Bayesian Networks with Imprecise Probabilities: Theory and Applications to Knowledge-based Systems and Classification

A Tutorial by

Alessandro Antonucci, Giorgio Corani and Denis Mauá

{alessandro,giorgio,denis}@idsia.ch

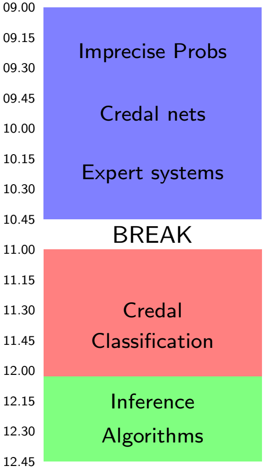
Istituto "Dalle Molle" di Studi sull'Intelligenza Artificiale - Lugano (Switzerland)

IJCAI-13

Beijing, August 5th, 2013

This morning

- From precise to imprecise probabilities
- Credal networks
- Applications to knowledge-based decision support systems
- Credal classifiers
- Application to data mining
- Inference algorithms on credal networks
- Discussion and general questions



09.00
09.15
09.30
09.45
10.00
10.15
10.30
10.45
11.00
11.15
11.30
11.45
12.00
12.15
12.30
12.45

Imprecise Probs

Credal nets

Expert systems

BREAK

Credal Classification

Inference Algorithms

Antonucci

Corani

giorgio

Mauá

denis

Just before starting . . .

Credal networks (i.e., Bayesian networks with imprecise probability) are drawing interest from the AI community in 2013

- ECSQARU 2013 Best Paper Award: *Approximating Credal Network Inferences by Linear Programming* by Alessandro Antonucci, Cassio de Campos, David Huber, and Marco Zaffalon
- UAI '13 Google Best Student Paper Award: *On the Complexity of Strong and Epistemic Credal Networks* by Denis Mauá, Cassio de Campos, Alessio Benavoli, and Alessandro Antonucci

More info and papers at `ipg.idsia.ch`

Outline (of the first part)

- A (first informal, then formal) introduction to IPs
    - Reasoning with (imprecise) fault trees
    - From determinism to imprecision (through uncertainty)
    - Motivations and coherence
- Credal sets
    - Basic concepts and operations
    - Modeling
- Credal networks
    - Background on Bayesian networks
    - From Bayesian to credal networks
    - Modeling (observations, missing data, information fusion, . . .)
- Applications to knowledge-based systems
    - Military decision making
    - Environmental risk analysis
    - (Imprecise) probabilistic description logic

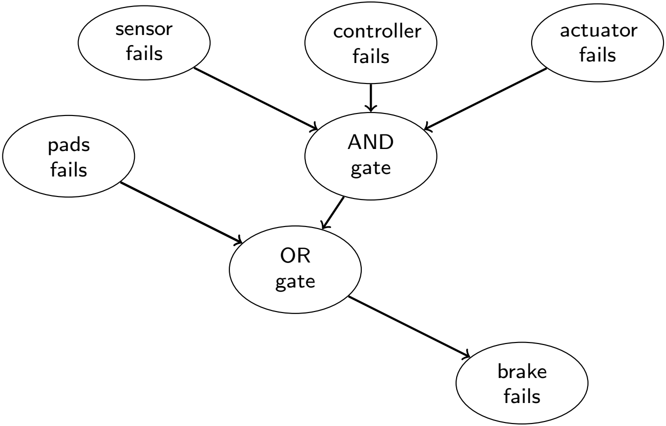
Reasoning: from Determinism to IPs

brake fails $= [$ pads $\vee$ (sensor $\wedge$ controller $\wedge$ actuator) $]$

devices failures are independent

Reasoning: from Determinism to IPs

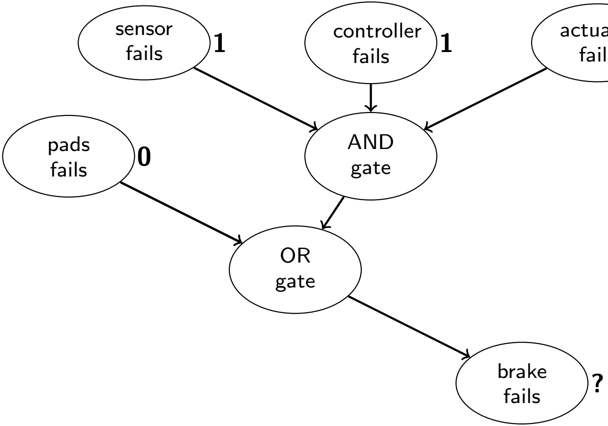
brake fails = [pads $\vee$ (sensor $\wedge$ controller $\wedge$ actuator)]



devices failures are independent

Reasoning: from Determinism to IPs

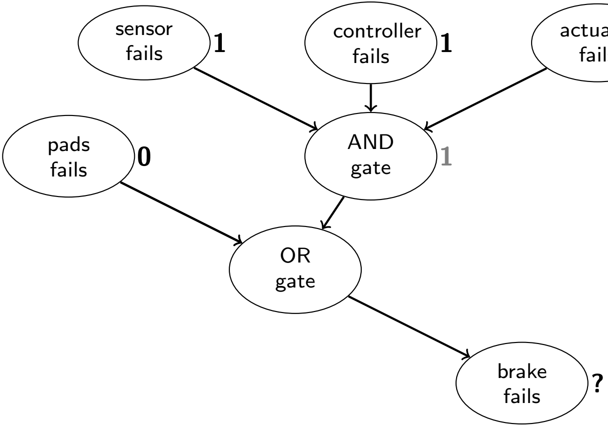
brake fails = [pads $\vee$ (sensor $\wedge$ controller $\wedge$ actuator)]



devices failures are independent

Reasoning: from Determinism to IPs

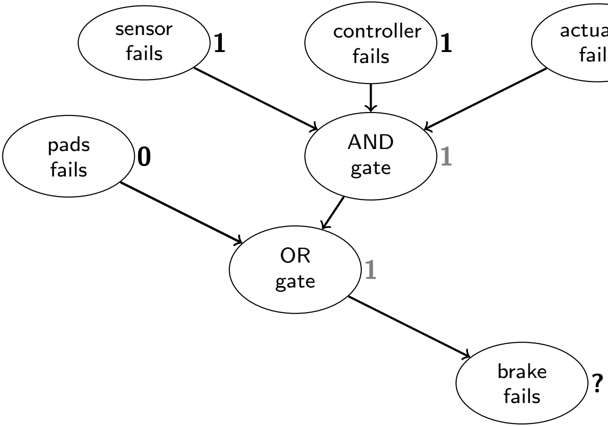
brake fails = [pads $\vee$ (sensor $\wedge$ controller $\wedge$ actuator)]



devices failures are independent

Reading: from Determinism to IPs

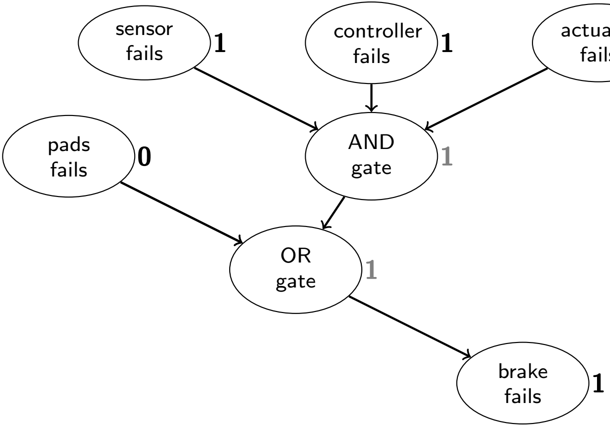
brake fails = [pads $\vee$ (sensor $\wedge$ controller $\wedge$ actuator)]



devices failures are independent

Reading: from Determinism to IPs

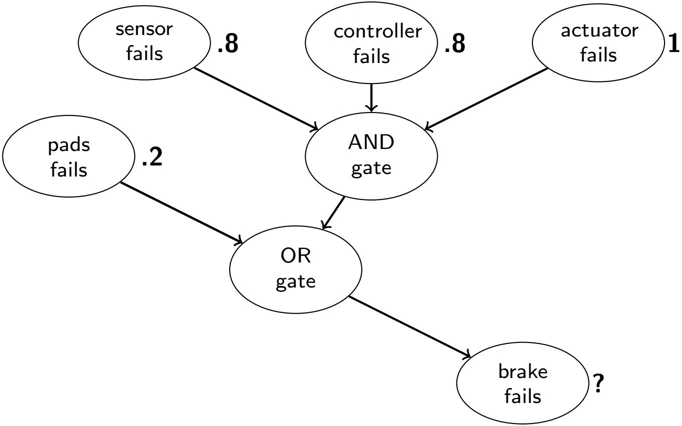
brake fails = [pads $\vee$ (sensor $\wedge$ controller $\wedge$ actuator)]



devices failures are independent

Reasoning: from Determinism to IPs

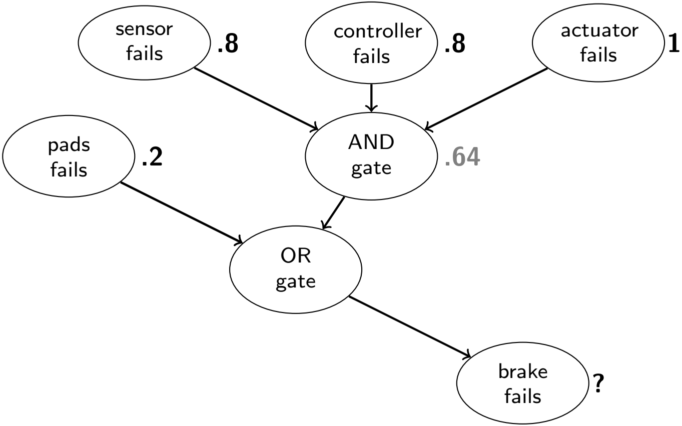
brake fails $=$ [pads $\vee$ (sensor $\wedge$ controller $\wedge$ actuator)]



devices failures are independent

Reasoning: from Determinism to IPs

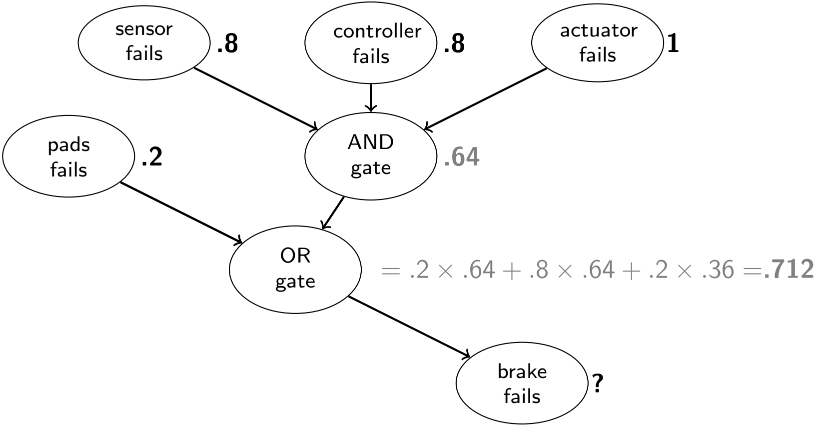
brake fails $=$ [pads $\vee$ (sensor $\wedge$ controller $\wedge$ actuator)]



devices failures are independent

Reasoning: from Determinism to IPs

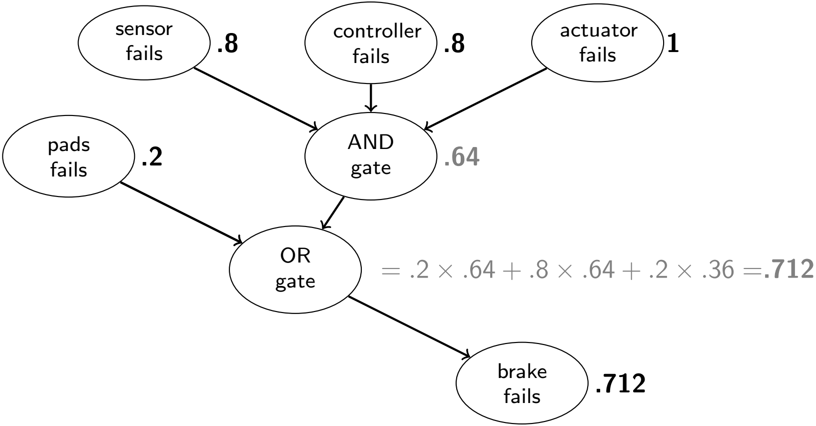
brake fails = [pads $\vee$ (sensor $\wedge$ controller $\wedge$ actuator)]



devices failures are independent

Reasoning: from Determinism to IPs

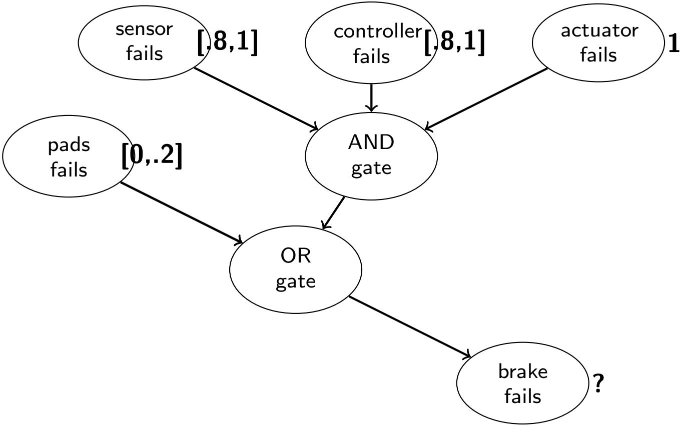
brake fails = [pads ∨ (sensor ∧ controller ∧ actuator)]



devices failures are independent

Reading: from Determinism to IPs

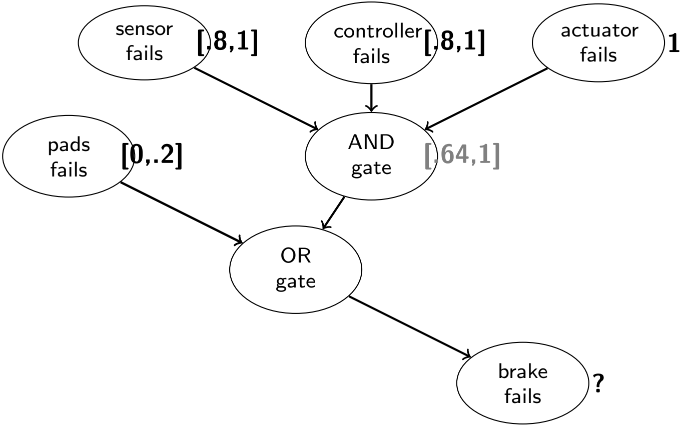
brake fails = [pads ∨ (sensor ∧ controller ∧ actuator)]



devices failures are independent

Reasoning: from Determinism to IPs

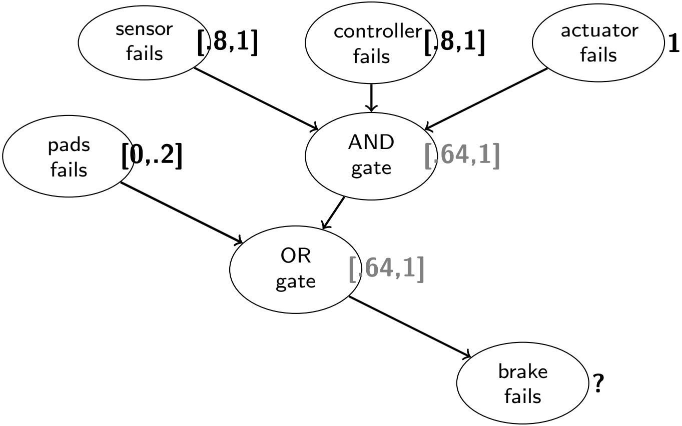
brake fails = [pads ∨ (sensor ∧ controller ∧ actuator)]



devices failures are independent

Reasoning: from Determinism to IPs

brake fails $= [$ pads $\vee$ (sensor $\wedge$ controller $\wedge$ actuator) $]$



sensor fails **[.8,1]**

controller fails **[.8,1]**

actuator fails **1**

pads fails **[0,.2]**

AND gate **[.64,1]**

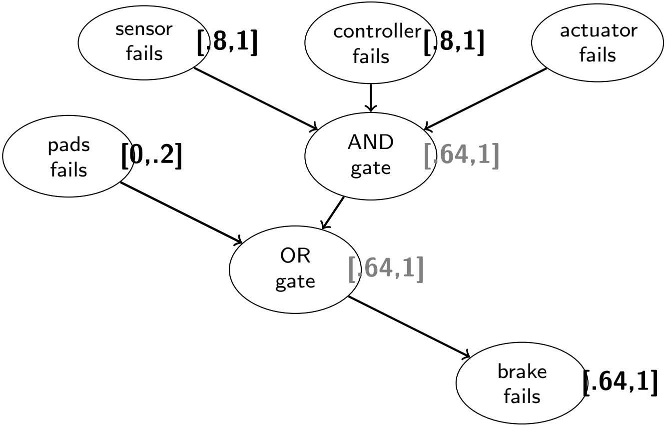
OR gate **[.64,1]**

brake fails **?**

devices failures are independent

Reasoning: from Determinism to IPs

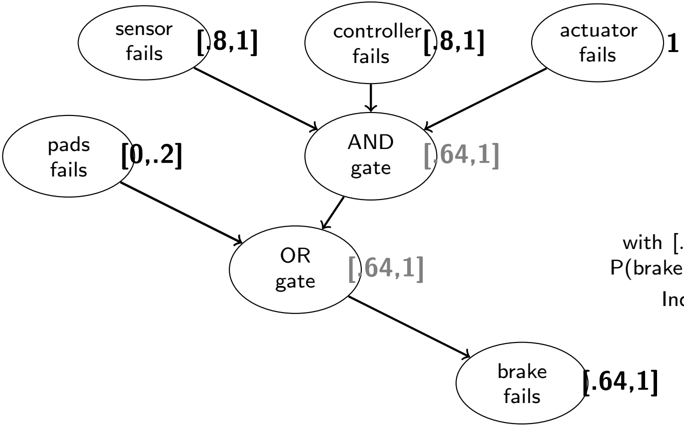
brake fails = [pads ∨ (sensor ∧ controller ∧ actuator)]



devices failures are independent

Reasoning: from Determinism to IPs

$$\text{brake fails} = [\ \text{pads} \vee (\ \text{sensor} \wedge \text{controller} \wedge \text{actuator}\ )\ ]$$



with $[.7, 1]$ instead
P(brake fails)∈[.49,1]
Indecision!

devices failures are independent

Three different levels of knowledge

- A football match between Italy and Spain
- Result of Spain after the regular time? Win, draw or loss?

DETERMINISM

The Spanish goalkeeper is unbeatable and Italy always receives a goal

Spain (certainly) wins

$$\begin{array}{l} P(Win) \\ P(Draw) \\ P(Loss) \end{array} = \begin{bmatrix} 1 \\ 0 \\ 0 \end{bmatrix}$$

UNCERTAINTY

Win is two times more probable than draw, and this being three times more probable than loss

$$\begin{array}{l} P(Win) \\ P(Draw) \\ P(Loss) \end{array} = \begin{bmatrix} .6 \\ .3 \\ .1 \end{bmatrix}$$

IMPRECISION

Win is more probable than draw, and this is more probable than loss

$P(Win) > P(Draw)$
$P(Draw) > P(Loss)$

$$\begin{array}{l} P(Win) \\ P(Draw) \\ P(Loss) \end{array} = \begin{bmatrix} \frac{\alpha}{3} + \beta + \frac{\gamma}{2} \\ \frac{\alpha}{3} + \frac{\gamma}{2} \\ \frac{\alpha}{3} \end{bmatrix}$$

$\forall \alpha, \beta, \gamma$ such that
$\alpha > 0, \beta > 0, \gamma > 0,$
$\alpha + \beta + \gamma = 1$

Three different levels of knowledge

- A football match between Italy and Spain
- Result of Spain after the regular time? Win, draw or loss?

DETERMINISM

The Spanish goalkeeper is unbeatable and Italy always receives a goal

Spain (certainly) wins

$$\begin{array}{l} P(Win) \\ P(Draw) \\ P(Loss) \end{array} = \begin{bmatrix} 1 \\ 0 \\ 0 \end{bmatrix}$$

UNCERTAINTY

Win is two times more probable than draw, and this being three times more probable than loss

$$\begin{array}{l} P(Win) \\ P(Draw) \\ P(Loss) \end{array} = \begin{bmatrix} .6 \\ .3 \\ .1 \end{bmatrix}$$

IMPRECISION

Win is more probable than draw, and this is more probable than loss

$$P(Win) > P(Draw)$$
$$P(Draw) > P(Loss)$$

$$\begin{array}{l} P(Win) \\ P(Draw) \\ P(Loss) \end{array} = \begin{bmatrix} \frac{\alpha}{3} + \beta + \frac{\gamma}{2} \\ \frac{\alpha}{3} + \frac{\gamma}{2} \\ \frac{\alpha}{3} \end{bmatrix}$$

$\forall \alpha, \beta, \gamma$ such that
$\alpha > 0, \beta > 0, \gamma > 0,$
$\alpha + \beta + \gamma = 1$

Three different levels of knowledge

- A football match between Italy and Spain
- Result of Spain after the regular time? Win, draw or loss?

DETERMINISM

The Spanish goalkeeper is unbeatable and Italy always receives a goal

Spain (certainly) wins

$$\begin{array}{l} P(Win) \\ P(Draw) \\ P(Loss) \end{array} = \begin{bmatrix} 1 \\ 0 \\ 0 \end{bmatrix}$$

UNCERTAINTY

Win is two times more probable than draw, and this being three times more probable than loss

$$\begin{array}{l} P(Win) \\ P(Draw) \\ P(Loss) \end{array} = \begin{bmatrix} .6 \\ .3 \\ .1 \end{bmatrix}$$

IMPRECISION

Win is more probable than draw, and this is more probable than loss

$P(Win) > P(Draw)$
$P(Draw) > P(Loss)$

$$\begin{array}{l} P(Win) \\ P(Draw) \\ P(Loss) \end{array} = \begin{bmatrix} \frac{\alpha}{3} + \beta + \frac{\gamma}{2} \\ \frac{\alpha}{3} + \frac{\gamma}{2} \\ \frac{\alpha}{3} \end{bmatrix}$$

$\forall \alpha, \beta, \gamma$ such that
$\alpha > 0, \beta > 0, \gamma > 0,$
$\alpha + \beta + \gamma = 1$

Three different levels of knowledge

- A football match between Italy and Spain
- Result of Spain after the regular time? Win, draw or loss?

