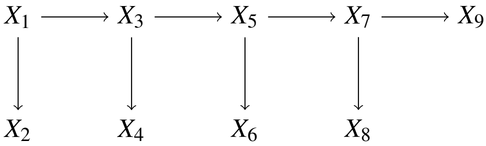
(ii) are compatible with the local information;

(iii) reflect all the epistemic irrelevancies embodied in the graphical structure.

UNIVERSITEIT GENT

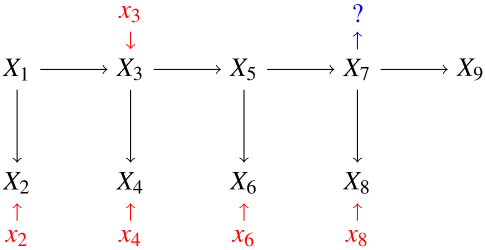
Explaining the basics of the algorithm

An example



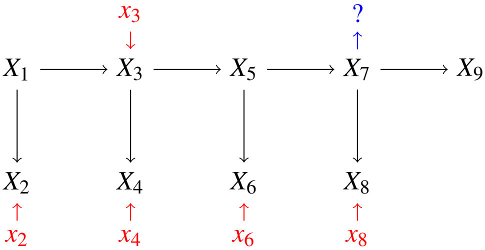
Explaining the basics of the algorithm

An example



Explaining the basics of the algorithm

An example

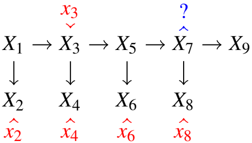


Regular extension

$$
\begin{aligned}
&\underline{R}(g(X_7)|X_2 = x_2, X_3 = x_3, X_4 = x_4, X_6 = x_6, X_8 = x_8) \\
&\qquad = \max\left\{\mu \in \mathbb{R} \colon \underline{P}_1\left(I_{\{x_2\}} I_{\{x_3\}} I_{\{x_4\}} I_{\{x_6\}} I_{\{x_8\}} [g(X_7) - \mu]\right)\right\}
\end{aligned}
$$

Explaining the basics of the algorithm

An example



We have to calculate $\underline{P}_1(f_1)$, where

$$f_1 = I_{\{x_2\}} I_{\{x_3\}} I_{\{x_4\}} I_{\{x_6\}} I_{\{x_8\}} [g(X_7) - \mu] = I_{\{x_2\}} f_3$$

Now use the recursion formula for $\underline{P}_1$:

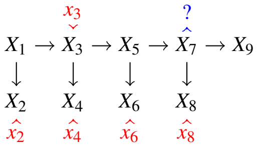
$$\underline{P}_1(f_1) = \underline{Q}_1(\underline{E}_1(f_1|X_1))$$

and the factorisation property of independent natural extension:

$$\underline{E}_1(f_1|X_1) = \underline{E}_1(I_{\{x_2\}} f_3|X_1) = \overline{\underline{Q}}_2(\{x_2\}|X_1) \odot \underline{P}_3(f_3|X_1)$$

Explaining the basics of the algorithm

An example



We have to calculate $\underline{P}_3(f_3|X_1)$, where

$$f_3 = I_{\{x_3\}}I_{\{x_4\}}I_{\{x_6\}}I_{\{x_8\}}[g(X_7)-\mu] = I_{\{x_3\}}I_{\{x_4\}}f_5$$

Now use the recursion formula for $\underline{P}_3(\cdot|X_1)$:

$$\underline{P}_3(f_3|X_1) = \underline{Q}_3(\underline{E}_3(f_3|X_3)|X_1)$$

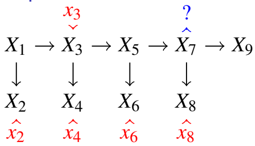
and the factorisation property of independent natural extension:

$$\underline{E}_3(f_3|X_3) = \underline{E}_3(I_{\{x_3\}}I_{\{x_4\}}f_5|X_3) = I_{\{x_3\}}\overline{Q}_4(\{x_4\}|x_3)\odot\underline{P}_5(f_5|x_3)$$

UNIVERSITEIT GENT

Explaining the basics of the algorithm

An example



In summary, so far we have:

$$\underline{P}_3(f_3|X_1) = \overline{\underline{Q}}_3(\{x_3\}|X_1) \odot \overline{\underline{Q}}_4(\{x_4\}|x_3) \odot \underline{P}_5(f_5|x_3)$$

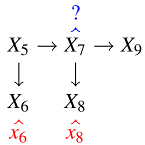
$$\underline{E}_1(f_1|X_1) = \overline{\underline{Q}}_2(\{x_2\}|X_1) \odot \overline{\underline{Q}}_3(\{x_3\}|X_1) \odot \overline{\underline{Q}}_4(\{x_4\}|x_3) \odot \underline{P}_5(f_5|x_3)$$

$$\underline{P}_1(f_1) = \begin{cases} \underline{Q}_1\left(\underline{Q}_2(\{x_2\}|X_1)\underline{Q}_3(\{x_3\}|X_1)\right)\underline{Q}_4(\{x_4\}|x_3)\underline{P}_5(f_5|x_3) & \underline{P}_5(f_5|x_3) \geq 0 \\ \overline{Q}_1\left(\overline{Q}_2(\{x_2\}|X_1)\overline{Q}_3(\{x_3\}|X_1)\right)\overline{Q}_4(\{x_4\}|x_3)\underline{P}_5(f_5|x_3) & \underline{P}_5(f_5|x_3) \leq 0 \end{cases}$$

$$= \overline{\underline{P}}_1(\{(x_2,x_3,x_4)\}) \odot \underline{P}_5(f_5|x_3)$$