DETERMINISM

The Spanish goalkeeper is unbeatable and Italy always receives a goal

Spain (certainly) wins

$$\begin{matrix} P(Win) \\ P(Draw) \\ P(Loss) \end{matrix} = \begin{bmatrix} 1 \\ 0 \\ 0 \end{bmatrix}$$

UNCERTAINTY

Win is two times more probable than draw, and this being three times more probable than loss

$$\begin{matrix} P(Win) \\ P(Draw) \\ P(Loss) \end{matrix} = \begin{bmatrix} .6 \\ .3 \\ .1 \end{bmatrix}$$

IMPRECISION

Win is more probable than draw, and this is more probable than loss

$P(Win) > P(Draw)$
$P(Draw) > P(Loss)$

$$\begin{matrix} P(Win) \\ P(Draw) \\ P(Loss) \end{matrix} = \begin{bmatrix} \frac{\alpha}{3} + \beta + \frac{\gamma}{2} \\ \frac{\alpha}{3} + \frac{\gamma}{2} \\ \frac{\alpha}{3} \end{bmatrix}$$

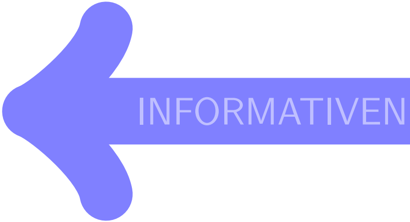
$\forall \alpha, \beta, \gamma$ such that
$\alpha > 0, \beta > 0, \gamma > 0,$
$\alpha + \beta + \gamma = 1$

Three different levels of knowledge

DETERMINISM UNCERTAINT

Three different levels of knowledge

DETERMINISM UNCERTAINT



INFORMATIVEN

Three different levels of knowledge

DETERMINISM UNCERTAINT

EXPRESSIVENE

Three different levels of knowledge

DETERMINISM

UNCERTAINTY

Three different levels of knowledge

DETERMINISM

UNCERTAINT

limit of
of avail
(e.g., ver

Three different levels of knowledge

DETERMINISM

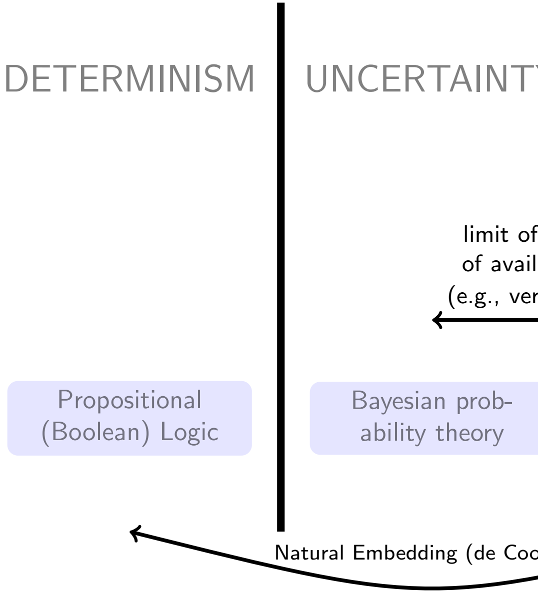
UNCERTAINT

limit of
of avail
(e.g., ver

Propositional (Boolean) Logic

Bayesian probability theory

Three different levels of knowledge



DETERMINISM

UNCERTAINT

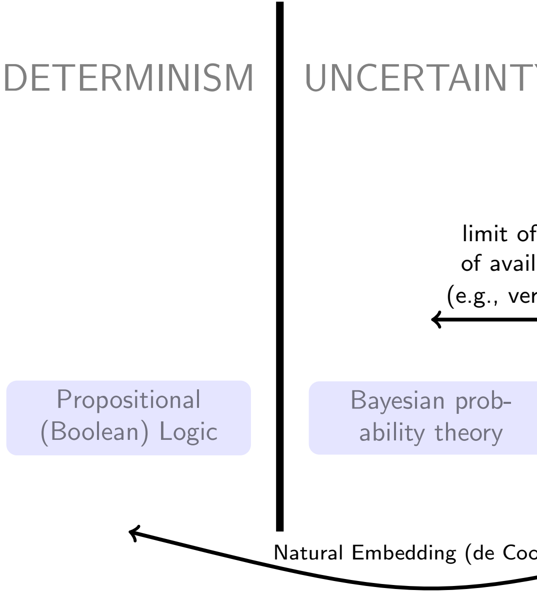
limit of
of avail
(e.g., ver

Propositional (Boolean) Logic

Bayesian probability theory

Natural Embedding (de Coo

Three different levels of knowledge



DETERMINISM

UNCERTAINT

limit of
of avail
(e.g., ver

Propositional (Boolean) Logic

Bayesian probability theory

Natural Embedding (de Coo

[. . .] Bayesian inference will always be a basic tool for practical everyday statistics,

if only because questions must be answered and decisions must be taken, so that a statistician must always stand ready to upgrade his vaguer forms of belief into precisely additive probabilities

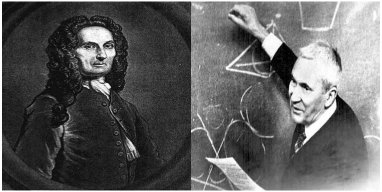
Art Dempster (in his foreword to Shafer's book)

Probability: one word for two (not exclusive) things

Randomness

Variability captured through repeated observations

De Moivre and Kolmogorov



- Chances
- Feature of the world
- Aleatory or objective
- Frequentist theory
- Limiting frequencies

Partial knowledge

Incomplete information about issues of interest

Bayes and De Finetti

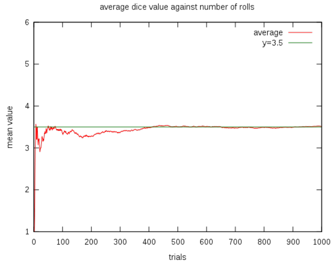
- Beliefs
- Feature of the observer
- Epistemic or subjective
- Bayesian theory
- Behaviour (bets dispositions)

Objective probability

- X taking its values in (finite set) Ω
- Value $X = x \in \Omega$ as the output of an experiment which can be iterated
- Prob $P(x)$ as limiting frequency

$$P(x) := \lim_{N \to +\infty} \frac{\#(X = x)}{N}$$

- Kolmogorov's axioms follow from this
- Probability as a property of the world
- Not only (statistical and quantum) mechanics, hazard games (coins, dices, cards), but also economics, bio/psycho/sociology, linguistics, etc.
- But not all events can be iterated ...



1. $\forall A \in 2^{\Omega},\ 0 \leq P(A) \leq 1$
2. $P(\Omega) = 1$
3. $\forall A, B \in 2^{\Omega} : A \wedge B = \emptyset$
   $P(A \vee B) = P(A) + P(B)$

Probability in everyday life

Bersani: Il voto a primavera? "Ci sono buone probabilità"

Il segretario del Pd rilancia il progetto del Nuovo Ulivo e offre una proposta di alleanza di governo rivolta al cosiddetto Terzo Polo. Poi sdrammatizza con una battuta le voci sul futuro leader del partito e rilancia le primarie: "Una cosa bellissima"

10 ottobre, 2010

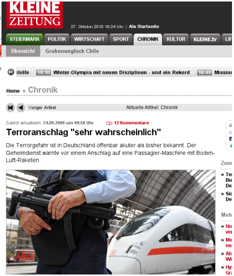
Il segretario del Pd Pier Luigi Bersani

Chile: probabilidades de encontrar mineros con vida 'son bajas'

Publicado el 12/Agosto/2010 | 09:10

SANTIAGO. **Las probabilidades de encontrar con vida a los 33 mineros atrapados desde hace una semana en un mina en el norte de Chile "son bajas"**, admitió el ministro de Minería, Laurance Golborne, que encabeza las labores de búsqueda.

"Las probabilidades de encontrarlos con vida son bajas", señaló el ministro en una entrevista con Canal 13, siete días después que un derrumbe sepultó a 33 mineros en el yacimiento San José, en las cercanías de Copiapó, a unos 800 km al norte de Santiago.



Zuletzt aktualisiert: **24.09.2009 um 09:58 Uhr**

Terroranschlag "sehr wahrscheinlich"

Die Terrorgefahr ist in Deutschland offenbar akuter als bisher bekannt. Der Geheimdienst warnte vor einem Anschlag auf eine Passagier-Maschine mit Boden-Luft-Raketen.

Xi Jinping, très probable successeur de Hu Jintao et "fils de prince"

AFP | 18.10.10 | 10h36

Xi Jinping, 57 ans, qui vient d'être nommé vice-président de la puissante Commission militaire

SPORTS

Wednesday October 20, 2010

Anderson and Sullivan both will 'probably' play for Herd

by Chuck McGill
Daily Mail sports writer

HUNTINGTON, W.Va. -- Regarding the quarterback

Economia

02/02/2010 às 16h05 - Atualizada em 25/10/2011 às 00h09

Fecombustíveis alerta para provável aumento de gasolina

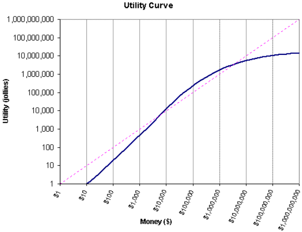
Probabilities often pertains to singular events

not necessarily related to statistics

Subjective probability

- Probability p of me smoking
- Singular event: frequency unavailable
- Subjective probability
  - models (partial) knowledge of a subject
  - feature of the subject not of the world
  - two subjects can assess different probs
- Quantitative measure of knowledge?
  - Behavioural approach
  - Subjective betting dispositions
  - A (linear) utility function is needed

- Money?
- Big money not linear!
- Small, somehow yes



lottery tickets

$\propto$

winning chance

$\propto$

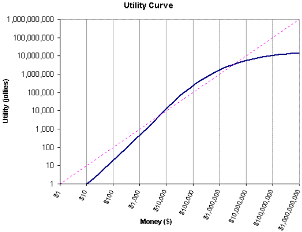
benefit

infinite number of tickets
makes utility real-valued

Subjective probability

- Probability p of me smoking
- Singular event: frequency unavailable
- Subjective probability
  - models (partial) knowledge of a subject
  - feature of the subject not of the world
  - two subjects can assess different probs
- Quantitative measure of knowledge?
  - Behavioural approach
  - Subjective betting dispositions
  - A (linear) utility function is needed

- Money?
- Big money not linear!
- Small, somehow yes



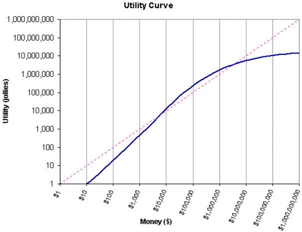
lottery tickets
$\propto$
winning chance
$\propto$
benefit

infinite number of tickets
makes utility real-valued

Subjective probability

- Probability *p* of me smoking
- Singular event: frequency unavailable
- Subjective probability
  - models (partial) knowledge of a subject
  - feature of the subject not of the world
  - two subjects can assess different probs
- Quantitative measure of knowledge?
  - Behavioural approach
  - Subjective betting dispositions
  - A (linear) utility function is needed

- Money?
- Big money not linear!
- Small, somehow yes



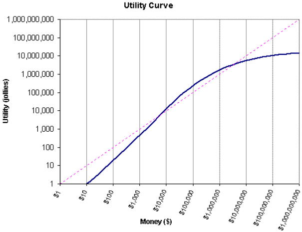
lottery tickets
$\propto$
winning chance
$\propto$
benefit

infinite number of tickets
makes utility real-valued

Subjective probability

- Probability p of me smoking
- Singular event: frequency unavailable
- Subjective probability
  - models (partial) knowledge of a subject
  - feature of the subject not of the world
  - two subjects can assess different probs
- Quantitative measure of knowledge?
  - Behavioural approach
  - Subjective betting dispositions
  - A (linear) utility function is needed

- Money?
- Big money not linear!
- Small, somehow yes



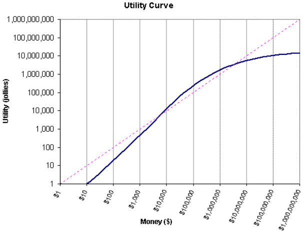
lottery tickets
$\propto$
winning chance
$\propto$
benefit

infinite number of tickets
makes utility real-valued

Subjective probability

- Probability *p* of me smoking
- Singular event: frequency unavailable
- Subjective probability
  - models (partial) knowledge of a subject
  - feature of the subject not of the world
  - two subjects can assess different probs
- Quantitative measure of knowledge?
  - Behavioural approach
  - Subjective betting dispositions
  - A (linear) utility function is needed

- Money?
- Big money not linear!
- Small, somehow yes



$$\begin{gathered}\text{lottery tickets}\\ \propto\\ \text{winning chance}\\ \propto\\ \text{benefit}\end{gathered}$$

infinite number of tickets makes utility real-valued

(Rationally) betting on gambles

- Probabilities as dispositions to buy/sell gambles
- Gambles as checks whose amount is uncertain/unknown

This check has a value of 100 EUR
if Alessandro is a smoker
zero otherwise

- The bookie sells this gamble
- Probability *p* as a *price* for the gamble
  - $\frac{\text{maximum price}}{\text{100EUR}}$ for which you buy the gamble
  - $\frac{\text{minimum price}}{\text{100EUR}}$ for which you (bookie) sell it
- Interpretation + rationality produce axioms

(Rationally) betting on gambles

- Probabilities as dispositions to buy/sell gambles
- Gambles as checks whose amount is uncertain/unknown

This check has a value of 100 EUR
if Alessandro is a smoker
zero otherwise

- The bookie sells this gamble
- Probability *p* as a *price* for the gamble
  - $\frac{\text{maximum price}}{\text{100EUR}}$ for which you buy the gamble
  - $\frac{\text{minimum price}}{\text{100EUR}}$ for which you (bookie) sell it
- Interpretation + rationality produce axioms

(Rationally) betting on gambles

- Probabilities as dispositions to buy/sell gambles
- Gambles as checks whose amount is uncertain/unknown

This check has a value of 100 EUR
if Alessandro is a smoker
zero otherwise

- The bookie sells this gamble
- Probability p as a *price* for the gamble
  - $\frac{\text{maximum price}}{100\text{EUR}}$ for which you buy the gamble
  - $\frac{\text{minimum price}}{100\text{EUR}}$ for which you (bookie) sell it
- Interpretation + rationality produce axioms

(Rationally) betting on gambles

- Probabilities as dispositions to buy/sell gambles
- Gambles as checks whose amount is uncertain/unknown

This check has a value of 100 EUR
if Alessandro is a smoker
zero otherwise

- The bookie sells this gamble
- Probability *p* as a *price* for the gamble
  - $\frac{\text{maximum price}}{100\text{EUR}}$ for which you buy the gamble
  - $\frac{\text{minimum price}}{100\text{EUR}}$ for which you (bookie) sell it
- Interpretation + rationality produce axioms

(Rationally) betting on gambles

- Probabilities as dispositions to buy/sell gambles
- Gambles as checks whose amount is uncertain/unknown

This check has a value of 100 EUR
if Alessandro is a smoker
zero otherwise

- The bookie sells this gamble
- Probability *p* as a *price* for the gamble
  - $\frac{\text{maximum price}}{100\text{EUR}}$ for which you buy the gamble
  - $\frac{\text{minimum price}}{100\text{EUR}}$ for which you (bookie) sell it
- Interpretation + rationality produce axioms

(Rationally) betting on gambles

- Probabilities as dispositions to buy/sell gambles
- Gambles as checks whose amount is uncertain/unknown

This check has a value of 100 EUR
if Alessandro is a smoker
zero otherwise

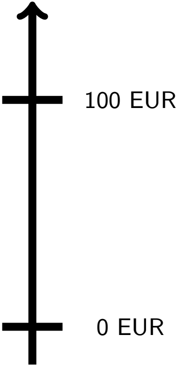
- The bookie sells this gamble
- Probability p as a *price* for the gamble
  - $\frac{\text{maximum price}}{100\text{EUR}}$ for which you buy the gamble
  - $\frac{\text{minimum price}}{100\text{EUR}}$ for which you (bookie) sell it
- Interpretation + rationality produce axioms

(Rationally) betting on gambles

- Probabilities as dispositions to buy/sell gambles
- Gambles as checks whose amount is uncertain/unknown

> This check has a value of 100 EUR
> if Alessandro is a smoker
> zero otherwise

- The bookie sells this gamble
- Probability *p* as a *price* for the gamble
  - $\frac{\text{maximum price}}{\text{100EUR}}$ for which you buy the gamble
  - $\frac{\text{minimum price}}{\text{100EUR}}$ for which you (bookie) sell it
- Interpretation + rationality produce axioms



(Rationally) betting on gambles

- Probabilities as dispositions to buy/sell gambles
- Gambles as checks whose amount is uncertain/unknown

> This check has a value of 100 EUR
> if Alessandro is a smoker
> zero otherwise

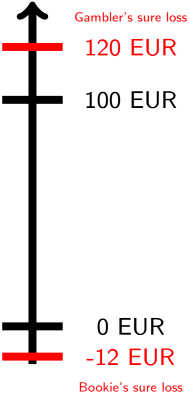
- The bookie sells this gamble
- Probability *p* as a *price* for the gamble
  - $\frac{\text{maximum price}}{\text{100EUR}}$ for which you buy the gamble
  - $\frac{\text{minimum price}}{\text{100EUR}}$ for which you (bookie) sell it
- Interpretation + rationality produce axioms

(Rationally) betting on gambles

- Probabilities as dispositions to buy/sell gambles
- Gambles as checks whose amount is uncertain/unknown

> This check has a value of 100 EUR
> if Alessandro is a smoker
> zero otherwise

- The bookie sells this gamble
- Probability *p* as a *price* for the gamble
  - $\frac{\text{maximum price}}{\text{100EUR}}$ for which you buy the gamble
  - $\frac{\text{minimum price}}{\text{100EUR}}$ for which you (bookie) sell it
- Interpretation + rationality produce axioms



Gambler's sure loss

120 EUR

100 EUR

0 EUR

-12 EUR

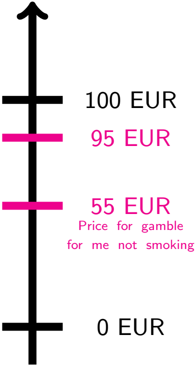
Bookie's sure loss

(Rationally) betting on gambles

- Probabilities as dispositions to buy/sell gambles
- Gambles as checks whose amount is uncertain/unknown

This check has a value of 100 EUR
if Alessandro is a smoker
zero otherwise

- The bookie sells this gamble
- Probability *p* as a *price* for the gamble
  - $\frac{\text{maximum price}}{100\text{EUR}}$ for which you buy the gamble
  - $\frac{\text{minimum price}}{100\text{EUR}}$ for which you (bookie) sell it
- Interpretation + rationality produce axioms

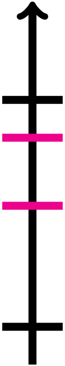


(Rationally) betting on gambles

- Probabilities as dispositions to buy/sell gambles
- Gambles as checks whose amount is uncertain/unknown

> This check has a value of 100 EUR
> if Alessandro is a smoker
> zero otherwise

- The bookie sells this gamble
- Probability *p* as a *price* for the gamble
  - $\frac{\text{maximum price}}{\text{100EUR}}$ for which you buy the gamble
  - $\frac{\text{minimum price}}{\text{100EUR}}$ for which you (bookie) sell it
- Interpretation + rationality produce axioms



100 EUR

95 EUR

55 EUR

Price for gamble for me not smoking

0 EUR

You spend 150 EUR to (certainly) win 100 EUR!

Coherence and linear previsions

Don't be crazy: choose prices s.t.
there is always a chance to win
(whatever the stakes set by the bookie)

Prices $\{P_{A_i}\}_{i=1}^N$ for events $A_i \subseteq \Omega$, $i = 1, \ldots, N$ are coherent iff

$$\max_{x\in\Omega} \sum_{i=1}^{N} c_i [I_{A_i}(x) - P_{A_i}] \geq 0$$

Moreover, assessments $\{P_{A_i}\}_{i=1}^N$ are coherent iff

- Exists probability mass function $P(X)$: $P(A_i) = P_{A_i}$
- Or, for general gambles, linear functional $P(f_i) := P_{f_i}$

$$P(f) = \sum_{x\in\Omega} P(x)\, f(x)$$

linear prevision
to be extended to a
coherent lower prevision

probability mass function
to be extended
to a credal set

Coherence and linear previsions

Don't be crazy: choose prices s.t.
there is always a chance to win
(whatever the stakes set by the bookie)

Prices $\{P_{A_i}\}_{i=1}^N$ for events $A_i \subseteq \Omega$, $i = 1, \ldots, N$ are coherent iff

$$\max_{x\in\Omega} \sum_{i=1}^{N} c_i[I_{A_i}(x) - P_{A_i}] \geq 0$$

Moreover, assessments $\{P_{A_i}\}_{i=1}^N$ are coherent iff

- Exists probability mass function $P(X)$: $P(A_i) = P_{A_i}$
- Or, for general gambles, linear functional $P(f_i) := P_{f_i}$

$$\longrightarrow P(f) = \sum_{x\in\Omega} P(x) \leftarrow f(x)$$

linear prevision
to be extended to a
coherent lower prevision

probability mass function
to be extended
to a credal set

Coherence and linear previsions

Don't be crazy: choose prices s.t.
there is always a chance to win
(whatever the stakes set by the bookie)

Prices $\{P_{A_i}\}_{i=1}^N$ for events $A_i \subseteq \Omega$, $i = 1, \ldots, N$ are **coherent** iff

$$\max_{x\in\Omega} \sum_{i=1}^{N} c_i [I_{A_i}(x) - P_{A_i}] \geq 0$$

Moreover, assessments $\{P_{A_i}\}_{i=1}^N$ are coherent iff

- Exists **probability mass function** $P(X)$: $P(A_i) = P_{A_i}$
- Or, for general gambles, **linear functional** $P(f_i) := P_{f_i}$

$$P(f) = \sum_{x\in\Omega} P(x)\, f(x)$$

linear prevision
to be extended to a
coherent lower prevision

probability mass function
to be extended
to a credal set

Coherence and linear previsions

Don't be crazy: choose prices s.t.
there is always a chance to win
(whatever the stakes set by the bookie)

Prices $\{P_{A_i}\}_{i=1}^N$ for events $A_i \subseteq \Omega$, $i = 1, \dots, N$ are coherent iff

$$\max_{x\in\Omega} \sum_{i=1}^{N} c_i [I_{A_i}(x) - P_{A_i}] \geq 0$$

Moreover, assessments $\{P_{A_i}\}_{i=1}^N$ are coherent iff

- Exists probability mass function $P(X)$: $P(A_i) = P_{A_i}$
- Or, for general gambles, linear functional $P(f_i) := P_{f_i}$

$$P(f) = \sum_{x\in\Omega} P(x)\, f(x)$$

linear prevision
to be extended to a
coherent lower prevision

probability mass function
to be extended
to a credal set

Coherence and linear previsions

Don't be crazy: choose prices s.t.
there is always a chance to win
(whatever the stakes set by the bookie)

Prices $\{P_{A_i}\}_{i=1}^{N}$ for events $A_i \subseteq \Omega$, $i = 1, \ldots, N$ are coherent iff

$$\max_{x\in\Omega} \sum_{i=1}^{N} c_i [I_{A_i}(x) - P_{A_i}] \geq 0$$

Moreover, assessments $\{P_{A_i}\}_{i=1}^{N}$ are coherent iff

- Exists probability mass function $P(X)$: $P(A_i) = P_{A_i}$
- Or, for general gambles, linear functional $P(f_i) := P_{f_i}$

$$P(f) = \sum_{x\in\Omega} P(x)\, f(x)$$

linear prevision
to be extended to a
coherent lower prevision

probability mass function
to be extended
to a credal set

Coherence and linear previsions

Don't be crazy: choose prices s.t.
there is always a chance to win
(whatever the stakes set by the bookie)

Prices $\{P_{A_i}\}_{i=1}^N$ for events $A_i \subseteq \Omega$, $i = 1, \ldots, N$ are coherent iff

$$\max_{x \in \Omega} \sum_{i=1}^{N} c_i [I_{A_i}(x) - P_{A_i}] \geq 0$$

Moreover, assessments $\{P_{A_i}\}_{i=1}^N$ are coherent iff

- Exists probability mass function $P(X)$: $P(A_i) = P_{A_i}$
- Or, for general gambles, linear functional $P(f_i) := P_{f_i}$

$$\rightarrow P(f) = \textstyle\sum_{x \in \Omega} P(x) \leftarrow f(x)$$

linear prevision
to be extended to a
coherent lower prevision

probability mass function
to be extended
to a credal set

Coherence and linear previsions

Don't be crazy: choose prices s.t.
there is always a chance to win
(whatever the stakes set by the bookie)

Prices $\{P_{A_i}\}_{i=1}^{N}$ for events $A_i \subseteq \Omega$, $i = 1, \ldots, N$ are coherent iff

$$\max_{x\in\Omega} \sum_{i=1}^{N} c_i [I_{A_i}(x) - P_{A_i}] \geq 0$$

Moreover, assessments $\{P_{A_i}\}_{i=1}^{N}$ are coherent iff

- Exists probability mass function $P(X)$: $P(A_i) = P_{A_i}$
- Or, for general gambles, linear functional $P(f_i) := P_{f_i}$

$$P(f) = \textstyle\sum_{x\in\Omega} P(x)\, f(x)$$

linear prevision

to be extended to a
coherent lower prevision

probability mass function

to be extended
to a credal set

Coherence and linear previsions

Don't be crazy: choose prices s.t.
there is always a chance to win
(whatever the stakes set by the bookie)

Prices $\{P_{A_i}\}_{i=1}^N$ for events $A_i \subseteq \Omega$, $i = 1, \ldots, N$ are coherent iff

$$\max_{x\in\Omega} \sum_{i=1}^{N} c_i [I_{A_i}(x) - P_{A_i}] \geq 0$$

Moreover, assessments $\{P_{A_i}\}_{i=1}^N$ are coherent iff

- Exists probability mass function $P(X)$: $P(A_i) = P_{A_i}$
- Or, for general gambles, linear functional $P(f_i) := P_{f_i}$

$$P(f) = \textstyle\sum_{x\in\Omega} P(x)\, f(x)$$

linear prevision
to be extended to a
coherent lower prevision

probability mass function
to be extended
to a credal set

(subjective, behavioural) imprecise probabilities

De Finetti's precision dogma

$\overline{P}(x)$		$\underline{P}(x)$
minimum selling price	$=$	maximum buying price

Walley's proposal for imprecision

No strong reasons for that rationality only requires
$\underline{P}(x) \leq \overline{P}(x)$

- Avoid sure loss! With max buying prices $\underline{P}(A)$ and $\underline{P}(A^c)$, you can buy both gambles and earn one for sure:

$$\underline{P}(A) + \underline{P}(A^c) \leq 1$$

- Be coherent! When buying both A and B, you pay $\underline{P}(A) + \underline{P}(B)$ and you have a gamble which gives one if $A \cup B$ occurs:

$$\underline{P}(A \cup B) \geq \underline{P}(A) + \underline{P}(B)$$

coherence self-consistency (beliefs revised if unsatisfied)
less problematic than a.s.l.

(subjective, behavioural) imprecise probabilities

De Finetti's precision dogma

$\overline{P}(x)$		$\underline{P}(x)$
minimum selling price	$\equiv$	maximum buying price

Walley's proposal for imprecision

No strong reasons for that rationality only requires $\underline{P}(x) \leq \overline{P}(x)$

- Avoid sure loss! With max buying prices $\underline{P}(A)$ and $\underline{P}(A^c)$, you can buy both gambles and earn one for sure:

$$\underline{P}(A) + \underline{P}(A^c) \leq 1$$

- Be coherent! When buying both A and B, you pay $\underline{P}(A) + \underline{P}(B)$ and you have a gamble which gives one if $A \cup B$ occurs:

$$\underline{P}(A \cup B) \geq \underline{P}(A) + \underline{P}(B)$$

coherence self-consistency (beliefs revised if unsatisfied)
less problematic than a.s.l.

(subjective, behavioural) imprecise probabilities

De Finetti's precision dogma

$\overline{P}(x)$		$\underline{P}(x)$
minimum selling price	$\equiv$	maximum buying price

Walley's proposal for imprecision

No strong reasons for that rationality only requires
$\underline{P}(x) \leq \overline{P}(x)$

- Avoid sure loss! With max buying prices $\underline{P}(A)$ and $\underline{P}(A^c)$, you can buy both gambles and earn one for sure:

$$\underline{P}(A) + \underline{P}(A^c) \leq 1$$

- Be coherent! When buying both A and B, you pay $\underline{P}(A) + \underline{P}(B)$ and you have a gamble which gives one if $A \cup B$ occurs:

$$\underline{P}(A \cup B) \geq \underline{P}(A) + \underline{P}(B)$$

coherence self-consistency (beliefs revised if unsatisfied)
less problematic than a.s.l.

(subjective, behavioural) imprecise probabilities

De Finetti's precision dogma

$\overline{P}(x)$		$\underline{P}(x)$
minimum selling price	$\equiv$	maximum buying price

Walley's proposal for imprecision

No strong reasons for that rationality only requires $\underline{P}(x) \leq \overline{P}(x)$

- Avoid sure loss! With max buying prices $\underline{P}(A)$ and $\underline{P}(A^c)$, you can buy both gambles and earn one for sure:

$$\underline{P}(A) + \underline{P}(A^c) \leq 1$$

- Be coherent! When buying both A and B, you pay $\underline{P}(A) + \underline{P}(B)$ and you have a gamble which gives one if $A \cup B$ occurs:

$$\underline{P}(A \cup B) \geq \underline{P}(A) + \underline{P}(B)$$

coherence self-consistency (beliefs revised if unsatisfied)
less problematic than a.s.l.

(subjective, behavioural) imprecise probabilities

De Finetti's precision dogma

$\overline{P}(x)$ minimum selling price $\equiv$ $\underline{P}(x)$ maximum buying price

Walley's proposal for imprecision

No strong reasons for that rationality only requires $\underline{P}(x) \leq \overline{P}(x)$

- Avoid sure loss! With max buying prices $\underline{P}(A)$ and $\underline{P}(A^c)$, you can buy both gambles and earn one for sure:

$$\underline{P}(A) + \underline{P}(A^c) \leq 1$$

- Be coherent! When buying both A and B, you pay $\underline{P}(A) + \underline{P}(B)$ and you have a gamble which gives one if $A \cup B$ occurs:

$$\underline{P}(A \cup B) \geq \underline{P}(A) + \underline{P}(B)$$

coherence self-consistency (beliefs revised if unsatisfied)
less problematic than a.s.l.

(subjective, behavioural) imprecise probabilities

De Finetti's precision dogma

$\overline{P}(x)$ minimum selling price $\equiv$ $\underline{P}(x)$ maximum buying price

Walley's proposal for imprecision

No strong reasons for that rationality only requires $\underline{P}(x) \leq \overline{P}(x)$

- Avoid sure loss! With max buying prices $\underline{P}(A)$ and $\underline{P}(A^c)$, you can buy both gambles and earn one for sure:

$$\underline{P}(A) + \underline{P}(A^c) \leq 1$$

- Be coherent! When buying both A and B, you pay $\underline{P}(A) + \underline{P}(B)$ and you have a gamble which gives one if $A \cup B$ occurs:

$$\underline{P}(A \cup B) \geq \underline{P}(A) + \underline{P}(B)$$

coherence self-consistency (beliefs revised if unsatisfied)
less problematic than a.s.l.

(Some) Reasons for imprecise probabilities

- Reflect the amount of information on which probs are based
- Uniform probs model indifference not ignorance
- When doing introspection, sometimes indecision/indeterminacy
- Easier to assess (e.g., qualitative knowledge, combining beliefs) Assessing precise probs could be possible in principle, but not in practice because of our bounded rationality
- Natural extension of precise models defined on some events determine only imprecise probabilities for events outside
- Robustness in statistics (multiple priors/likelihoods) and decision problems (multiple prob distributions/utilities)

Credal sets (Levi, 1980) as IP models

- Without the precision dogma, incomplete knowledge described by (credal) sets of probability mass functions
- Induced by a finite number of assessments (l/u gambles prices) which are linear constraints on the consistent probabilities
- Sets of consistent (precise) probability mass functions convex with a finite number of extremes (if $|\Omega| < +\infty$)
- E.g., no constraints ⇒ vacuous credal set (model of ignorance)

$$K(X) = \left\{ P(X) \left| \begin{array}{l} \sum_{x\in\Omega} P(x) = 1 \\ P(x) \geq 0 \end{array} \right. \right\}$$

Natural extension

- Price assessments are linear constraints on probabilities (e.g., $\underline{P}(f) = .21$ means $\sum_x P(x)f(x) \geq .21$)
- Compute the extremes $\{P_j(X)\}_{j=1}^{v}$ of the feasible region
- The credal set $K(X)$ is $\mathrm{ConvHull}\{P_j(X)\}_{j=1}^{v}$
- Lower prices/expectations of any gamble/function of/on X

$$\underline{P}(h) = \min_{P(X) \in K(X)} \sum_{x \in \mathcal{X}} P(x) \cdot h(x)$$

LP task: optimum on the extremes of $K(X)$

Computing expectations (inference) on credal sets

- Constrained optimization problem, or
- Combinatorial optimization on the extremes space (# of extremes can be exponential in # of constraints)

Natural extension

- Price assessments are linear constraints on probabilities (e.g., $\underline{P}(f) = .21$ means $\sum_x P(x)f(x) \geq .21$)
- Compute the extremes $\{P_j(X)\}_{j=1}^{v}$ of the feasible region
- The credal set $K(X)$ is $\mathrm{ConvHull}\{P_j(X)\}_{j=1}^{v}$
- Lower prices/expectations of any gamble/function of/on X

$$\underline{P}(h) = \min_{P(X)\in K(X)} \sum_{x\in\mathcal{X}} P(x) \cdot h(x)$$

LP task: optimum on the extremes of $K(X)$

Computing expectations (inference) on credal sets

- Constrained optimization problem, or
- Combinatorial optimization on the extremes space (# of extremes can be exponential in # of constraints)

Natural extension

- Price assessments are linear constraints on probabilities (e.g., $\underline{P}(f) = .21$ means $\sum_x P(x)f(x) \geq .21$)
- Compute the extremes $\{P_j(X)\}_{j=1}^{v}$ of the feasible region
- The credal set $K(X)$ is $\mathrm{ConvHull}\{P_j(X)\}_{j=1}^{v}$
- Lower prices/expectations of any gamble/function of/on X

$$\underline{P}(h) = \min_{P(X) \in K(X)} \sum_{x \in \mathcal{X}} P(x) \cdot h(x)$$

LP task: optimum on the extremes of $K(X)$

Computing expectations (inference) on credal sets

- Constrained optimization problem, or
- Combinatorial optimization on the extremes space (# of extremes can be exponential in # of constraints)

Natural extension

- Price assessments are linear constraints on probabilities (e.g., $\underline{P}(f) = .21$ means $\sum_x P(x)f(x) \geq .21$)
- Compute the extremes $\{P_j(X)\}_{j=1}^v$ of the feasible region
- The credal set $K(X)$ is $\mathrm{ConvHull}\{P_j(X)\}_{j=1}^v$
- Lower prices/expectations of any gamble/function of/on X

$$\underline{P}(h) = \min_{P(X) \in K(X)} \sum_{x \in \mathcal{X}} P(x) \cdot h(x)$$

LP task: optimum on the extremes of $K(X)$

Computing expectations (inference) on credal sets

- Constrained optimization problem, or
- Combinatorial optimization on the extremes space (# of extremes can be exponential in # of constraints)

Natural extension

- Price assessments are linear constraints on probabilities (e.g., $\underline{P}(f) = .21$ means $\sum_x P(x)f(x) \geq .21$)
- Compute the extremes $\{P_j(X)\}_{j=1}^{v}$ of the feasible region
- The credal set $K(X)$ is $\mathrm{ConvHull}\{P_j(X)\}_{j=1}^{v}$
- Lower prices/expectations of any gamble/function of/on X

$$\underline{P}(h) = \min_{P(X) \in K(X)} \sum_{x \in \mathcal{X}} P(x) \cdot h(x)$$

LP task: optimum on the extremes of $K(X)$

Computing expectations (inference) on credal sets

- Constrained optimization problem, or
- Combinatorial optimization on the extremes space (# of extremes can be exponential in # of constraints)

Natural extension

- Price assessments are linear constraints on probabilities (e.g., $\underline{P}(f) = .21$ means $\sum_x P(x)f(x) \geq .21$)
- Compute the extremes $\{P_j(X)\}_{j=1}^{v}$ of the feasible region
- The credal set $K(X)$ is $\mathrm{ConvHull}\{P_j(X)\}_{j=1}^{v}$
- Lower prices/expectations of any gamble/function of/on X

$$\underline{P}(h) = \min_{P(X) \in K(X)} \sum_{x \in \mathcal{X}} P(x) \cdot h(x)$$

LP task: optimum on the extremes of $K(X)$

Computing expectations (inference) on credal sets

- Constrained optimization problem, or
- Combinatorial optimization on the extremes space (# of extremes can be exponential in # of constraints)

Natural extension

- Price assessments are linear constraints on probabilities (e.g., $\underline{P}(f) = .21$ means $\sum_x P(x)f(x) \geq .21$)
- Compute the extremes $\{P_j(X)\}_{j=1}^{v}$ of the feasible region
- The credal set $K(X)$ is $\mathrm{ConvHull}\{P_j(X)\}_{j=1}^{v}$
- Lower prices/expectations of any gamble/function of/on X

$$\underline{P}(h) = \min_{P(X) \in K(X)} \sum_{x \in \mathcal{X}} P(x) \cdot h(x)$$

LP task: optimum on the extremes of $K(X)$

Computing expectations (inference) on credal sets

- Constrained optimization problem, or
- Combinatorial optimization on the extremes space
  (# of extremes can be exponential in # of constraints)

Lower-upper conjugacy

E.g., with events

$$\underline{P}(A) = \min_{P(X)\in K(X)} \sum_{x\in A} P(x)$$

$$\overline{P}(A^c) = \max_{P(X)\in K(X)} \sum_{x\notin A} P(x) = \max_{P(X)\in K(X)} \left[1 - \sum_{x\in A} P(x)\right] = 1 - \underline{P}(A)$$

For gambles, similarly,

$$\overline{P}(-f) = -\underline{P}(f)$$

$$\underline{P}(f) = \min_{P(X)\in K(X)} \sum_{x} P(x)f(x)$$

$$\overline{P}(-f) = \max_{P(X)\in K(X)} \sum_{x} [-P(x)f(x)] = -\min_{P(X)\in K(X)} \sum_{x} P(x)f(x)$$

Self-conjugacy $\equiv$ single-point (precise) credal set $\equiv$ linear functional

Lower-upper conjugacy

E.g., with events

$$\underline{P}(A) = \min_{P(X)\in K(X)} \sum_{x\in A} P(x)$$

$$\overline{P}(A^c) = \max_{P(X)\in K(X)} \sum_{x\notin A} P(x) = \max_{P(X)\in K(X)} \left[1 - \sum_{x\in A} P(x)\right] = 1 - \underline{P}(A)$$

For gambles, similarly,

$$\overline{P}(-f) = -\underline{P}(f)$$

$$\underline{P}(f) = \min_{P(X)\in K(X)} \sum_{x} P(x)f(x)$$

$$\overline{P}(-f) = \max_{P(X)\in K(X)} \sum_{x} [-P(x)f(x)] = -\min_{P(X)\in K(X)} \sum_{x} P(x)f(x)$$

Self-conjugacy ≡ single-point (precise) credal set ≡ linear functional

Lower-upper conjugacy

E.g., with events

$$\underline{P}(A) = \min_{P(X)\in K(X)} \sum_{x\in A} P(x)$$

$$\overline{P}(A^c) = \max_{P(X)\in K(X)} \sum_{x\notin A} P(x) = \max_{P(X)\in K(X)} \left[1 - \sum_{x\in A} P(x)\right] = 1 - \underline{P}(A)$$

For gambles, similarly,

$$\overline{P}(-f) = -\underline{P}(f)$$

$$\underline{P}(f) = \min_{P(X)\in K(X)} \sum_{x} P(x)f(x)$$

$$\overline{P}(-f) = \max_{P(X)\in K(X)} \sum_{x} [-P(x)f(x)] = - \min_{P(X)\in K(X)} \sum_{x} P(x)f(x)$$

Self-conjugacy ≡ single-point (precise) credal set ≡ linear functional

Lower-upper conjugacy

E.g., with events

$$\underline{P}(A) = \min_{P(X) \in K(X)} \sum_{x \in A} P(x)$$

$$\overline{P}(A^c) = \max_{P(X) \in K(X)} \sum_{x \notin A} P(x) = \max_{P(X) \in K(X)} \left[1 - \sum_{x \in A} P(x)\right] = 1 - \underline{P}(A)$$

For gambles, similarly,

$$\overline{P}(-f) = -\underline{P}(f)$$

$$\underline{P}(f) = \min_{P(X) \in K(X)} \sum_{x} P(x) f(x)$$

$$\overline{P}(-f) = \max_{P(X) \in K(X)} \sum_{x} [-P(x) f(x)] = - \min_{P(X) \in K(X)} \sum_{x} P(x) f(x)$$

Self-conjugacy $\equiv$ *single-point (precise) credal set* $\equiv$ *linear functional*

Credal Sets over Boolean Variables

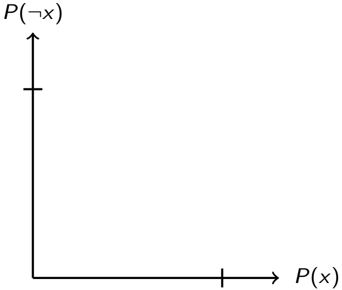
Credal Sets over Boolean Variables

- Boolean X, values in $\Omega_X = \{x, \neg x\}$

Credal Sets over Boolean Variables

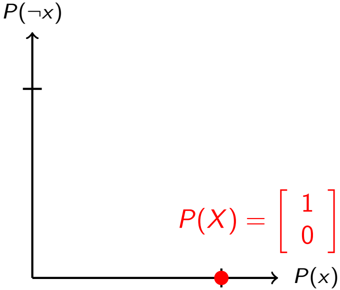
- Boolean X, values in $\Omega_X = \{x, \neg x\}$
- Determinism $\equiv$ degenerate mass f E.g.,

  $X = x \iff P(X) = \left[\begin{array}{c} 1 \\ 0 \end{array} \right]$



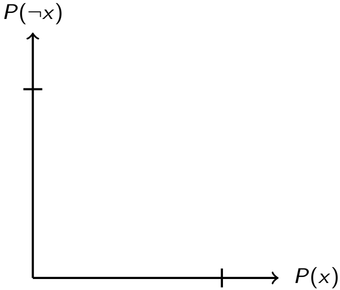
Credal Sets over Boolean Variables

- Boolean X, values in $\Omega_X = \{x, \neg x\}$
- Determinism $\equiv$ degenerate mass f E.g., $X = x \iff P(X) = \left[\begin{array}{c} 1 \\ 0 \end{array} \right]$



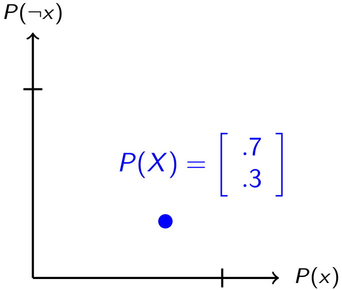
Credal Sets over Boolean Variables

- Boolean X, values in $\Omega_X = \{x, \neg x\}$
- Determinism $\equiv$ degenerate mass f E.g., $X = x \iff P(X) = \left[\begin{array}{c} 1 \\ 0 \end{array} \right]$
- Uncertainty $\equiv$ prob mass function $P(X) = \left[\begin{array}{c} p \\ 1-p \end{array} \right]$ with $p \in [0,1]$



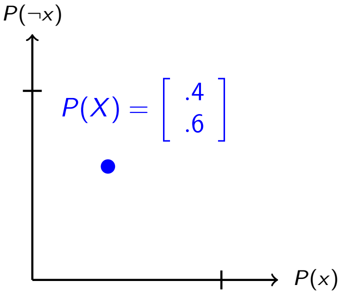
Credal Sets over Boolean Variables

- Boolean X, values in $\Omega_X = \{x, \neg x\}$
- Determinism $\equiv$ degenerate mass f E.g.,
  $X = x \iff P(X) = \left[\begin{array}{c} 1 \\ 0 \end{array}\right]$
- Uncertainty $\equiv$ prob mass function
  $P(X) = \left[\begin{array}{c} p \\ 1-p \end{array}\right]$ with $p \in [0, 1]$



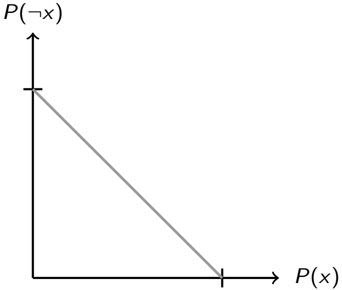
Credal Sets over Boolean Variables

- Boolean X, values in $\Omega_X = \{x, \neg x\}$
- Determinism $\equiv$ degenerate mass f E.g., $X = x \iff P(X) = \left[\begin{array}{c} 1 \\ 0 \end{array} \right]$
- Uncertainty $\equiv$ prob mass function $P(X) = \left[\begin{array}{c} p \\ 1-p \end{array} \right]$ with $p \in [0,1]$



Credal Sets over Boolean Variables

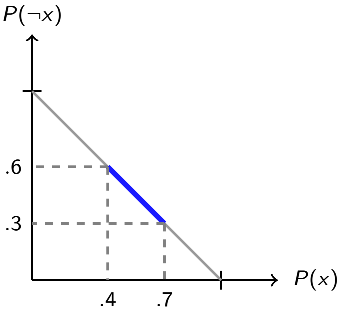
- Boolean X, values in $\Omega_X = \{x, \neg x\}$
- Determinism $\equiv$ degenerate mass f E.g.,
  $X = x \iff P(X) = \left[\begin{array}{c} 1 \\ 0 \end{array}\right]$
- Uncertainty $\equiv$ prob mass function
  $P(X) = \left[\begin{array}{c} p \\ 1-p \end{array}\right]$ with $p \in [0, 1]$
- Imprecision credal set
  on the *probability simplex*



Credal Sets over Boolean Variables

- Boolean X, values in $\Omega_X = \{x, \neg x\}$
- Determinism $\equiv$ degenerate mass f E.g., $X = x \iff P(X) = \left[\begin{array}{c} 1 \\ 0 \end{array} \right]$
- Uncertainty $\equiv$ prob mass function $P(X) = \left[\begin{array}{c} p \\ 1-p \end{array} \right]$ with $p \in [0,1]$
- Imprecision credal set on the *probability simplex*

$$K(X) \equiv \left\{ P(X) = \left[\begin{array}{c} p \\ 1-p \end{array} \right] \middle| .4 \leq p \leq .7 \right\}$$

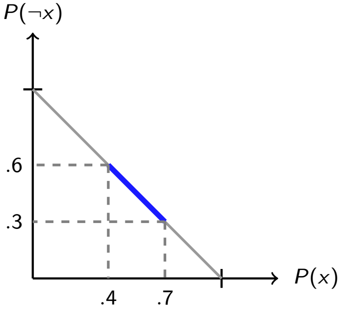


Credal Sets over Boolean Variables

- Boolean X, values in $\Omega_X = \{x, \neg x\}$
- Determinism $\equiv$ degenerate mass f E.g.,
  $X = x \iff P(X) = \left[\begin{array}{c} 1 \\ 0 \end{array} \right]$
- Uncertainty $\equiv$ prob mass function
  $P(X) = \left[\begin{array}{c} p \\ 1-p \end{array} \right]$ with $p \in [0,1]$
- Imprecision credal set
  on the *probability simplex*

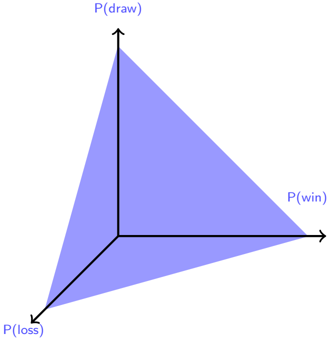
  $$K(X) \equiv \left\{ P(X) = \left[\begin{array}{c} p \\ 1-p \end{array} \right] \middle| .4 \leq p \leq .7 \right\}$$
- A CS over a Boolean variable cannot
  have more than two vertices!