Explaining the basics of the algorithm

An example



Because $\underline{P}_1(f_1) \geq 0 \Leftrightarrow \underline{P}_5(f_5|x_3) \geq 0$, we have to calculate $\underline{P}_5(f_5|x_3)$, where

$$f_5 = I_{\{x_6\}} I_{\{x_8\}}[g(X_7) - \mu] = I_{\{x_6\}} f_7$$

Now use the recursion formula for $\underline{P}_5(\cdot|x_3)$:

$$\underline{P}_5(f_5|x_3) = \underline{Q}_5(\underline{E}_5(f_5|X_5)|x_3)$$

and the factorisation property of independent natural extension:

$$\underline{E}_5(f_5|X_5) = \underline{E}_5(I_{\{x_6\}} f_7|X_5) = \underline{Q}_6(\{x_6\}|X_5) \odot \underline{P}_7(f_7|X_5)$$

Explaining the basics of the algorithm

An example

$$\begin{array}{ccccc} & & ? & & \\ & & \hat{} & & \\ X_5 & \to & X_7 & \to & X_9 \\ \downarrow & & \downarrow & & \\ X_6 & & X_8 & & \\ \hat{x}_6 & & \hat{x}_8 & & \end{array}$$

We have to calculate $\underline{P}_7(f_7|X_5)$, where

$$f_7 = I_{\{x_8\}}[g(X_7) - \mu]$$

Now use the recursion formula for $\underline{P}_7(\cdot|X_5)$:

$$\underline{P}_7(f_7|X_5) = \underline{Q}_7(\underline{E}_7(f_7|X_7)|X_5)$$

and apply independent natural extension:

$$\underline{E}_7(I_{\{x_8\}}[g(X_7) - \mu]|X_7) = \overline{\underline{Q}}_8(I_{\{x_8\}}|X_7) \odot [g(X_7) - \mu]$$

Explaining the basics of the algorithm

An example

$$
\begin{array}{ccccc}
 & & ? & & \\
 & & \wedge & & \\
X_5 & \rightarrow & X_7 & \rightarrow & X_9 \\
\downarrow & & \downarrow & & \\
X_6 & & X_8 & & \\
\wedge & & \wedge & & \\
x_6 & & x_8 & &
\end{array}
$$

In summary, we have that

$$\underline{P}_1(f_1) \geq 0 \Leftrightarrow \underline{P}_5(f_5|x_3) \geq 0$$

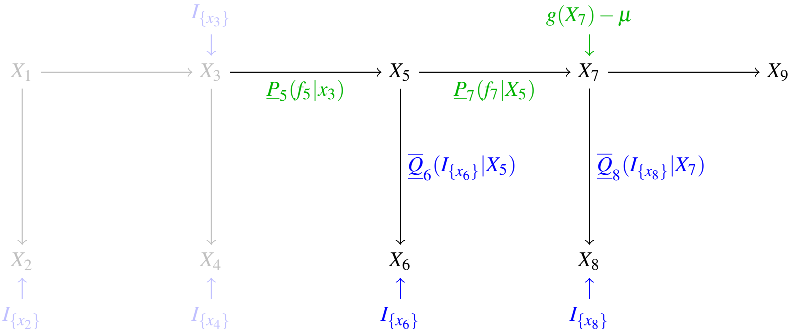
where

$$\underline{P}_5(f_5|x_3) = \underline{Q}_5(\underline{Q}_6(\{x_6\}|X_5) \odot \underline{P}_7(f_7|X_5)|x_3)$$

$$\underline{P}_7(f_7|X_5) = \underline{Q}_7(\overline{\underline{Q}}_8(I_{\{x_8\}}|X_7) \odot [g(X_7) - \mu]|X_5)$$

Explaining the basics of the algorithm

An example



$$\underline{P}_7(f_7|X_5) = \underline{Q}_7(\overline{Q}_8(I_{\{x_8\}}|X_7) \odot [g(X_7)-\mu]|X_5)$$

$$\underline{P}_5(f_5|x_3) = \underline{Q}_5(\underline{Q}_6(\{x_6\}|X_5) \odot \underline{P}_7(f_7|X_5)|x_3)$$

UNIVERSITEIT GENT

Conclusion

On the positive side

- very efficient, essentially linear in number of nodes
- not much more conservative than strong independence:
  - the same for forward inference
  - dilation for backward inference

On the negative side

- the algorithm works only for trees, not for polytrees or more general acyclic directed nets
- the algorithm works only for one function g at a time, it doesn't provide the entire credal set

UNIVERSITEIT GENT