  $$\mathrm{ext}[K(X)] = \left\{ \left[\begin{array}{c} .7 \\ .3 \end{array} \right], \left[\begin{array}{c} .4 \\ .6 \end{array} \right] \right\}$$



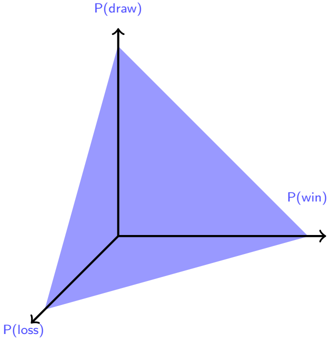
Geometric Representation of CSs (ternary variables)

- Ternary X (e.g., $\Omega = \{$win,draw,loss$\}$)
- $P(X) \equiv$ point in the space (simplex)
- No bounds to $|\mathrm{ext}[K(X)]|$
- Modeling ignorance
  - Uniform models indifference
  - Vacuous credal set
- Expert qualitative knowledge
  - Comparative judgements: win is more probable than draw, which more probable than loss
  - Qualitative judgements: adjective $\equiv$ IP statements



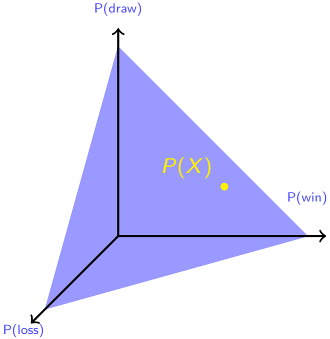
Geometric Representation of CSs (ternary variables)

- Ternary X (e.g., $\Omega = \{\text{win,draw,loss}\}$)
- $P(X) \equiv$ point in the space (simplex)
- No bounds to $|\mathrm{ext}[K(X)]|$
- Modeling ignorance
  - Uniform models indifference
  - Vacuous credal set
- Expert qualitative knowledge
  - Comparative judgements: win is more probable than draw, which more probable than loss
  - Qualitative judgements: adjective $\equiv$ IP statements



Geometric Representation of CSs (ternary variables)

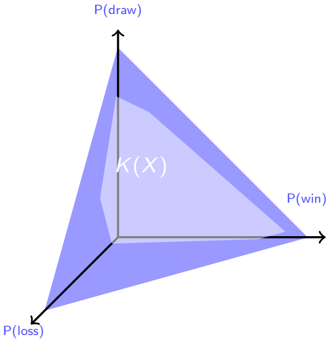
- Ternary X (e.g., $\Omega = \{$win,draw,loss$\}$)
- $P(X) \equiv$ point in the space (simplex)
- No bounds to $|\mathrm{ext}[K(X)]|$
- Modeling ignorance
  - Uniform models indifference
  - Vacuous credal set
- Expert qualitative knowledge
  - Comparative judgements: win is more probable than draw, which more probable than loss
  - Qualitative judgements: adjective $\equiv$ IP statements



$$P(X) = \left[\begin{array}{c} .6 \\ .3 \\ .1 \end{array}\right]$$

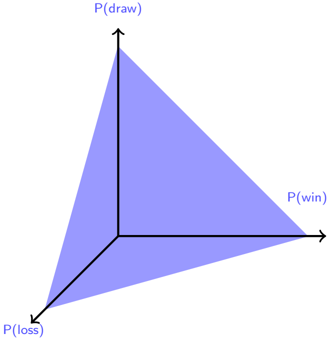
Geometric Representation of CSs (ternary variables)

- Ternary X (e.g., $\Omega = \{$win,draw,loss$\}$)
- $P(X) \equiv$ point in the space (simplex)
- No bounds to $|\mathrm{ext}[K(X)]|$
- Modeling ignorance
  - Uniform models indifference
  - Vacuous credal set
- Expert qualitative knowledge
  - Comparative judgements: win is more probable than draw, which more probable than loss
  - Qualitative judgements: adjective $\equiv$ IP statements



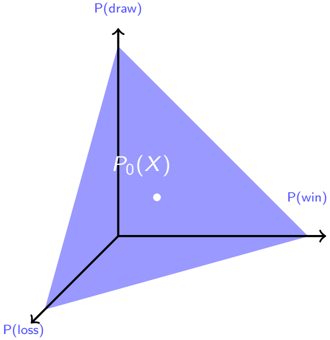
Geometric Representation of CSs (ternary variables)

- Ternary X (e.g., $\Omega = \{$win,draw,loss$\}$)
- $P(X) \equiv$ point in the space (simplex)
- No bounds to $|\mathrm{ext}[K(X)]|$
- Modeling ignorance
  - Uniform models indifference
  - Vacuous credal set
- Expert qualitative knowledge
  - Comparative judgements: win is more probable than draw, which more probable than loss
  - Qualitative judgements: adjective $\equiv$ IP statements



Geometric Representation of CSs (ternary variables)

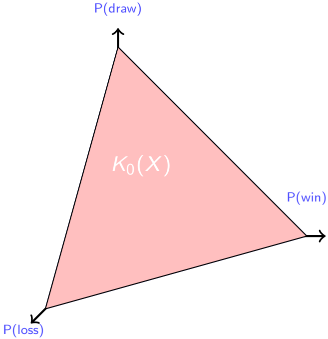
- Ternary X (e.g., $\Omega = \{$win,draw,loss$\}$)
- $P(X) \equiv$ point in the space (simplex)
- No bounds to $|\mathrm{ext}[K(X)]|$
- Modeling ignorance
  - Uniform models indifference
  - Vacuous credal set
- Expert qualitative knowledge
  - Comparative judgements: win is more probable than draw, which more probable than loss
  - Qualitative judgements: adjective $\equiv$ IP statements



$$P_0(x) = \frac{1}{|\Omega_X|}$$

Geometric Representation of CSs (ternary variables)

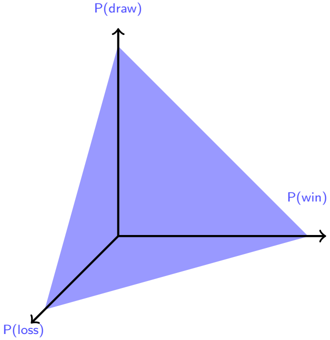
- Ternary X (e.g., $\Omega = \{$win,draw,loss$\}$)
- $P(X) \equiv$ point in the space (simplex)
- No bounds to $|\mathrm{ext}[K(X)]|$
- Modeling ignorance
  - Uniform models indifference
  - Vacuous credal set
- Expert qualitative knowledge
  - Comparative judgements: win is more probable than draw, which more probable than loss
  - Qualitative judgements: adjective $\equiv$ IP statements



$$K_0(X)=\left\{P(X) \;\middle|\; \begin{array}{l} \sum_x P(x) = 1, \\ P(x) \geq 0 \end{array} \right\}$$

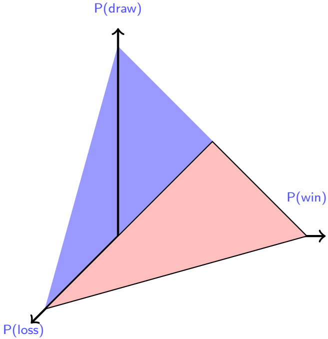
Geometric Representation of CSs (ternary variables)

- Ternary X (e.g., $\Omega = \{$win,draw,loss$\}$)
- $P(X) \equiv$ point in the space (simplex)
- No bounds to $|\mathrm{ext}[K(X)]|$
- Modeling ignorance
  - Uniform models indifference
  - Vacuous credal set
- Expert qualitative knowledge
  - Comparative judgements: win is more probable than draw, which more probable than loss
  - Qualitative judgements: adjective $\equiv$ IP statements



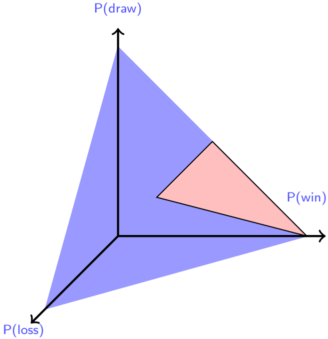
Geometric Representation of CSs (ternary variables)

- Ternary X (e.g., $\Omega = \{$win,draw,loss$\}$)
- $P(X) \equiv$ point in the space (simplex)
- No bounds to $|\mathrm{ext}[K(X)]|$
- Modeling ignorance
  - Uniform models indifference
  - Vacuous credal set
- Expert qualitative knowledge
  - Comparative judgements: win is more probable than draw, which more probable than loss
  - Qualitative judgements: adjective $\equiv$ IP statements



Geometric Representation of CSs (ternary variables)

- Ternary X (e.g., $\Omega = \{$win,draw,loss$\}$)
- $P(X) \equiv$ point in the space (simplex)
- No bounds to $|\mathrm{ext}[K(X)]|$
- Modeling ignorance
  - Uniform models indifference
  - Vacuous credal set
- Expert qualitative knowledge
  - Comparative judgements: win is more probable than draw, which more probable than loss
  - Qualitative judgements: adjective $\equiv$ IP statements



Geometric Representation of CSs (ternary variables)

- Ternary X (e.g., $\Omega = \{$win,draw,loss$\}$)
- $P(X) \equiv$ point in the space (simplex)
- No bounds to $|\mathrm{ext}[K(X)]|$
- Modeling ignorance
  - Uniform models indifference
  - Vacuous credal set
- Expert qualitative knowledge
  - Comparative judgements: win is more probable than draw, which more probable than loss
  - Qualitative judgements: adjective $\equiv$ IP statements

From natural language to linear constraints on probabilities

(Walley, 1991)

extremely probable $P(x) \geq 0.98$
very high probability $P(x) \geq 0.9$
highly probable $P(x) \geq 0.85$
very probable $P(x) \geq 0.75$
has a very good chance $P(x) \geq 0.65$
quite probable $P(x) \geq 0.6$
probable $P(x) \geq 0.5$
has a good chance $0.4 \leq P(x) \leq 0.85$
is improbable (unlikely) $P(x) \leq 0.5$
is somewhat unlikely $P(x) \leq 0.4$
is very unlikely $P(x) \leq 0.25$
has little chance $P(x) \leq 0.2$
is highly improbable $P(x) \leq 0.15$
is has very low probability $P(x) \leq 0.1$
is extremely unlikely $P(x) \leq 0.02$

Marginalization (and credal sets in 4D)

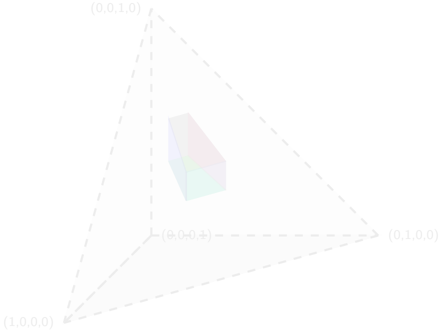
- Two Boolean variables: **S**moker, Lung **C**ancer
- 8 “Bayesian” physicians, each assessing $P_j(S, C)$

$$K(S, C) = \mathrm{CH}\left\{P_j(S, C)\right\}_{j=1}^{8}$$

j	$P_j(s, c)$	$P_j(s, \bar{c})$	$P_j(\bar{s}, c)$	$P_j(\bar{s}, \bar{c})$
1	1/8	1/8	3/8	3/8
2	1/8	1/8	9/16	3/16
3	3/16	1/16	3/8	3/8
4	3/16	1/16	9/16	3/16
5	1/4	1/4	1/4	1/4
6	1/4	1/4	3/8	1/8
7	3/8	1/8	1/4	1/4
8	3/8	1/8	3/8	1/8

Marginals elementwise (on extremes)

$$K(C) = CH\left\{\sum_s P_j(C, s)\right\}_{j=1}^{8} \qquad \frac{1}{2} \leq P(c) \leq \frac{3}{4}$$



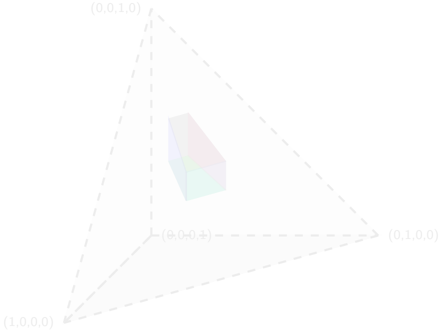
Marginalization (and credal sets in 4D)

- Two Boolean variables:
  Smoker, Lung **C**ancer
- 8 “Bayesian” physicians,
  each assessing $P_j(S, C)$
  $K(S, C) = \mathrm{CH}\left\{P_j(S, C)\right\}_{j=1}^{8}$

j	$P_j(s, c)$	$P_j(s, \bar{c})$	$P_j(\bar{s}, c)$	$P_j(\bar{s}, \bar{c})$
1	1/8	1/8	3/8	3/8
2	1/8	1/8	9/16	3/16
3	3/16	1/16	3/8	3/8
4	3/16	1/16	9/16	3/16
5	1/4	1/4	1/4	1/4
6	1/4	1/4	3/8	1/8
7	3/8	1/8	1/4	1/4
8	3/8	1/8	3/8	1/8

Marginals elementwise (on extremes)

$$K(C) = CH\left\{\sum_s P_j(C, s)\right\}_{j=1}^{8} \quad \frac{1}{2} \leq P(c) \leq \frac{3}{4}$$



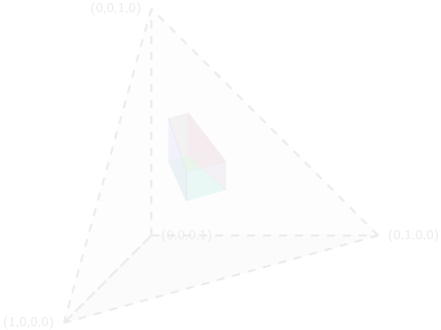
Marginalization (and credal sets in 4D)

- Two Boolean variables: **S**moker, Lung **C**ancer
- 8 “Bayesian” physicians, each assessing $P_j(S, C)$
  $K(S, C) = \mathrm{CH}\,\{P_j(S, C)\}_{j=1}^{8}$

j	$P_j(s, c)$	$P_j(s, \bar{c})$	$P_j(\bar{s}, c)$	$P_j(\bar{s}, \bar{c})$
1	1/8	1/8	3/8	3/8
2	1/8	1/8	9/16	3/16
3	3/16	1/16	3/8	3/8
4	3/16	1/16	9/16	3/16
5	1/4	1/4	1/4	1/4
6	1/4	1/4	3/8	1/8
7	3/8	1/8	1/4	1/4
8	3/8	1/8	3/8	1/8

Marginals elementwise (on extremes)

$$K(C) = CH\left\{\sum_s P_j(C, s)\right\}_{j=1}^{8} \quad \frac{1}{2} \leq P(c) \leq \frac{3}{4}$$



Marginalization (and credal sets in 4D)

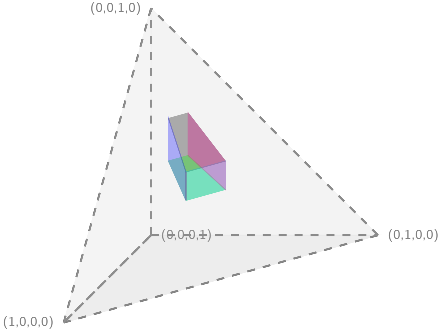
- Two Boolean variables: **S**moker, Lung **C**ancer
- 8 "Bayesian" physicians, each assessing $P_j(S,C)$

$$K(S,C) = \mathrm{CH}\left\{P_j(S,C)\right\}_{j=1}^{8}$$

j	$P_j(s,c)$	$P_j(s,\bar{c})$	$P_j(\bar{s},c)$	$P_j(\bar{s},\bar{c})$
1	1/8	1/8	3/8	3/8
2	1/8	1/8	9/16	3/16
3	3/16	1/16	3/8	3/8
4	3/16	1/16	9/16	3/16
5	1/4	1/4	1/4	1/4
6	1/4	1/4	3/8	1/8
7	3/8	1/8	1/4	1/4
8	3/8	1/8	3/8	1/8

Marginals elementwise (on extremes)

$$K(C) = CH\left\{\sum_s P_j(C,s)\right\}_{j=1}^{8} \quad \frac{1}{2} \leq P(c) \leq \frac{3}{4}$$



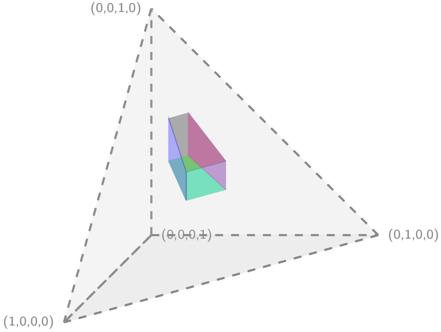
Marginalization (and credal sets in 4D)

- Two Boolean variables: **S**moker, Lung **C**ancer
- 8 “Bayesian” physicians, each assessing $P_j(S, C)$
  $K(S, C) = \mathrm{CH}\left\{P_j(S, C)\right\}_{j=1}^{8}$

j	$P_j(s, c)$	$P_j(s, \bar{c})$	$P_j(\bar{s}, c)$	$P_j(\bar{s}, \bar{c})$
1	1/8	1/8	3/8	3/8
2	1/8	1/8	9/16	3/16
3	3/16	1/16	3/8	3/8
4	3/16	1/16	9/16	3/16
5	1/4	1/4	1/4	1/4
6	1/4	1/4	3/8	1/8
7	3/8	1/8	1/4	1/4
8	3/8	1/8	3/8	1/8

Marginals elementwise (on extremes)

$$K(C) = CH\left\{\sum_{s} P_j(C, s)\right\}_{j=1}^{8} \quad \frac{1}{2} \leq P(c) \leq \frac{3}{4}$$



Credal sets induced by probability intervals

- Assessing lower and upper probabilities: $[l_x, u_x]$, for each $x \in \Omega$
- The consistent credal set is

$$K(X) := \left\{ P(X) \,\middle|\, \begin{array}{l} l_x \leq P(x) \leq u_x \\ P(x) \geq 0 \\ \sum_x P(x) = 1 \end{array} \right\}$$

- Avoiding sure loss implies non-emptiness of the credal set

$$\sum_x l_x \leq 1 \leq \sum_x u_x$$

- Coherence implies the reachability (bounds are tight)

$$u_x + \sum_{x' \neq x} l_x \leq 1 \qquad l_x + \sum_{x' \neq x} u_x \geq 1$$

Credal sets induced by probability intervals

- Assessing lower and upper probabilities: $[l_x, u_x]$, for each $x \in \Omega$
- The consistent credal set is

$$K(X) := \left\{ P(X) \;\middle|\; \begin{array}{l} l_x \leq P(x) \leq u_x \\ P(x) \geq 0 \\ \sum_x P(x) = 1 \end{array} \right\}$$

- Avoiding sure loss implies non-emptiness of the credal set

$$\sum_x l_x \leq 1 \leq \sum_x u_x$$

- Coherence implies the reachability (bounds are tight)

$$u_x + \sum_{x' \neq x} l_x \leq 1 \qquad l_x + \sum_{x' \neq x} u_x \geq 1$$

Credal sets induced by probability intervals

- Assessing lower and upper probabilities: $[l_x, u_x]$, for each $x \in \Omega$
- The consistent credal set is

$$K(X) := \left\{ P(X) \,\middle|\, \begin{array}{l} l_x \leq P(x) \leq u_x \\ P(x) \geq 0 \\ \sum_x P(x) = 1 \end{array} \right\}$$

- Avoiding sure loss implies non-emptiness of the credal set

$$\sum_x l_x \leq 1 \leq \sum_x u_x$$

- Coherence implies the reachability (bounds are tight)

$$u_x + \sum_{x' \neq x} l_x \leq 1 \qquad l_x + \sum_{x' \neq x} u_x \geq 1$$

Credal sets induced by probability intervals

- Assessing lower and upper probabilities: $[l_x, u_x]$, for each $x \in \Omega$
- The consistent credal set is

$$K(X) := \left\{ P(X) \;\middle|\; \begin{array}{l} l_x \leq P(x) \leq u_x \\ P(x) \geq 0 \\ \sum_x P(x) = 1 \end{array} \right\}$$

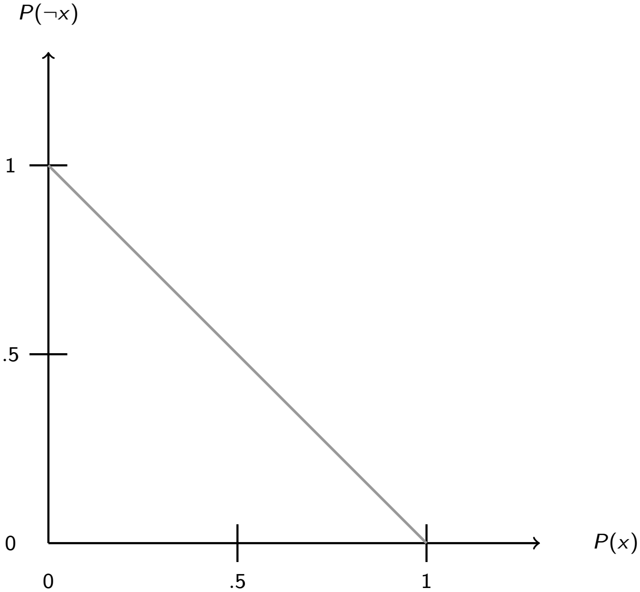
- Avoiding sure loss implies non-emptiness of the credal set

$$\sum_x l_x \leq 1 \leq \sum_x u_x$$

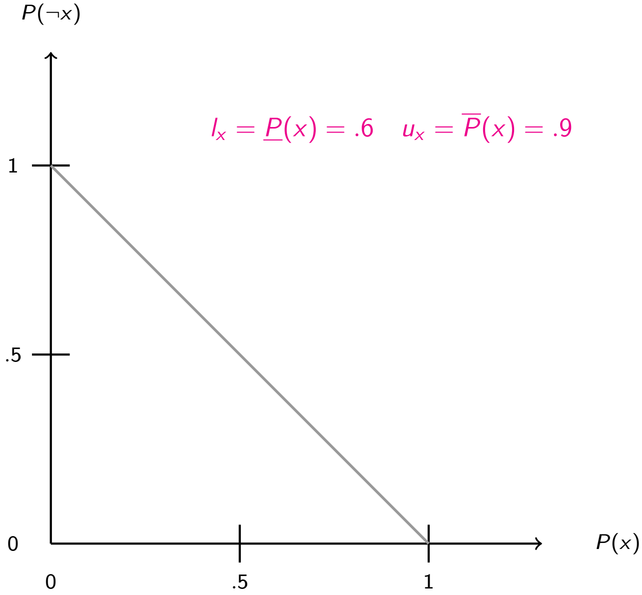
- Coherence implies the reachability (bounds are tight)

$$u_x + \sum_{x' \neq x} l_x \leq 1 \qquad l_x + \sum_{x' \neq x} u_x \geq 1$$

Refining assessments (when possible)

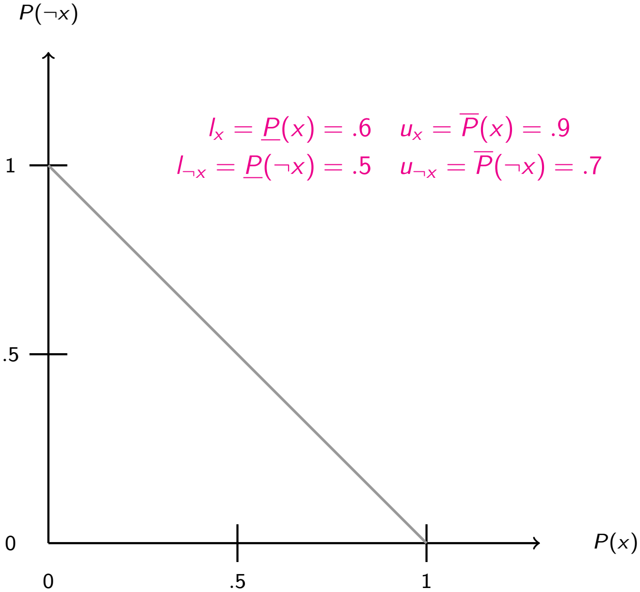


Refining assessments (when possible)



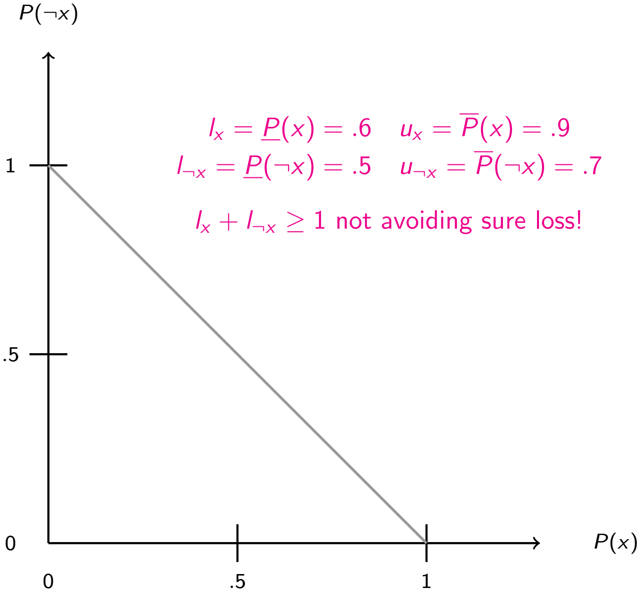
$l_x = \underline{P}(x) = .6 \quad u_x = \overline{P}(x) = .9$

Refining assessments (when possible)



$$l_x = \underline{P}(x) = .6 \quad u_x = \overline{P}(x) = .9$$
$$l_{\neg x} = \underline{P}(\neg x) = .5 \quad u_{\neg x} = \overline{P}(\neg x) = .7$$

Refining assessments (when possible)

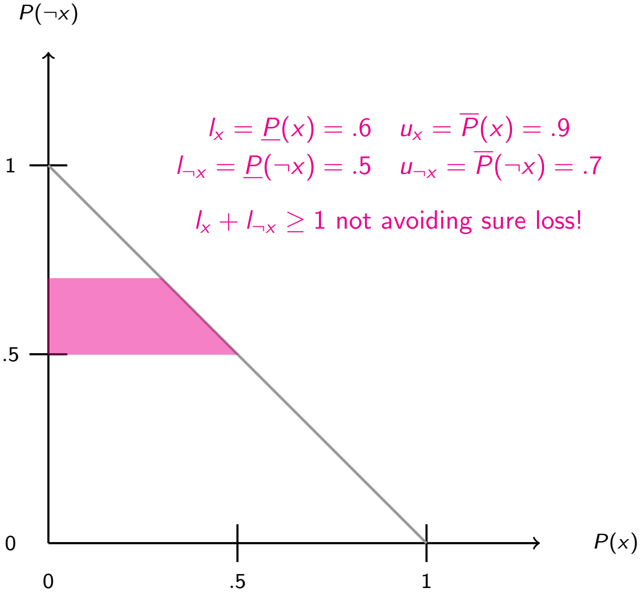


$l_x = \underline{P}(x) = .6 \quad u_x = \overline{P}(x) = .9$

$l_{\neg x} = \underline{P}(\neg x) = .5 \quad u_{\neg x} = \overline{P}(\neg x) = .7$

$l_x + l_{\neg x} \geq 1$ not avoiding sure loss!

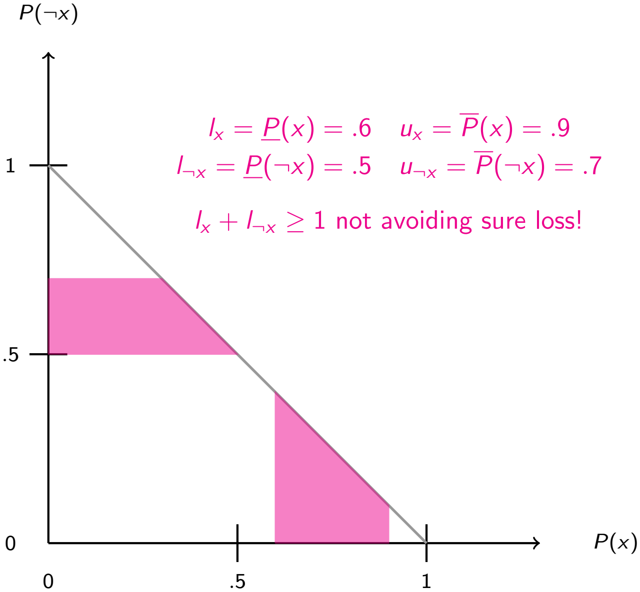
Refining assessments (when possible)



$$l_x = \underline{P}(x) = .6 \quad u_x = \overline{P}(x) = .9$$
$$l_{\neg x} = \underline{P}(\neg x) = .5 \quad u_{\neg x} = \overline{P}(\neg x) = .7$$

$l_x + l_{\neg x} \geq 1$ not avoiding sure loss!

Refining assessments (when possible)

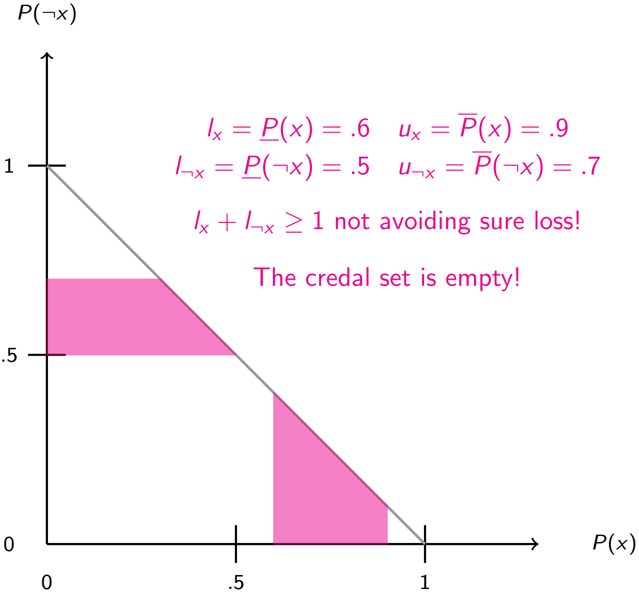


$l_x = \underline{P}(x) = .6 \quad u_x = \overline{P}(x) = .9$

$l_{\neg x} = \underline{P}(\neg x) = .5 \quad u_{\neg x} = \overline{P}(\neg x) = .7$

$l_x + l_{\neg x} \geq 1$ not avoiding sure loss!

Refining assessments (when possible)



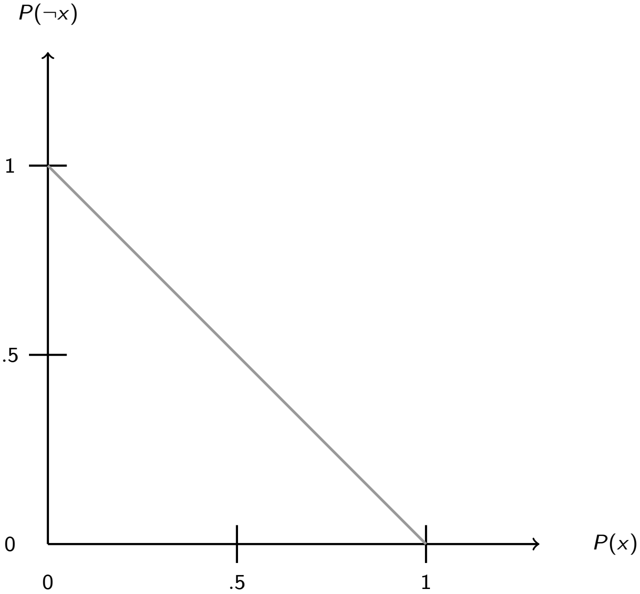
$$l_x = \underline{P}(x) = .6 \quad u_x = \overline{P}(x) = .9$$

$$l_{\neg x} = \underline{P}(\neg x) = .5 \quad u_{\neg x} = \overline{P}(\neg x) = .7$$

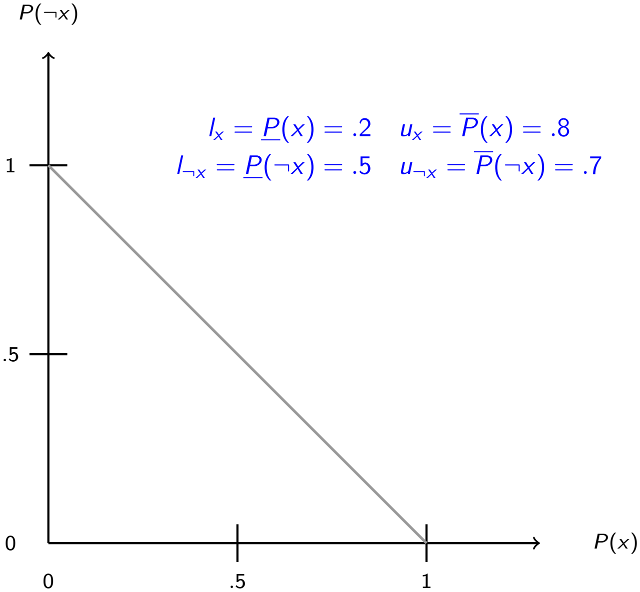
$l_x + l_{\neg x} \geq 1$ not avoiding sure loss!

The credal set is empty!

Refining assessments (when possible)

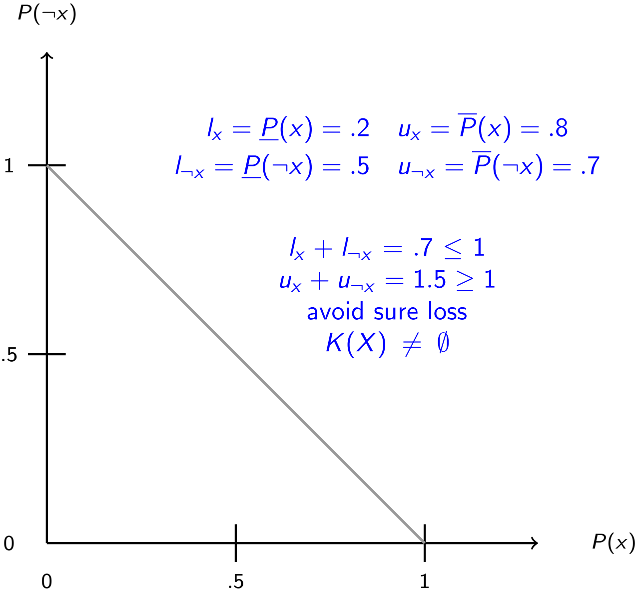


Refining assessments (when possible)



$$l_x = \underline{P}(x) = .2 \quad u_x = \overline{P}(x) = .8$$
$$l_{\neg x} = \underline{P}(\neg x) = .5 \quad u_{\neg x} = \overline{P}(\neg x) = .7$$

Refining assessments (when possible)



$$l_x = \underline{P}(x) = .2 \quad u_x = \overline{P}(x) = .8$$

$$l_{\neg x} = \underline{P}(\neg x) = .5 \quad u_{\neg x} = \overline{P}(\neg x) = .7$$

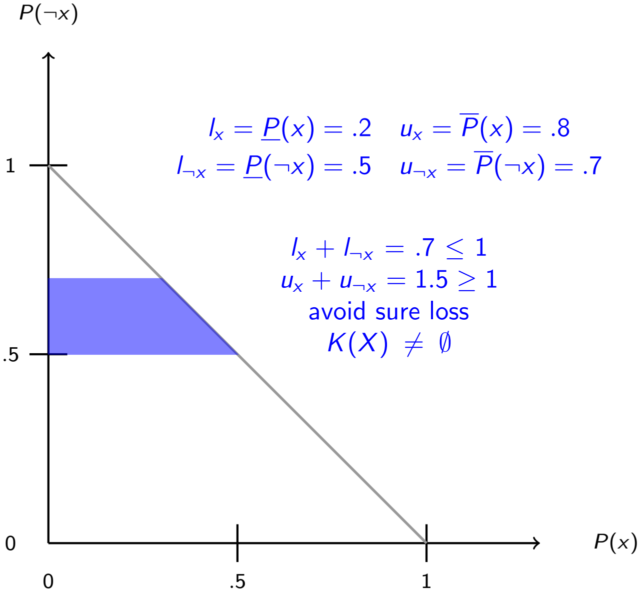
$$l_x + l_{\neg x} = .7 \leq 1$$

$$u_x + u_{\neg x} = 1.5 \geq 1$$

avoid sure loss

$$K(X) \neq \emptyset$$

Refining assessments (when possible)



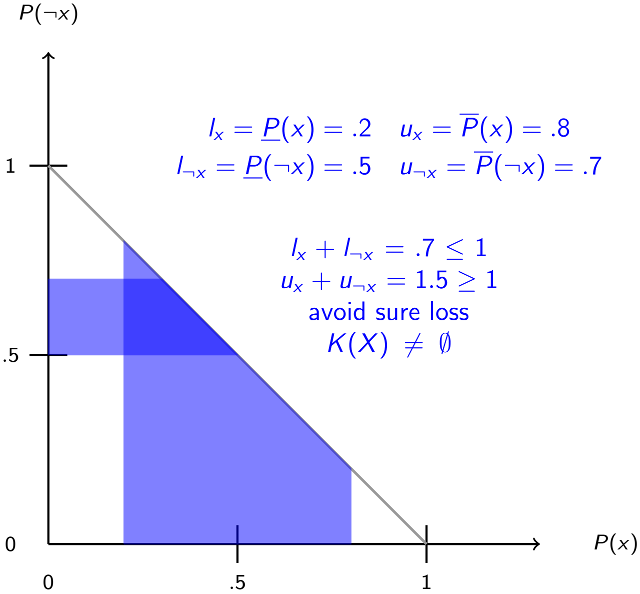
$$l_x = \underline{P}(x) = .2 \quad u_x = \overline{P}(x) = .8$$
$$l_{\neg x} = \underline{P}(\neg x) = .5 \quad u_{\neg x} = \overline{P}(\neg x) = .7$$

$$l_x + l_{\neg x} = .7 \leq 1$$
$$u_x + u_{\neg x} = 1.5 \geq 1$$

avoid sure loss

$$K(X) \neq \emptyset$$

Refining assessments (when possible)



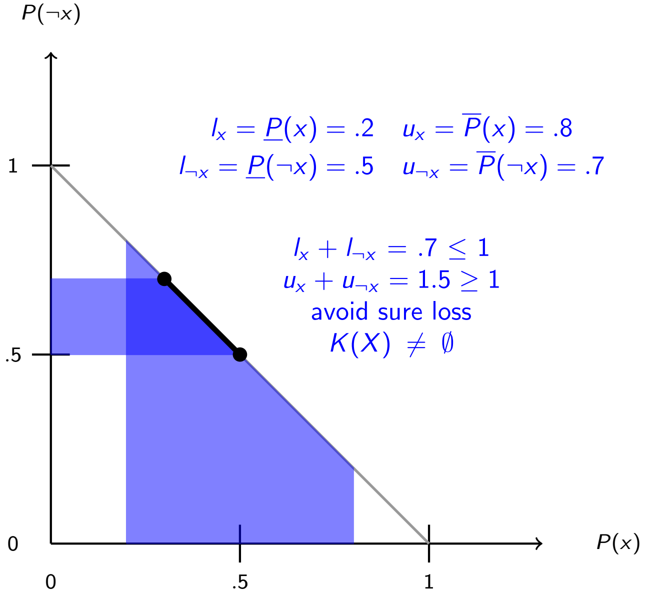
$$l_x = \underline{P}(x) = .2 \quad u_x = \overline{P}(x) = .8$$
$$l_{\neg x} = \underline{P}(\neg x) = .5 \quad u_{\neg x} = \overline{P}(\neg x) = .7$$

$$l_x + l_{\neg x} = .7 \leq 1$$
$$u_x + u_{\neg x} = 1.5 \geq 1$$

avoid sure loss

$$K(X) \neq \emptyset$$

Refining assessments (when possible)



$l_x = \underline{P}(x) = .2 \quad u_x = \overline{P}(x) = .8$

$l_{\neg x} = \underline{P}(\neg x) = .5 \quad u_{\neg x} = \overline{P}(\neg x) = .7$

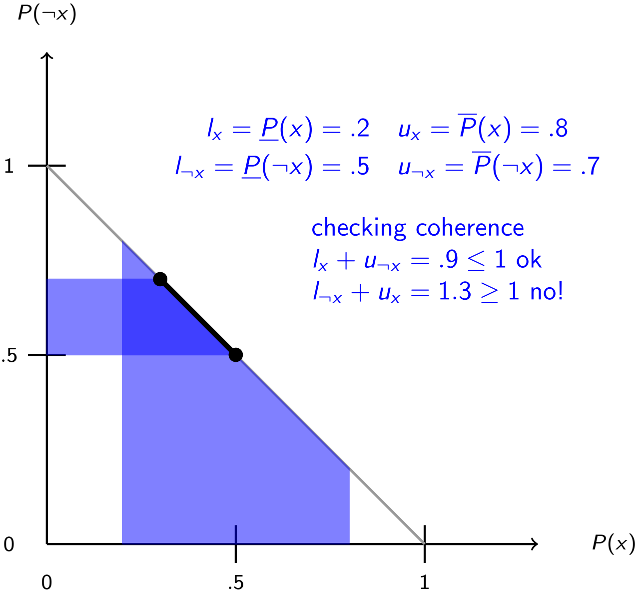
$l_x + l_{\neg x} = .7 \leq 1$

$u_x + u_{\neg x} = 1.5 \geq 1$

avoid sure loss

$K(X) \neq \emptyset$

Refining assessments (when possible)



$l_x = \underline{P}(x) = .2 \quad u_x = \overline{P}(x) = .8$

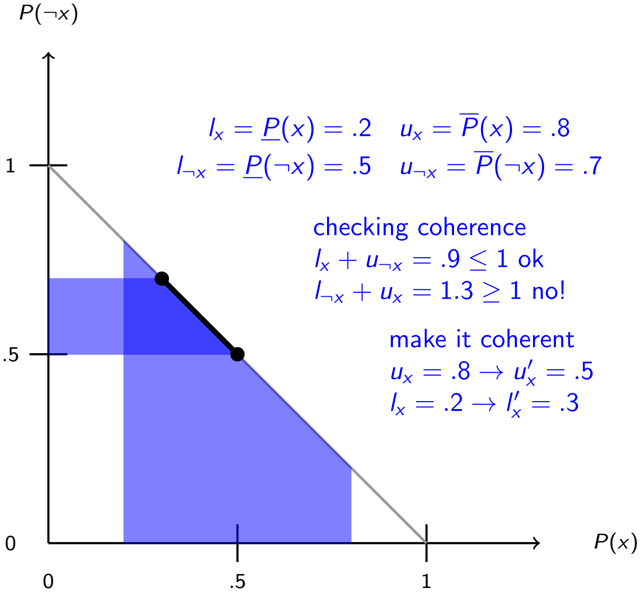
$l_{\neg x} = \underline{P}(\neg x) = .5 \quad u_{\neg x} = \overline{P}(\neg x) = .7$

checking coherence

$l_x + u_{\neg x} = .9 \leq 1$ ok

$l_{\neg x} + u_x = 1.3 \geq 1$ no!

Refining assessments (when possible)



$$l_x = \underline{P}(x) = .2 \quad u_x = \overline{P}(x) = .8$$
$$l_{\neg x} = \underline{P}(\neg x) = .5 \quad u_{\neg x} = \overline{P}(\neg x) = .7$$

checking coherence

$l_x + u_{\neg x} = .9 \leq 1$ ok

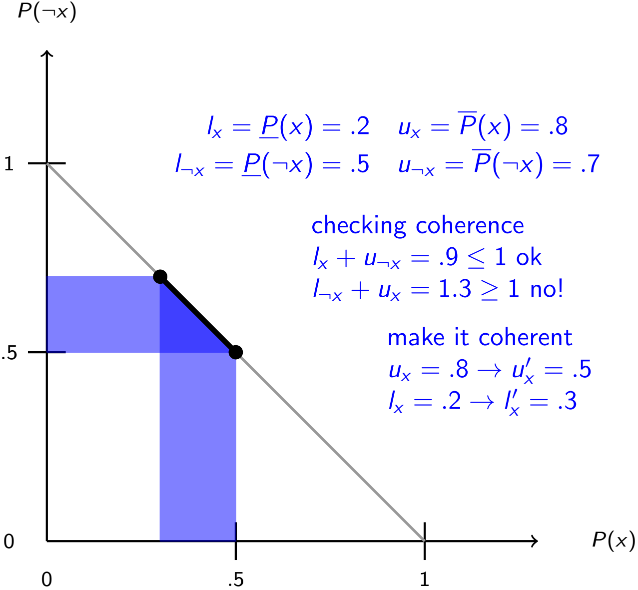
$l_{\neg x} + u_x = 1.3 \geq 1$ no!

make it coherent

$u_x = .8 \rightarrow u'_x = .5$

$l_x = .2 \rightarrow l'_x = .3$

Refining assessments (when possible)



$l_x = \underline{P}(x) = .2 \quad u_x = \overline{P}(x) = .8$

$l_{\neg x} = \underline{P}(\neg x) = .5 \quad u_{\neg x} = \overline{P}(\neg x) = .7$

checking coherence

$l_x + u_{\neg x} = .9 \leq 1$ ok

$l_{\neg x} + u_x = 1.3 \geq 1$ no!

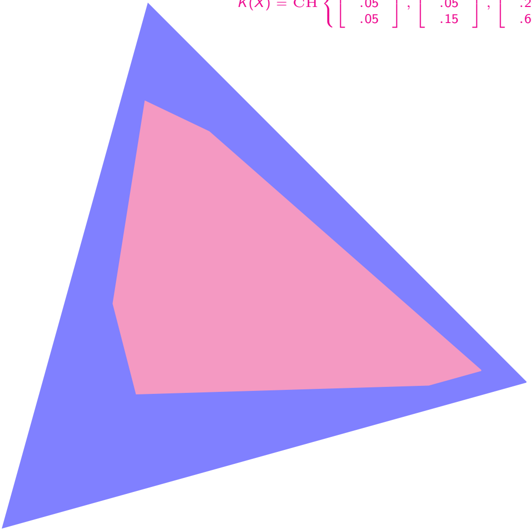
make it coherent

$u_x = .8 \rightarrow u'_x = .5$

$l_x = .2 \rightarrow l'_x = .3$

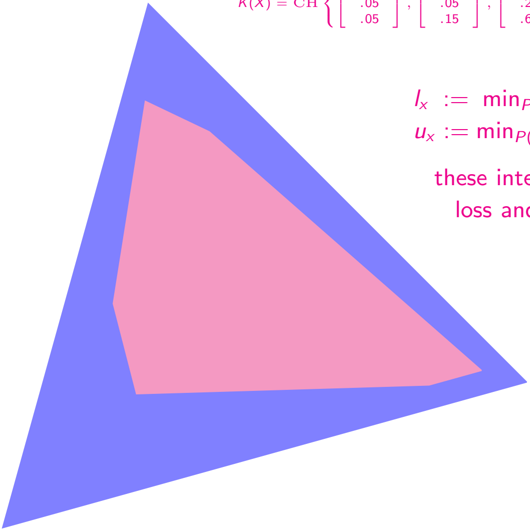
Probability intervals are not fully general

$$K(X) = \mathrm{CH}\left\{ \begin{bmatrix} .90 \\ .05 \\ .05 \end{bmatrix}, \begin{bmatrix} .80 \\ .05 \\ .15 \end{bmatrix}, \begin{bmatrix} .20 \\ .20 \\ .60 \end{bmatrix}, \begin{bmatrix} .10 \\ .40 \\ .50 \end{bmatrix}, \begin{bmatrix} .05 \\ .80 \\ .15 \end{bmatrix}, \begin{bmatrix} .20 \\ .70 \\ .10 \end{bmatrix} \right\}$$



Probability intervals are not fully general

$$K(X) = \mathrm{CH}\left\{\begin{bmatrix} .90 \\ .05 \\ .05 \end{bmatrix}, \begin{bmatrix} .80 \\ .05 \\ .15 \end{bmatrix}, \begin{bmatrix} .20 \\ .20 \\ .60 \end{bmatrix}, \begin{bmatrix} .10 \\ .40 \\ .50 \end{bmatrix}, \begin{bmatrix} .05 \\ .80 \\ .15 \end{bmatrix}, \begin{bmatrix} .20 \\ .70 \\ .10 \end{bmatrix}\right\}$$

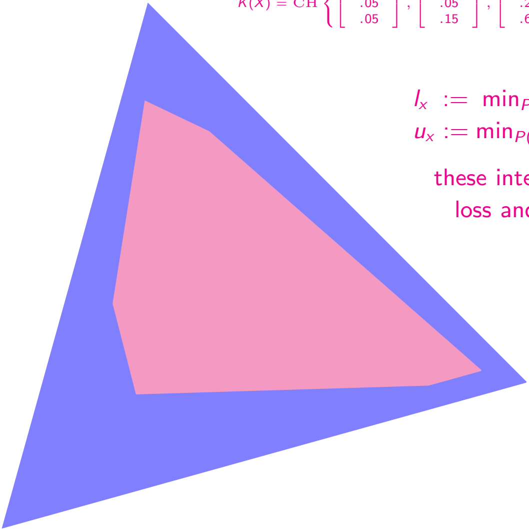


$$l_x := \min_{P(X)\in K(X)} p(x)$$
$$u_x := \min_{P(X)\in K(X)} p(x)$$

these intervals avoid sure loss and are coherent

Probability intervals are not fully general

$$K(X) = \mathrm{CH}\left\{ \begin{bmatrix} .90 \\ .05 \\ .05 \end{bmatrix}, \begin{bmatrix} .80 \\ .05 \\ .15 \end{bmatrix}, \begin{bmatrix} .20 \\ .20 \\ .60 \end{bmatrix}, \begin{bmatrix} .10 \\ .40 \\ .50 \end{bmatrix}, \begin{bmatrix} .05 \\ .80 \\ .15 \end{bmatrix}, \begin{bmatrix} .20 \\ .70 \\ .10 \end{bmatrix} \right\}$$



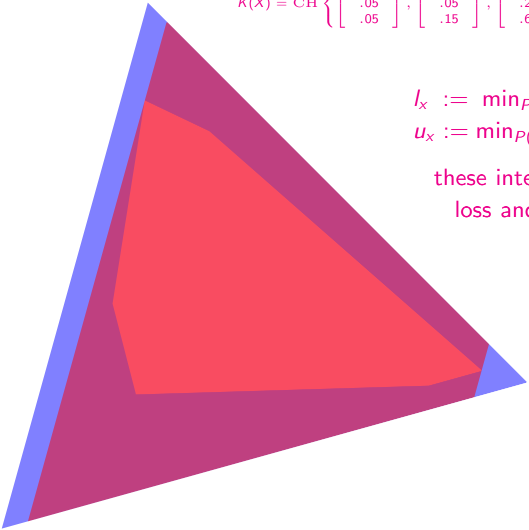
$$\begin{aligned} l_x &:= \min_{P(X) \in K(X)} p(x) \\ u_x &:= \min_{P(X) \in K(X)} p(x) \end{aligned}$$

these intervals avoid sure loss and are coherent

$$\begin{aligned} [l_{x'}, u_{x'}] &= [.05, .90] \\ [l_{x''}, u_{x''}] &= [.05, .80] \\ [l_{x'''}, u_{x'''}] &= [.05, .60] \end{aligned}$$

Probability intervals are not fully general

$$K(X) = \mathrm{CH}\left\{\begin{bmatrix} .90 \\ .05 \\ .05 \end{bmatrix}, \begin{bmatrix} .80 \\ .05 \\ .15 \end{bmatrix}, \begin{bmatrix} .20 \\ .20 \\ .60 \end{bmatrix}, \begin{bmatrix} .10 \\ .40 \\ .50 \end{bmatrix}, \begin{bmatrix} .05 \\ .80 \\ .15 \end{bmatrix}, \begin{bmatrix} .20 \\ .70 \\ .10 \end{bmatrix}\right\}$$



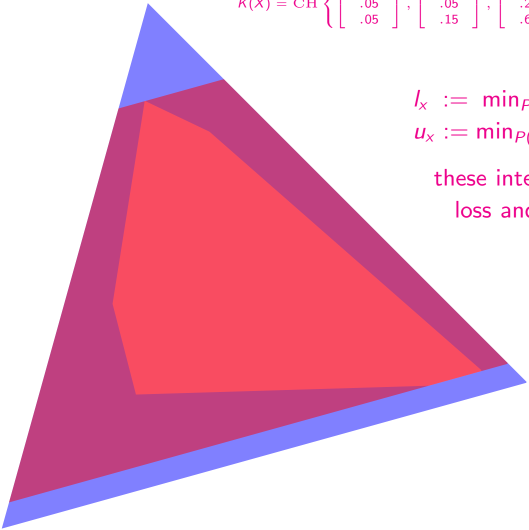
$$l_x := \min_{P(X)\in K(X)} p(x)$$
$$u_x := \min_{P(X)\in K(X)} p(x)$$

these intervals avoid sure loss and are coherent

$$[l_{x'}, u_{x'}] = [.05, .90]$$
$$[l_{x''}, u_{x''}] = [.05, .80]$$
$$[l_{x'''}, u_{x'''}] = [.05, .60]$$

Probability intervals are not fully general

$$K(X) = \mathrm{CH}\left\{ \begin{bmatrix} .90 \\ .05 \\ .05 \end{bmatrix}, \begin{bmatrix} .80 \\ .05 \\ .15 \end{bmatrix}, \begin{bmatrix} .20 \\ .20 \\ .60 \end{bmatrix}, \begin{bmatrix} .10 \\ .40 \\ .50 \end{bmatrix}, \begin{bmatrix} .05 \\ .80 \\ .15 \end{bmatrix}, \begin{bmatrix} .20 \\ .70 \\ .10 \end{bmatrix} \right\}$$



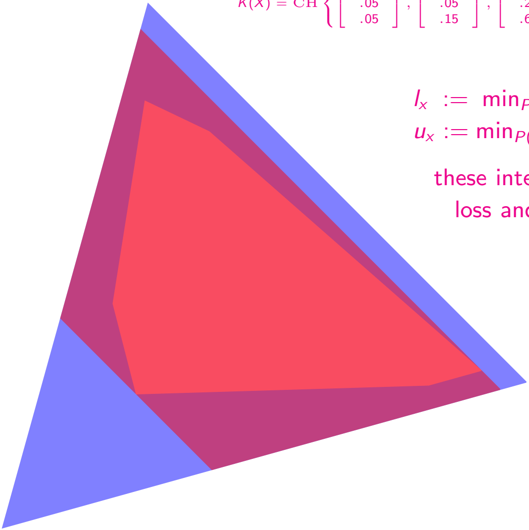
$$l_x \;:=\; \min_{P(X)\in K(X)} p(x)$$
$$u_x := \min_{P(X)\in K(X)} p(x)$$

these intervals avoid sure loss and are coherent

$$[l_{x'}, u_{x'}] \;=\; [.05, .90]$$
$$[l_{x''}, u_{x''}] \;=\; [.05, .80]$$
$$[l_{x'''}, u_{x'''}] = [.05, .60]$$

Probability intervals are not fully general

$$K(X) = \mathrm{CH}\left\{ \begin{bmatrix} .90 \\ .05 \\ .05 \end{bmatrix}, \begin{bmatrix} .80 \\ .05 \\ .15 \end{bmatrix}, \begin{bmatrix} .20 \\ .20 \\ .60 \end{bmatrix}, \begin{bmatrix} .10 \\ .40 \\ .50 \end{bmatrix}, \begin{bmatrix} .05 \\ .80 \\ .15 \end{bmatrix}, \begin{bmatrix} .20 \\ .70 \\ .10 \end{bmatrix} \right\}$$



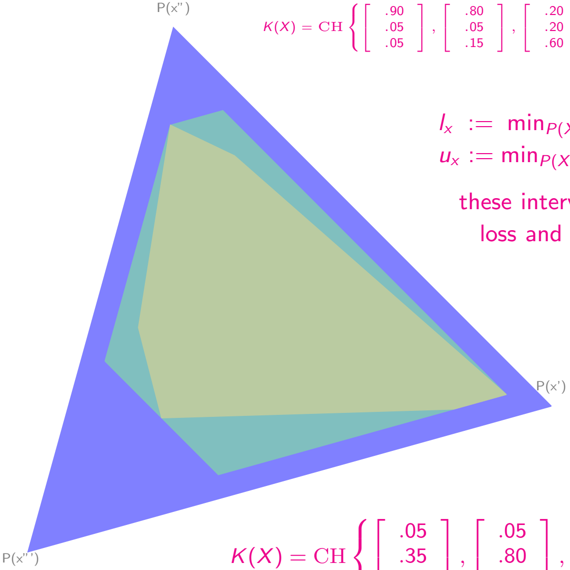
$$l_x \ := \ \min_{P(X) \in K(X)} p(x)$$
$$u_x \ := \min_{P(X) \in K(X)} p(x)$$

these intervals avoid sure loss and are coherent

$$[l_{x'}, u_{x'}] \ = \ [.05, .90]$$
$$[l_{x''}, u_{x''}] \ = \ [.05, .80]$$
$$[l_{x'''}, u_{x'''}] = [.05, .60]$$

Probability intervals are not fully general

$$K(X) = \mathrm{CH}\left\{ \begin{bmatrix} .90 \\ .05 \\ .05 \end{bmatrix}, \begin{bmatrix} .80 \\ .05 \\ .15 \end{bmatrix}, \begin{bmatrix} .20 \\ .20 \\ .60 \end{bmatrix}, \begin{bmatrix} .10 \\ .40 \\ .50 \end{bmatrix}, \begin{bmatrix} .05 \\ .80 \\ .15 \end{bmatrix}, \begin{bmatrix} .20 \\ .70 \\ .10 \end{bmatrix} \right\}$$



$$l_x := \min_{P(X)\in K(X)} p(x)$$
$$u_x := \min_{P(X)\in K(X)} p(x)$$

these intervals avoid sure loss and are coherent

$$[l_{x'}, u_{x'}] = [.05, .90]$$
$$[l_{x''}, u_{x''}] = [.05, .80]$$
$$[l_{x'''}, u_{x'''}] = [.05, .60]$$

$$K(X) = \mathrm{CH}\left\{ \begin{bmatrix} .05 \\ .35 \\ .60 \end{bmatrix}, \begin{bmatrix} .05 \\ .80 \\ .15 \end{bmatrix}, \begin{bmatrix} .15 \\ .80 \\ .05 \end{bmatrix}, \begin{bmatrix} .35 \\ .05 \\ .60 \end{bmatrix}, \begin{bmatrix} .90 \\ .05 \\ .05 \end{bmatrix} \right\}$$

Learning credal sets from (few) data

- Learning from data about X
- Max lik estimate $P(x) = \frac{n(x)}{N}$
- Bayesian (ESS $s = 2$) $\frac{n(x)+st(x)}{N}$
- Imprecise: set of priors (vacuous t)

$$\frac{n(x)}{N+s} \leq P(x) \leq \frac{n(x)+s}{N+s}$$

imprecise Dirichlet model
(Walley & Bernard)

- They a.s.l. and are coherent
- Non-negligible size of intervals only for small N
  (Bayesian for $N \to \infty$)

Learning credal sets from (few) data

- Learning from data about X
- Max lik estimate $P(x) = \frac{n(x)}{N}$
- Bayesian (ESS $s = 2$) $\frac{n(x)+st(x)}{N}$
- Imprecise: set of priors (vacuous t)

$$\frac{n(x)}{N+s} \leq P(x) \leq \frac{n(x)+s}{N+s}$$

imprecise Dirichlet model
(Walley & Bernard)

- They a.s.l. and are coherent
- Non-negligible size of intervals only for small N
(Bayesian for $N \to \infty$)

$$\begin{array}{c} n(win) \\ n(draw) \\ n(loss) \end{array} = \begin{bmatrix} 4 \\ 1 \\ 3 \end{bmatrix}$$

1957: Spain vs. Italy 5 — 1
1973: Italy vs. Spain 3 — 2
1980: Spain vs. Italy 1 — 0
1983: Spain vs. Italy 1 — 0
1983: Italy vs. Spain 2 — 1
1987: Spain vs. Italy 1 — 1
2000: Spain vs. Italy 1 — 2
2001: Italy vs. Spain 1 — 0

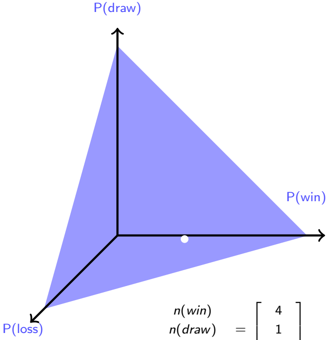
Learning credal sets from (few) data

- Learning from data about X
- Max lik estimate $P(x) = \frac{n(x)}{N}$
- Bayesian (ESS $s = 2$) $\frac{n(x)+st(x)}{N}$
- Imprecise: set of priors (vacuous t)

$$\frac{n(x)}{N+s} \leq P(x) \leq \frac{n(x)+s}{N+s}$$

imprecise Dirichlet model
(Walley & Bernard)

- They a.s.l. and are coherent
- Non-negligible size of intervals only for small N
(Bayesian for $N \to \infty$)



$$\begin{array}{c} n(win) \\ n(draw) \\ n(loss) \end{array} = \begin{bmatrix} 4 \\ 1 \\ 3 \end{bmatrix}$$

1957:	Spain	vs.	Italy	5 — 1
1973:	Italy	vs.	Spain	3 — 2
1980:	Spain	vs.	Italy	1 — 0
1983:	Spain	vs.	Italy	1 — 0
1983:	Italy	vs.	Spain	2 — 1
1987:	Spain	vs.	Italy	1 — 1
2000:	Spain	vs.	Italy	1 — 2
2001:	Italy	vs.	Spain	1 — 0

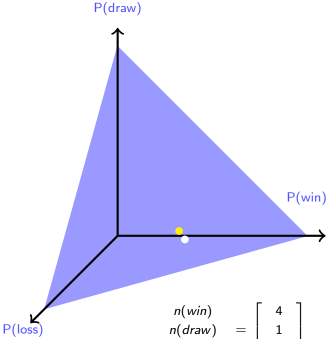
Learning credal sets from (few) data

- Learning from data about X
- Max lik estimate $P(x) = \frac{n(x)}{N}$
- Bayesian (ESS $s = 2$) $\frac{n(x)+st(x)}{N}$
- Imprecise: set of priors (vacuous t)

$$\frac{n(x)}{N+s} \leq P(x) \leq \frac{n(x)+s}{N+s}$$

imprecise Dirichlet model
(Walley & Bernard)

- They a.s.l. and are coherent
- Non-negligible size of intervals only for small N
  (Bayesian for $N \to \infty$)



$$\begin{array}{c} n(win) \\ n(draw) \\ n(loss) \end{array} = \begin{bmatrix} 4 \\ 1 \\ 3 \end{bmatrix}$$

1957:	Spain	vs.	Italy	5 — 1
1973:	Italy	vs.	Spain	3 — 2
1980:	Spain	vs.	Italy	1 — 0
1983:	Spain	vs.	Italy	1 — 0
1983:	Italy	vs.	Spain	2 — 1
1987:	Spain	vs.	Italy	1 — 1
2000:	Spain	vs.	Italy	1 — 2
2001:	Italy	vs.	Spain	1 — 0

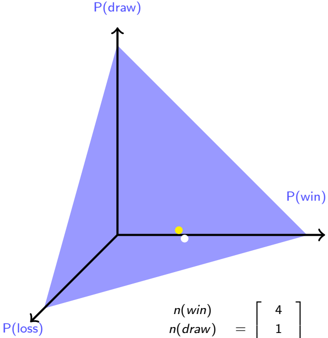
Learning credal sets from (few) data

- Learning from data about X
- Max lik estimate $P(x) = \frac{n(x)}{N}$
- Bayesian (ESS $s = 2$) $\frac{n(x)+st(x)}{N}$
- Imprecise: set of priors (vacuous t)

$$\frac{n(x)}{N+s} \leq P(x) \leq \frac{n(x)+s}{N+s}$$

imprecise Dirichlet model
(Walley & Bernard)

- They a.s.l. and are coherent
- Non-negligible size of intervals only for small N
(Bayesian for $N \to \infty$)



$$\begin{array}{c} n(win) \\ n(draw) \\ n(loss) \end{array} = \begin{bmatrix} 4 \\ 1 \\ 3 \end{bmatrix}$$

1957: Spain vs. Italy 5 — 1
1973: Italy vs. Spain 3 — 2
1980: Spain vs. Italy 1 — 0
1983: Spain vs. Italy 1 — 0
1983: Italy vs. Spain 2 — 1
1987: Spain vs. Italy 1 — 1
2000: Spain vs. Italy 1 — 2
2001: Italy vs. Spain 1 — 0

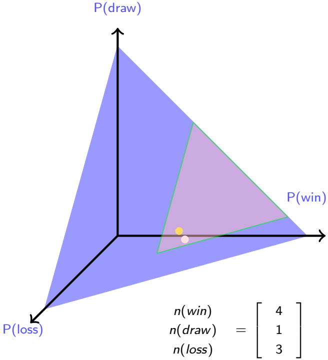
Learning credal sets from (few) data

- Learning from data about X
- Max lik estimate $P(x) = \frac{n(x)}{N}$
- Bayesian (ESS $s = 2$) $\frac{n(x)+st(x)}{N}$
- Imprecise: set of priors (vacuous t)

$$\frac{n(x)}{N+s} \leq P(x) \leq \frac{n(x)+s}{N+s}$$

imprecise Dirichlet model
(Walley & Bernard)

- They a.s.l. and are coherent
- Non-negligible size of intervals only for small N
(Bayesian for $N \to \infty$)



$$\begin{array}{c} n(win) \\ n(draw) \\ n(loss) \end{array} = \begin{bmatrix} 4 \\ 1 \\ 3 \end{bmatrix}$$

1957:	Spain	vs.	Italy	5 — 1
1973:	Italy	vs.	Spain	3 — 2
1980:	Spain	vs.	Italy	1 — 0
1983:	Spain	vs.	Italy	1 — 0
1983:	Italy	vs.	Spain	2 — 1
1987:	Spain	vs.	Italy	1 — 1
2000:	Spain	vs.	Italy	1 — 2
2001:	Italy	vs.	Spain	1 — 0

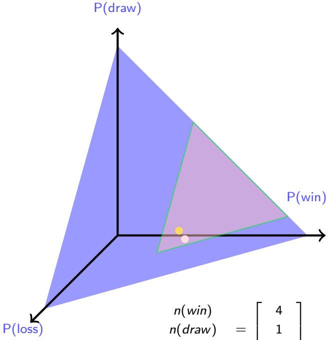
Learning credal sets from (few) data

- Learning from data about X
- Max lik estimate $P(x) = \frac{n(x)}{N}$
- Bayesian (ESS $s = 2$) $\frac{n(x)+st(x)}{N}$
- Imprecise: set of priors (vacuous t)

$$\frac{n(x)}{N+s} \leq P(x) \leq \frac{n(x)+s}{N+s}$$

imprecise Dirichlet model
(Walley & Bernard)

- They a.s.l. and are coherent
- Non-negligible size of intervals only for small N
(Bayesian for $N \to \infty$)



$$\begin{array}{r} n(win) \\ n(draw) \\ n(loss) \end{array} = \begin{bmatrix} 4 \\ 1 \\ 3 \end{bmatrix}$$

1957: Spain vs. Italy 5 — 1
1973: Italy vs. Spain 3 — 2
1980: Spain vs. Italy 1 — 0
1983: Spain vs. Italy 1 — 0
1983: Italy vs. Spain 2 — 1
1987: Spain vs. Italy 1 — 1
2000: Spain vs. Italy 1 — 2
2001: Italy vs. Spain 1 — 0

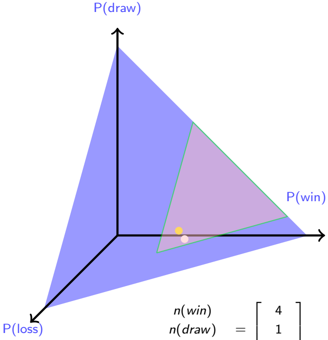
Learning credal sets from (few) data

- Learning from data about X
- Max lik estimate $P(x) = \frac{n(x)}{N}$
- Bayesian (ESS $s = 2$) $\frac{n(x)+st(x)}{N}$
- Imprecise: set of priors (vacuous t)

$$\frac{n(x)}{N+s} \leq P(x) \leq \frac{n(x)+s}{N+s}$$

imprecise Dirichlet model
(Walley & Bernard)

- They a.s.l. and are coherent
- Non-negligible size of intervals only for small N
  (Bayesian for $N \to \infty$)



$$\begin{array}{c} n(win) \\ n(draw) \\ n(loss) \end{array} = \begin{bmatrix} 4 \\ 1 \\ 3 \end{bmatrix}$$

1957: Spain vs. Italy 5 — 1
1973: Italy vs. Spain 3 — 2
1980: Spain vs. Italy 1 — 0
1983: Spain vs. Italy 1 — 0
1983: Italy vs. Spain 2 — 1
1987: Spain vs. Italy 1 — 1
2000: Spain vs. Italy 1 — 2
2001: Italy vs. Spain 1 — 0

Learning credal sets from (missing) data

- Coping with missing data?
- Missing at random (MAR) $P(O = *|X = x)$ indep of X Ignore missing data
- Not always the case!
- Conservative updating (de Cooman & Zaffalon) ignorance about the process $P(O|X)$ as a vacuous model
- Consider all the explanations (and take the convex hull)

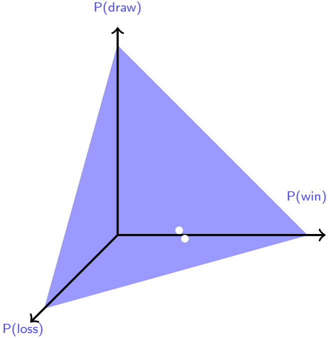
Learning credal sets from (missing) data

- Coping with missing data?
- Missing at random (MAR) $P(O = *|X = x)$ indep of X Ignore missing data
- Not always the case!
- Conservative updating (de Cooman & Zaffalon) ignorance about the process $P(O|X)$ as a vacuous model
- Consider all the explanations (and take the convex hull)

1957:	Spain	vs.	Italy	$5 - 1$
1973:	Italy	vs.	Spain	$3 - 2$
1980:	Spain	vs.	Italy	$1 - 0$
1983:	Spain	vs.	Italy	$1 - 0$
1983:	Italy	vs.	Spain	$2 - 1$
1987:	Spain	vs.	Italy	$1 - 1$
2000:	Spain	vs.	Italy	$1 - 2$
2001:	Italy	vs.	Spain	$1 - 0$
2003:	Spain	vs.	Italy	$* - *$
2011:	Italy	vs.	Spain	$* - *$

Learning credal sets from (missing) data

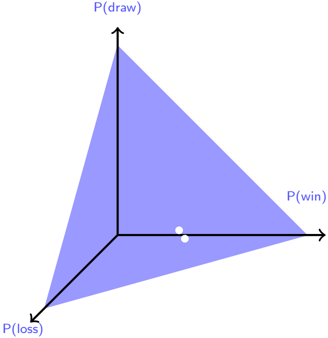
- Coping with missing data?
- Missing at random (MAR)
  $P(O = *|X = x)$ indep of X Ignore
  missing data
- Not always the case!
- Conservative updating
  (de Cooman & Zaffalon) ignorance
  about the process $P(O|X)$ as a
  vacuous model
- Consider all the explanations (and
  take the convex hull)



1957:	Spain	vs.	Italy	5 — 1
1973:	Italy	vs.	Spain	3 — 2
1980:	Spain	vs.	Italy	1 — 0
1983:	Spain	vs.	Italy	1 — 0
1983:	Italy	vs.	Spain	2 — 1
1987:	Spain	vs.	Italy	1 — 1
2000:	Spain	vs.	Italy	1 — 2
2001:	Italy	vs.	Spain	1 — 0
2003:	Spain	vs.	Italy	$*$ — $*$
2011:	Italy	vs.	Spain	$*$ — $*$

Learning credal sets from (missing) data

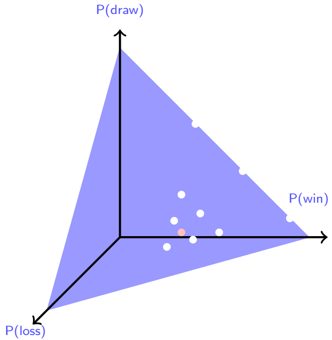
- Coping with missing data?
- Missing at random (MAR) $P(O = *|X = x)$ indep of X Ignore missing data
- Not always the case!
- Conservative updating (de Cooman & Zaffalon) ignorance about the process $P(O|X)$ as a vacuous model
- Consider all the explanations (and take the convex hull)



1957:	Spain	vs.	Italy	5 — 1
1973:	Italy	vs.	Spain	3 — 2
1980:	Spain	vs.	Italy	1 — 0
1983:	Spain	vs.	Italy	1 — 0
1983:	Italy	vs.	Spain	2 — 1
1987:	Spain	vs.	Italy	1 — 1
2000:	Spain	vs.	Italy	1 — 2
2001:	Italy	vs.	Spain	1 — 0
2003:	Spain	vs.	Italy	$*$ — $*$
2011:	Italy	vs.	Spain	$*$ — $*$

Learning credal sets from (missing) data

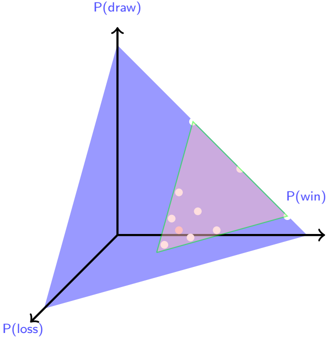
- Coping with missing data?
- Missing at random (MAR) $P(O = *|X = x)$ indep of X Ignore missing data
- Not always the case!
- Conservative updating (de Cooman & Zaffalon) ignorance about the process $P(O|X)$ as a vacuous model
- Consider all the explanations (and take the convex hull)



1957:	Spain	vs.	Italy	5 — 1
1973:	Italy	vs.	Spain	3 — 2
1980:	Spain	vs.	Italy	1 — 0
1983:	Spain	vs.	Italy	1 — 0
1983:	Italy	vs.	Spain	2 — 1
1987:	Spain	vs.	Italy	1 — 1
2000:	Spain	vs.	Italy	1 — 2
2001:	Italy	vs.	Spain	1 — 0
2003:	Spain	vs.	Italy	$*$ — $*$
2011:	Italy	vs.	Spain	$*$ — $*$

Learning credal sets from (missing) data

- Coping with missing data?
- Missing at random (MAR) $P(O = *|X = x)$ indep of X Ignore missing data
- Not always the case!
- Conservative updating (de Cooman & Zaffalon) ignorance about the process $P(O|X)$ as a vacuous model
- Consider all the explanations (and take the convex hull)



1957:	Spain	vs.	Italy	5 — 1
1973:	Italy	vs.	Spain	3 — 2
1980:	Spain	vs.	Italy	1 — 0
1983:	Spain	vs.	Italy	1 — 0
1983:	Italy	vs.	Spain	2 — 1
1987:	Spain	vs.	Italy	1 — 1
2000:	Spain	vs.	Italy	1 — 2
2001:	Italy	vs.	Spain	1 — 0
2003:	Spain	vs.	Italy	$*$ — $*$
2011:	Italy	vs.	Spain	$*$ — $*$

Basic operations with credal sets

	PRECISE Mass functions	IMPRECISE Credal sets
Joint	$P(X,Y)$	$K(X,Y)$
Marginalization	$P(X)$ s.t. $p(x) = \sum_y p(x,y)$	$K(X) =$ $\left\{ P(X) \,\middle\vert\, \begin{array}{l} P(x) = \sum_y P(x,y) \\ P(X,Y) \in K(X,Y) \end{array} \right\}$
Conditioning	$P(X\vert y)$ s.t. $p(x\vert y) = \frac{P(x,y)}{\sum_y P(x,y)}$	$K(X\vert y) =$ $\left\{ P(X\vert y) \,\middle\vert\, \begin{array}{l} P(x\vert y) = \frac{P(x,y)}{\sum_y P(x,y)} \\ P(X,Y) \in K(X,Y) \end{array} \right\}$
Combination	$P(x,y) = P(x\vert y)P(y)$	$K(X\vert Y) \otimes K(Y) =$ $\left\{ P(X,Y) \,\middle\vert\, \begin{array}{l} P(x,y)=P(x\vert y)P(y) \\ P(X\vert y) \in K(X\vert y) \\ P(Y) \in K(Y) \end{array} \right\}$

Basic operations with credal sets

	PRECISE Mass functions	IMPRECISE Credal sets
Joint	$P(X,Y)$	$K(X,Y)$
Marginalization	$P(X)$ s.t. $p(x) = \sum_y p(x,y)$	$K(X) = \left\{ P(X) \middle\vert \begin{array}{l} P(x) = \sum_y P(x,y) \\ P(X,Y) \in K(X,Y) \end{array} \right\}$
Conditioning	$P(X\vert y)$ s.t. $p(x\vert y) = \frac{P(x,y)}{\sum_y P(x,y)}$	$K(X\vert y) = \left\{ P(X\vert y) \middle\vert \begin{array}{l} P(x\vert y) = \frac{P(x,y)}{\sum_y P(x,y)} \\ P(X,Y) \in K(X,Y) \end{array} \right\}$
Combination	$P(x,y) = P(x\vert y)P(y)$	$K(X\vert Y) \otimes K(Y) = \left\{ P(X,Y) \middle\vert \begin{array}{l} P(x,y)=P(x\vert y)P(y) \\ P(X\vert y) \in K(X\vert y) \\ P(Y) \in K(Y) \end{array} \right\}$

Basic operations with credal sets

	PRECISE Mass functions	IMPRECISE Credal sets
Joint	$P(X,Y)$	$K(X,Y)$
Marginalization	$P(X)$ s.t. $p(x) = \sum_y p(x,y)$	$K(X) = \left\{ P(X) \,\middle\vert\, \begin{array}{l} P(x) = \sum_y P(x,y) \\ P(X,Y) \in K(X,Y) \end{array} \right\}$
Conditioning	$P(X\vert y)$ s.t. $p(x\vert y) = \frac{P(x,y)}{\sum_y P(x,y)}$	$K(X\vert y) = \left\{ P(X\vert y) \,\middle\vert\, \begin{array}{l} P(x\vert y) = \frac{P(x,y)}{\sum_y P(x,y)} \\ P(X,Y) \in K(X,Y) \end{array} \right\}$
Combination	$P(x,y) = P(x\vert y)P(y)$	$K(X\vert Y) \otimes K(Y) = \left\{ P(X,Y) \,\middle\vert\, \begin{array}{l} P(x,y) = P(x\vert y)P(y) \\ P(X\vert y) \in K(X\vert y) \\ P(Y) \in K(Y) \end{array} \right\}$

Basic operations with credal sets

	PRECISE Mass functions	IMPRECISE Credal sets
Joint	$P(X,Y)$	$K(X,Y)$
Marginalization	$P(X)$ s.t. $p(x) = \sum_y p(x,y)$	$K(X) = \left\{ P(X) \,\middle\vert\, \begin{array}{l} P(x) = \sum_y P(x,y) \\ P(X,Y) \in K(X,Y) \end{array} \right\}$
Conditioning	$P(X\vert y)$ s.t. $p(x\vert y) = \frac{P(x,y)}{\sum_y P(x,y)}$	$K(X\vert y) = \left\{ P(X\vert y) \,\middle\vert\, \begin{array}{l} P(x\vert y) = \frac{P(x,y)}{\sum_y P(x,y)} \\ P(X,Y) \in K(X,Y) \end{array} \right\}$
Combination	$P(x,y) = P(x\vert y)P(y)$	$K(X\vert Y) \otimes K(Y) = \left\{ P(X,Y) \,\middle\vert\, \begin{array}{l} P(x,y) = P(x\vert y)P(y) \\ P(X\vert y) \in K(X\vert y) \\ P(Y) \in K(Y) \end{array} \right\}$

Basic operations with credal sets

	PRECISE Mass functions	IMPRECISE Credal sets
Joint	$P(X,Y)$	$K(X,Y)$
Marginalization	$P(X)$ s.t. $p(x) = \sum_y p(x,y)$	$K(X) =$ $\left\{ P(X) \,\middle\vert\, \begin{array}{l} P(x) = \sum_y P(x,y) \\ P(X,Y) \in K(X,Y) \end{array} \right\}$
Conditioning	$P(X\vert y)$ s.t. $p(x\vert y) = \frac{P(x,y)}{\sum_y P(x,y)}$	$K(X\vert y) =$ $\left\{ P(X\vert y) \,\middle\vert\, \begin{array}{l} P(x\vert y) = \frac{P(x,y)}{\sum_y P(x,y)} \\ P(X,Y) \in K(X,Y) \end{array} \right\}$
Combination	$P(x,y) = P(x\vert y)P(y)$	$K(X\vert Y) \otimes K(Y) =$ $\left\{ P(X,Y) \,\middle\vert\, \begin{array}{l} P(x,y)=P(x\vert y)P(y) \\ P(X\vert y) \in K(X\vert y) \\ P(Y) \in K(Y) \end{array} \right\}$

Basic operations with credal sets

	PRECISE Mass functions	IMPRECISE Credal sets
Joint	$P(X, Y)$	$K(X, Y)$
Marginalization	$P(X)$ s.t. $p(x) = \sum_y p(x, y)$	$K(X) =$ $\left\{ P(X) \middle\vert \begin{array}{l} P(x) = \sum_y P(x, y) \\ P(X, Y) \in K(X, Y) \end{array} \right\}$
Conditioning	$P(X\vert y)$ s.t. $p(x\vert y) = \frac{P(x,y)}{\sum_y P(x,y)}$	$K(X\vert y) =$ $\left\{ P(X\vert y) \middle\vert \begin{array}{l} P(x\vert y) = \frac{P(x,y)}{\sum_y P(x,y)} \\ P(X, Y) \in K(X, Y) \end{array} \right\}$
Combination	$P(x, y) = P(x\vert y)P(y)$	$K(X\vert Y) \otimes K(Y) =$ $\left\{ P(X, Y) \middle\vert \begin{array}{l} P(x, y) = P(x\vert y)P(y) \\ P(X\vert y) \in K(X\vert y) \\ P(Y) \in K(Y) \end{array} \right\}$

Basic operations with credal sets

	PRECISE Mass functions	IMPRECISE Credal sets
Joint	$P(X, Y)$	$K(X, Y)$
Marginalization	$P(X)$ s.t. $p(x) = \sum_y p(x, y)$	$K(X) =$ $\left\{ P(X) \,\middle\vert\, \begin{array}{l} P(x) = \sum_y P(x, y) \\ P(X, Y) \in K(X, Y) \end{array} \right\}$
Conditioning	$P(X \vert y)$ s.t. $p(x \vert y) = \frac{P(x,y)}{\sum_y P(x,y)}$	$K(X \vert y) =$ $\left\{ P(X \vert y) \,\middle\vert\, \begin{array}{l} P(x \vert y) = \frac{P(x,y)}{\sum_y P(x,y)} \\ P(X, Y) \in K(X, Y) \end{array} \right\}$
Combination	$P(x, y) = P(x \vert y) P(y)$	$K(X \vert Y) \otimes K(Y) =$ $\left\{ P(X, Y) \,\middle\vert\, \begin{array}{l} P(x, y) = P(x \vert y) P(y) \\ P(X \vert y) \in K(X \vert y) \\ P(Y) \in K(Y) \end{array} \right\}$

Basic operations with credal sets

	PRECISE Mass functions	IMPRECISE Credal sets
Joint	$P(X,Y)$	$K(X,Y)$
Marginalization	$P(X)$ s.t. $p(x) = \sum_y p(x,y)$	$K(X) = \left\{ P(X) \middle\vert \begin{array}{l} P(x) = \sum_y P(x,y) \\ P(X,Y) \in K(X,Y) \end{array} \right\}$
Conditioning	$P(X\vert y)$ s.t. $p(x\vert y) = \frac{P(x,y)}{\sum_y P(x,y)}$	$K(X\vert y) = \left\{ P(X\vert y) \middle\vert \begin{array}{l} P(x\vert y) = \frac{P(x,y)}{\sum_y P(x,y)} \\ P(X,Y) \in K(X,Y) \end{array} \right\}$
Combination	$P(x,y) = P(x\vert y)P(y)$	$K(X\vert Y) \otimes K(Y) = \left\{ P(X,Y) \middle\vert \begin{array}{l} P(x,y)=P(x\vert y)P(y) \\ P(X\vert y) \in K(X\vert y) \\ P(Y) \in K(Y) \end{array} \right\}$

Basic operations with credal sets

	PRECISE Mass functions	IMPRECISE Credal sets
Joint	$P(X, Y)$	$K(X, Y)$
Marginalization	$P(X)$ s.t. $p(x) = \sum_y p(x, y)$	$K(X) = \left\{ P(X) \,\middle\vert\, \begin{array}{l} P(x) = \sum_y P(x, y) \\ P(X, Y) \in K(X, Y) \end{array} \right\}$
Conditioning	$P(X\vert y)$ s.t. $p(x\vert y) = \frac{P(x,y)}{\sum_y P(x,y)}$	$K(X\vert y) = \left\{ P(X\vert y) \,\middle\vert\, \begin{array}{l} P(x\vert y) = \frac{P(x,y)}{\sum_y P(x,y)} \\ P(X, Y) \in K(X, Y) \end{array} \right\}$
Combination	$P(x, y) = P(x\vert y)P(y)$	$K(X\vert Y) \otimes K(Y) = \left\{ P(X, Y) \,\middle\vert\, \begin{array}{l} P(x, y){=}P(x\vert y)P(y) \\ P(X\vert y) \in K(X\vert y) \\ P(Y) \in K(Y) \end{array} \right\}$

Basic operations with credal sets (vertices)

IMPRECISE
Credal sets

IMPRECISE
Extremes

Joint

$$K(X,Y)$$

Marginalization

$$K(X) \quad = \left\{ P(X) \,\middle|\, \begin{array}{l} P(x) = \sum_y P(x,y) \\ P(X,Y) \in K(X,Y) \end{array} \right\}$$

Conditioning

$$K(X|y) \quad = \left\{ P(X|y) \,\middle|\, \begin{array}{l} P(x|y) = \frac{P(x,y)}{\sum_y P(x,y)} \\ P(X,Y) \in K(X,Y) \end{array} \right\}$$

Combination

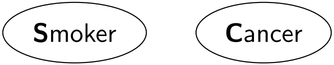
$$K(X|Y) \otimes K(Y) \quad = \left\{ P(X,Y) \,\middle|\, \begin{array}{l} P(x,y){=}P(x|y)P(y) \\ P(X|y) \in K(X|y) \\ P(Y) \in K(Y) \end{array} \right\}$$

Basic operations with credal sets (vertices)

	IMPRECISE Credal sets	IMPRECISE Extremes
Joint	$K(X,Y)$	$= \mathrm{CH}\left\{P_j(X,Y)\right\}_{j=1}^{n_v}$
Marginalization	$K(X) = \left\{P(X) \,\middle\vert\, \begin{array}{l} P(x)=\sum_y P(x,y) \\ P(X,Y) \in K(X,Y) \end{array}\right\}$	$= \mathrm{CH}\left\{P(X) \,\middle\vert\, \begin{array}{l} P(x)=\sum_y P(x,y) \\ P(X,Y) \in \mathrm{ext}[K(X,Y)] \end{array}\right\}$
Conditioning	$K(X\vert y) = \left\{P(X\vert y) \,\middle\vert\, \begin{array}{l} P(x\vert y)=\frac{P(x,y)}{\sum_y P(x,y)} \\ P(X,Y) \in K(X,Y) \end{array}\right\}$	$= \mathrm{CH}\left\{P(X\vert y) \,\middle\vert\, \begin{array}{l} P(x\vert y)=\frac{P(x,y)}{\sum_y P(x,y)} \\ P(X,Y) \in \mathrm{ext}[K(X,Y)] \end{array}\right\}$
Combination	$K(X\vert Y) \otimes K(Y) = \left\{P(X,Y) \,\middle\vert\, \begin{array}{l} P(x,y)=P(x\vert y)P(y) \\ P(X\vert y) \in K(X\vert y) \\ P(Y) \in K(Y) \end{array}\right\}$	$= \mathrm{CH}\left\{P(X,Y) \,\middle\vert\, \begin{array}{l} P(x,y)=P(x\vert y)P(y) \\ P(X\vert y) \in \mathrm{ext}[K(X\vert y)] \\ P(Y) \in \mathrm{ext}[K(Y)] \end{array}\right\}$

An imprecise bivariate (graphical?) model

- Two Boolean variables: **S**moker, Lung **C**ancer



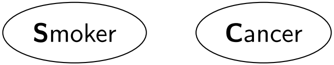
Smoker

Cancer

An imprecise bivariate (graphical?) model

- Two Boolean variables: **S**moker, Lung **C**ancer

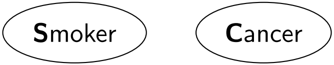
j	$P_j(s, c)$	$P_j(s, \neg c)$	$P_j(\neg s, c)$	$P_j(\neg s, \neg c)$
1	1/8	1/8	3/8	3/8
2	1/8	1/8	9/16	3/16
3	3/16	1/16	3/8	3/8
4	3/16	1/16	9/16	3/16
5	1/4	1/4	1/4	1/4
6	1/4	1/4	3/8	1/8
7	3/8	1/8	1/4	1/4
8	3/8	1/8	3/8	1/8



An imprecise bivariate (graphical?) model

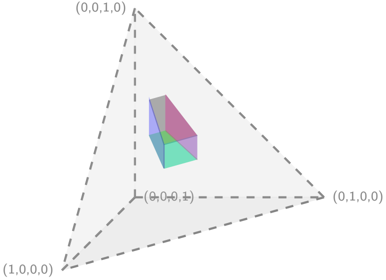
- Two Boolean variables: **S**moker, Lung **C**ancer

j	$P_j(s,c)$	$P_j(s,\neg c)$	$P_j(\neg s,c)$	$P_j(\neg s,\neg c)$
1	1/8	1/8	3/8	3/8
2	1/8	1/8	9/16	3/16
3	3/16	1/16	3/8	3/8
4	3/16	1/16	9/16	3/16
5	1/4	1/4	1/4	1/4
6	1/4	1/4	3/8	1/8
7	3/8	1/8	1/4	1/4
8	3/8	1/8	3/8	1/8



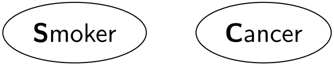
Smoker

Cancer

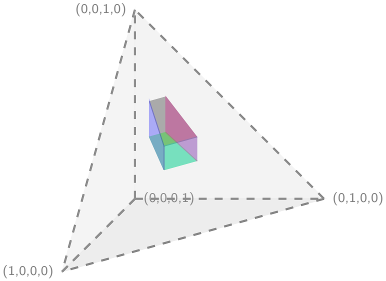


An imprecise bivariate (graphical?) model

- Two Boolean variables: **S**moker, Lung **C**ancer
- Compute:
  - Marginal $K(S)$
  - Conditioning $K(C|S) := \{K(C|s), K(C|s)\}$
  - Combination (marg ext) $K'(C, S) := K(C|S) \otimes K(S)$

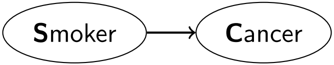


j	$P_j(s, c)$	$P_j(s, \neg c)$	$P_j(\neg s, c)$	$P_j(\neg s, \neg c)$
1	1/8	1/8	3/8	3/8
2	1/8	1/8	9/16	3/16
3	3/16	1/16	3/8	3/8
4	3/16	1/16	9/16	3/16
5	1/4	1/4	1/4	1/4
6	1/4	1/4	3/8	1/8
7	3/8	1/8	1/4	1/4
8	3/8	1/8	3/8	1/8

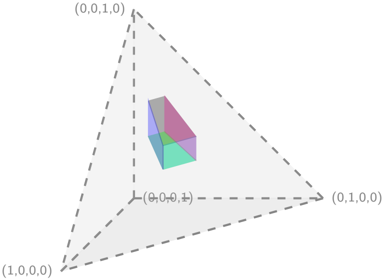


An imprecise bivariate (graphical?) model

- Two Boolean variables: **S**moker, Lung **C**ancer
- Compute:
  - Marginal $K(S)$
  - Conditioning $K(C|S) := \{K(C|s), K(C|s)\}$
  - Combination (marg ext) $K'(C,S) := K(C|S) \otimes K(S)$
- Is this a (I)PGM?

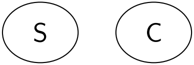


j	$P_j(s, c)$	$P_j(s, \neg c)$	$P_j(\neg s, c)$	$P_j(\neg s, \neg c)$
1	1/8	1/8	3/8	3/8
2	1/8	1/8	9/16	3/16
3	3/16	1/16	3/8	3/8
4	3/16	1/16	9/16	3/16
5	1/4	1/4	1/4	1/4
6	1/4	1/4	3/8	1/8
7	3/8	1/8	1/4	1/4
8	3/8	1/8	3/8	1/8



Cano-Cano-Moral Transformation

Cano-Cano-Moral Transformation



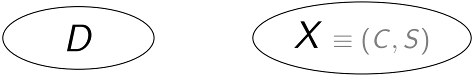
Cano-Cano-Moral Transformation

- Joint variable $X := (C, S)$, $K(X) = \{P_j(X)\}_{j=1}^{n_v}$ ($|\mathcal{X} = 4|$ and $n_v = 8$)

$$X \equiv (C, S)$$

Cano-Cano-Moral Transformation

- Joint variable $X := (C, S)$, $K(X) = \{P_j(X)\}_{j=1}^{n_v}$ ($|\mathcal{X} = 4|$ and $n_v = 8$)
- Auxiliary variable D to enumerate elements of $\mathrm{ext}[K(X)]$



Cano-Cano-Moral Transformation

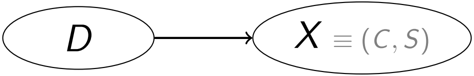
- Joint variable $X := (C, S)$, $K(X) = \{P_j(X)\}_{j=1}^{n_v}$ $(|\mathcal{X} = 4|$ and $n_v = 8)$
- Auxiliary variable D to enumerate elements of $\mathrm{ext}[K(X)]$
  - Precise $K(X|D)$, with $P(X|d_j) := P_j(X)$ with $|\mathcal{D}| = 8$

Cano-Cano-Moral Transformation

- Joint variable $X := (C, S)$, $K(X) = \{P_j(X)\}_{j=1}^{n_v}$ $(|\mathcal{X} = 4|$ and $n_v = 8)$
- Auxiliary variable D to enumerate elements of $\mathrm{ext}[K(X)]$
  - Precise $K(X|D)$, with $P(X|d_j) := P_j(X)$ with $|\mathcal{D}| = 8$
  - Vacuous $K(D)$ (8 vertices)

Cano-Cano-Moral Transformation

- Joint variable $X := (C, S)$, $K(X) = \{P_j(X)\}_{j=1}^{n_v}$ $(|\mathcal{X} = 4|$ and $n_v = 8)$
- Auxiliary variable D to enumerate elements of $\mathrm{ext}[K(X)]$
  - Precise $K(X|D)$, with $P(X|d_j) := P_j(X)$ with $|\mathcal{D}| = 8$
  - Vacuous $K(D)$ (8 vertices)
- Marginal extension $K(X, D) = K(X|D) \otimes K(D)$

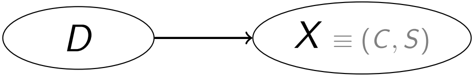


unconditional
vacuous

conditional
precise

Cano-Cano-Moral Transformation

- Joint variable $X := (C, S)$, $K(X) = \{P_j(X)\}_{j=1}^{n_v}$ ($|\mathcal{X} = 4|$ and $n_v = 8$)
- Auxiliary variable D to enumerate elements of $\mathrm{ext}[K(X)]$
  - Precise $K(X|D)$, with $P(X|d_j) := P_j(X)$ with $|\mathcal{D}| = 8$
  - Vacuous $K(D)$ (8 vertices)
- Marginal extension $K(X, D) = K(X|D) \otimes K(D)$
- After marginalization of D, $K(X, D)$ becomes $K(X)$



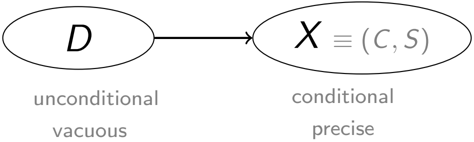
D $\rightarrow$ $X \equiv (C, S)$

unconditional
vacuous

conditional
precise

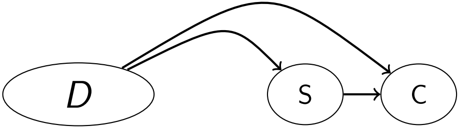
Cano-Cano-Moral Transformation

- Joint variable $X := (C, S)$, $K(X) = \{P_j(X)\}_{j=1}^{n_v}$ $(|\mathcal{X} = 4|$ and $n_v = 8)$
- Auxiliary variable D to enumerate elements of $\mathrm{ext}[K(X)]$
  - Precise $K(X|D)$, with $P(X|d_j) := P_j(X)$ with $|\mathcal{D}| = 8$
  - Vacuous $K(D)$ (8 vertices)
- Marginal extension $K(X, D) = K(X|D) \otimes K(D)$
- After marginalization of D, $K(X, D)$ becomes $K(X)$
- IP models as combination of precise models and vacuous priors
- Hyerarchical model with a vacuous second order knowledge



Cano-Cano-Moral Transformation

- Joint variable $X := (C, S)$, $K(X) = \{P_j(X)\}_{j=1}^{n_v}$ ($|\mathcal{X} = 4|$ and $n_v = 8$)
- Auxiliary variable D to enumerate elements of $\mathrm{ext}[K(X)]$
  - Precise $K(X|D)$, with $P(X|d_j) := P_j(X)$ with $|\mathcal{D}| = 8$
  - Vacuous $K(D)$ (8 vertices)
- Marginal extension $K(X, D) = K(X|D) \otimes K(D)$
- After marginalization of D, $K(X, D)$ becomes $K(X)$
- IP models as combination of precise models and vacuous priors
- Hyerarchical model with a vacuous second order knowledge



unconditional
vacuous

conditional
precise

Probabilistic Graphical Models

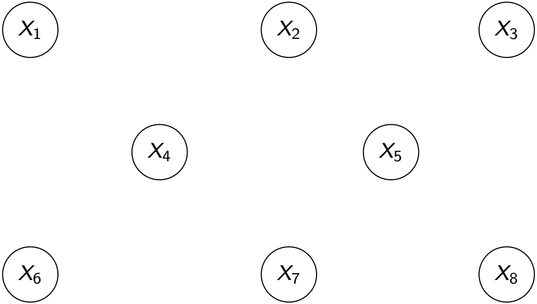
aka Decomposable Multivariate Probabilistic Models

(whose decomposability is induced by independence)

Probabilistic Graphical Models

aka Decomposable Multivariate Probabilistic Models

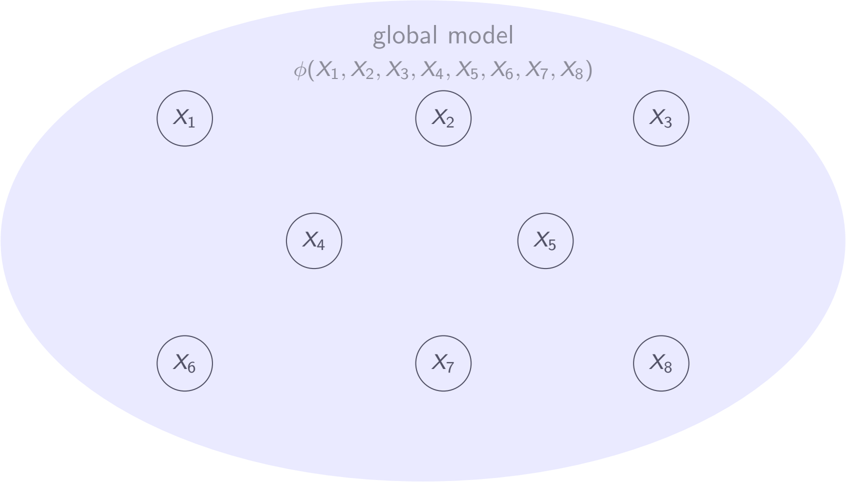
(whose decomposability is induced by independence)



Probabilistic Graphical Models

aka Decomposable Multivariate Probabilistic Models

(whose decomposability is induced by independence)

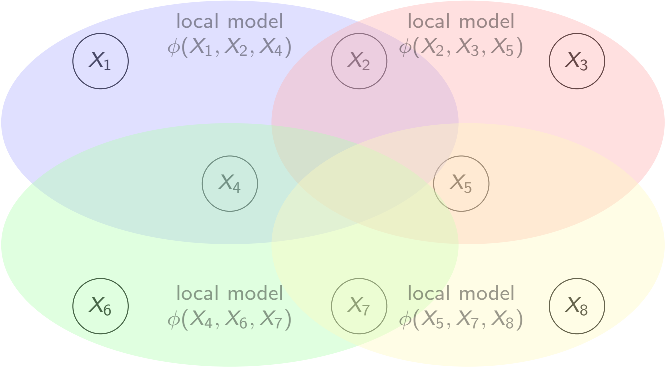


Probabilistic Graphical Models

aka Decomposable Multivariate Probabilistic Models

(whose decomposability is induced by independence)

$$\phi(X_1, X_2, X_3, X_4, X_5, X_6, X_7, X_8) = \phi(X_1, X_2, X_4) \otimes \phi(X_2, X_3, X_5) \otimes \phi(X_4, X_6, X_7) \otimes \phi(X_5, X_7, X_8)$$



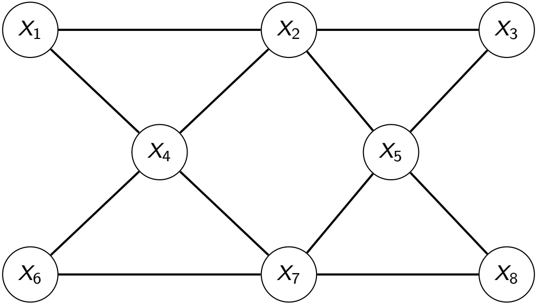
Probabilistic Graphical Models

aka Decomposable Multivariate Probabilistic Models

(whose decomposability is induced by independence)

undirected graphs

precise/imprecise Markov random fields



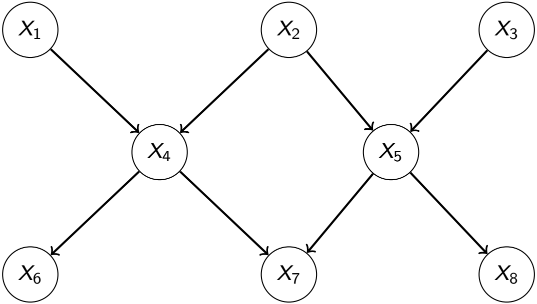
Probabilistic Graphical Models

aka Decomposable Multivariate Probabilistic Models

(whose decomposability is induced by independence)

directed graphs

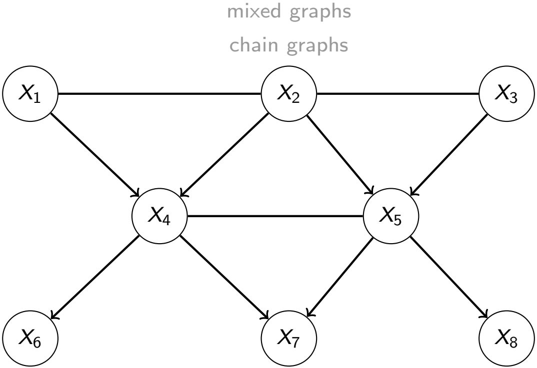
Bayesian/credal networks



Probabilistic Graphical Models

aka Decomposable Multivariate Probabilistic Models

(whose decomposability is induced by independence)



Independence

Independence

Stochastic independence/irrelevance (precise case)

- X and Y stochastically independent: $P(x, y) = P(x)P(y)$
- Y stochastically irrelevant to X: $P(X|y) = P(X)$
- independence $\equiv$ irrelevance

Independence

Stochastic independence/irrelevance (precise case)

- X and Y stochastically independent: $P(x, y) = P(x)P(y)$
- Y stochastically irrelevant to X: $P(X|y) = P(X)$
- independence $\equiv$ irrelevance

Strong independence (imprecise case)

- X and Y strongly independent: stochastic independence $\forall P(X, Y) \in \mathrm{ext}[K(X, Y)]$
- Equivalent to Y strongly irrelevant to X: $P(X|y) = P(X)$ $\forall P(X, Y) \in \mathrm{ext}[K(X, Y)]$

Independence

Stochastic independence/irrelevance (precise case)

- X and Y stochastically independent: $P(x,y) = P(x)P(y)$
- Y stochastically irrelevant to X: $P(X|y) = P(X)$
- independence $\equiv$ irrelevance

Strong independence (imprecise case)

- X and Y strongly independent: stochastic independence $\forall P(X,Y) \in \mathrm{ext}[K(X,Y)]$
- Equivalent to Y strongly irrelevant to X: $P(X|y) = P(X)$ $\forall P(X,Y) \in \mathrm{ext}[K(X,Y)]$

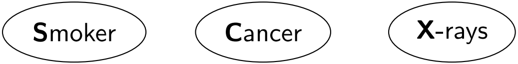
Epistemic irrelevance (imprecise case)

- Y epistemically irrelevant to X: $K(X|y) = K(X)$
- Asymmetric concept! Its simmetrization: epistemic indep

Every notion of independence/irrelevance admits a conditional

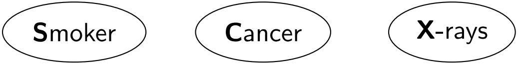
A tri-variate example

- 3 Boolean variables: **S**moker, Lung **C**ancer, **X**-rays



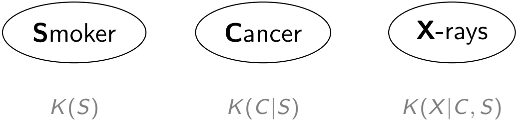
A tri-variate example

- 3 Boolean variables: **S**moker, Lung **C**ancer, **X**-rays
- Given cancer, no relation between smoker and X-rays
- IP language: given C, S and X strongly independent



A tri-variate example

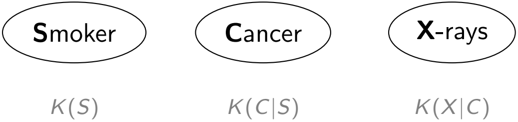
- 3 Boolean variables: **S**moker, Lung **C**ancer, **X**-rays
- Given cancer, no relation between smoker and X-rays
- IP language: given C, S and X strongly independent
- Marginal extension (iterated two times)
  $K(S,C,X) = K(X|C,S) \otimes K(C,S) = K(X|C,S) \otimes K(C|S) \otimes K(S)$



A tri-variate example

- 3 Boolean variables: **S**moker, Lung **C**ancer, **X**-rays
- Given cancer, no relation between smoker and X-rays
- IP language: given C, S and X strongly independent
- Marginal extension (iterated two times)
  $K(S,C,X) = K(X|C,S) \otimes K(C,S) = K(X|C,S) \otimes K(C|S) \otimes K(S)$
- Independence implies irrelevance: given C, S irrelevant to X

$$K(S,C,X) = K(X|C) \otimes K(C|S) \otimes K(S)$$



Smoker

Cancer

X-rays

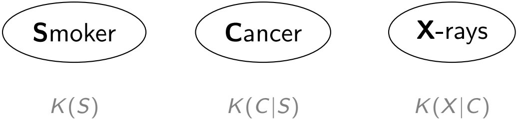
$K(S)$

$K(C|S)$

$K(X|C)$

A tri-variate example

- 3 Boolean variables: **S**moker, Lung **C**ancer, **X**-rays
- Given cancer, no relation between smoker and X-rays
- IP language: given C, S and X strongly independent
- Marginal extension (iterated two times)
  $K(S, C, X) = K(X|C, S) \otimes K(C, S) = K(X|C, S) \otimes K(C|S) \otimes K(S)$
- Independence implies irrelevance: given C, S irrelevant to X
  $$K(S, C, X) = K(X|C) \otimes K(C|S) \otimes K(S)$$
- Global model decomposed in 3 "local" models



Smoker

Cancer

X-rays

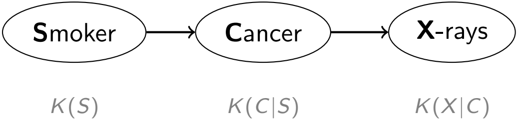
$K(S)$

$K(C|S)$

$K(X|C)$

A tri-variate example

- 3 Boolean variables: **S**moker, Lung **C**ancer, **X**-rays
- Given cancer, no relation between smoker and X-rays
- IP language: given C, S and X strongly independent
- Marginal extension (iterated two times)
  $K(S,C,X) = K(X|C,S) \otimes K(C,S) = K(X|C,S) \otimes K(C|S) \otimes K(S)$
- Independence implies irrelevance: given C, S irrelevant to X
  $$K(S,C,X) = K(X|C) \otimes K(C|S) \otimes K(S)$$
- Global model decomposed in 3 “local” models
- A true PGM! Needed: language to express independencies



Markov Condition

- Probabilistic model over set of variables $(X_1, \ldots, X_n)$ in one-to-one correspondence with the nodes of a graph

Markov Condition

- Probabilistic model over set of variables $(X_1, \ldots, X_n)$ in one-to-one correspondence with the nodes of a graph

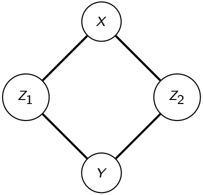
Undirected Graphs

X and Y are independent given Z
if any path between X and Y
containts an element of Z

Markov Condition

- Probabilistic model over set of variables $(X_1, \ldots, X_n)$ in one-to-one correspondence with the nodes of a graph

Undirected Graphs



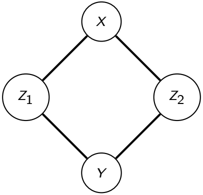
X and Y are independent given Z
if any path between X and Y
containts an element of Z

Markov Condition

- Probabilistic model over set of variables $(X_1, \ldots, X_n)$ in one-to-one correspondence with the nodes of a graph

Undirected Graphs

X and Y are independent given Z
if any path between X and Y
containts an element of Z



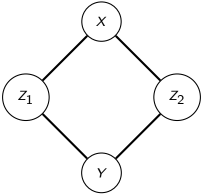
Directed Graphs

Given its parents, every node is independent of its
non-descendants non-parents

Markov Condition

- Probabilistic model over set of variables $(X_1, \ldots, X_n)$ in one-to-one correspondence with the nodes of a graph

Undirected Graphs



X and Y are independent given Z
if any path between X and Y
containts an element of Z

Directed Graphs

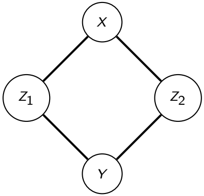
Given its parents, every node is independent of its non-descendants non-parents

X and Y are d-separated by Z if, along every path between X and Y there is a W such that either W has converging arrows and is not in Z and none of its descendants are in Z, or W has no converging arrows and is in Z

Markov Condition

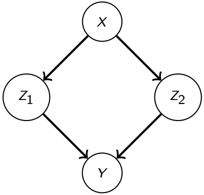
- Probabilistic model over set of variables $(X_1, \ldots, X_n)$ in one-to-one correspondence with the nodes of a graph

Undirected Graphs



X and Y are independent given Z
if any path between X and Y
containts an element of Z

Directed Graphs



Given its parents, every node is independent of its non-descendants non-parents

X and Y are d-separated by Z if, along every path between X and Y there is a W such that either W has converging arrows and is not in Z and none of its descendants are in Z, or W has no converging arrows and is in Z

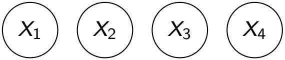
Bayesian networks *(Pearl, 1986)*

Bayesian networks *(Pearl, 1986)*

- Set of categorical variables $X_1, \ldots, X_n$

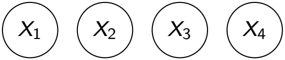
Bayesian networks *(Pearl, 1986)*

- Set of categorical variables $X_1, \ldots, X_n$



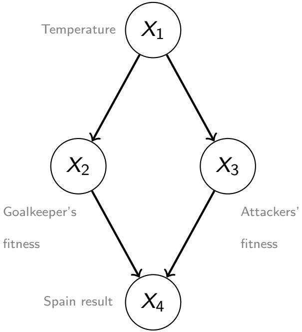
Bayesian networks *(Pearl, 1986)*

- Set of categorical variables $X_1, \ldots, X_n$
- Directed acyclic graph



Bayesian networks *(Pearl, 1986)*

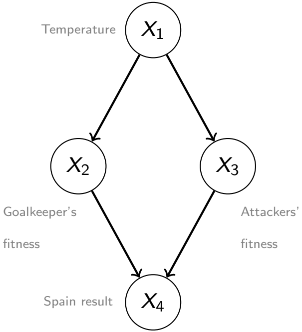
- Set of categorical variables $X_1, \ldots, X_n$
- Directed acyclic graph



Bayesian networks *(Pearl, 1986)*

- Set of categorical variables $X_1, \ldots, X_n$
- Directed acyclic graph
  - conditional (stochastic) independencies according to the Markov condition:

    "any node is conditionally independent of its non-descendents given its parents"

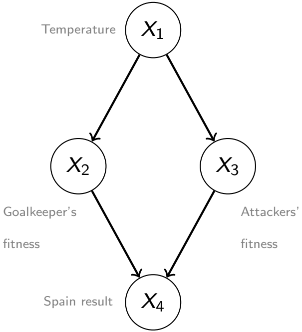


E.g., given temperature, fitnesses independent

Bayesian networks *(Pearl, 1986)*

- Set of categorical variables $X_1, \ldots, X_n$
- Directed acyclic graph
  - conditional (stochastic) independencies according to the Markov condition:

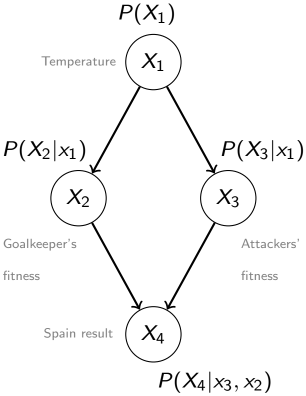
    "any node is conditionally independent of its non-descendents given its parents"
- A conditional mass function for each node and each possible value of the parents



Bayesian networks *(Pearl, 1986)*

- Set of categorical variables $X_1, \dots, X_n$
- Directed acyclic graph
  - conditional (stochastic) independencies according to the Markov condition:

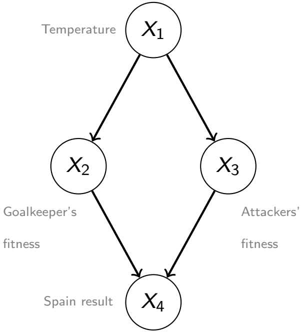
    “any node is conditionally independent of its non-descendents given its parents”
- A conditional mass function for each node and each possible value of the parents
  - $\{P(X_i|\mathrm{pa}(X_i)) \ , \forall i = 1, \dots, n \ , \forall \mathrm{pa}(X_i) \ \}$



Bayesian networks *(Pearl, 1986)*

- Set of categorical variables $X_1, \ldots, X_n$
- Directed acyclic graph
  - conditional (stochastic) independencies according to the Markov condition:

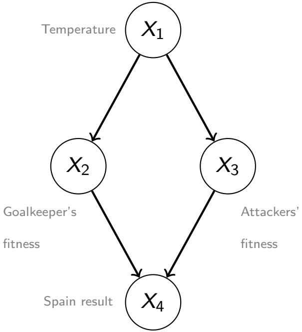
    "any node is conditionally independent of its non-descendents given its parents"
- A conditional mass function for each node and each possible value of the parents
  - $\{P(X_i|\mathrm{pa}(X_i))\ , \forall i = 1, \ldots, n\ , \ \forall \mathrm{pa}(X_i)\ \}$
- Defines a joint probability mass function



Bayesian networks *(Pearl, 1986)*

- Set of categorical variables $X_1, \ldots, X_n$
- Directed acyclic graph
  - conditional (stochastic) independencies according to the Markov condition:

    "any node is conditionally independent of its non-descendents given its parents"
- A conditional mass function for each node and each possible value of the parents
  - $\{P(X_i|\mathrm{pa}(X_i)) \ , \forall i = 1, \ldots, n \ , \ \forall \mathrm{pa}(X_i) \ \}$
- Defines a joint probability mass function
  - $P(x_1, \ldots, x_n) = \prod_{i=1}^{n} P(x_i|\mathrm{pa}(X_i))$



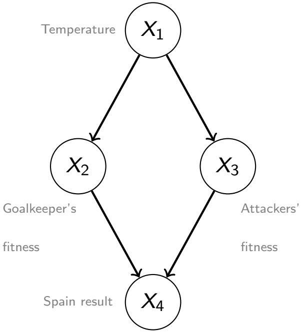
$$P(x_1, x_2, x_3, x_4) = P(x_1)P(x_2|x_1)P(x_3|x_1)P(x_4|x_3, x_2)$$

Credal networks *(Cozman, 2000)*

- Generalization of BNs to imprecise probabilities

Credal networks *(Cozman, 2000)*

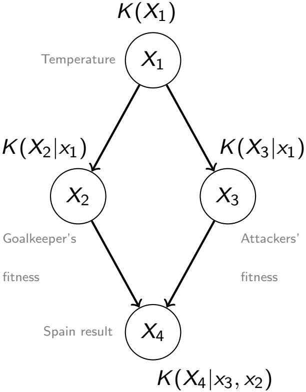
- Generalization of BNs to imprecise probabilities



Credal networks *(Cozman, 2000)*

- Generalization of BNs to imprecise probabilities
- Credal sets instead of prob mass functions

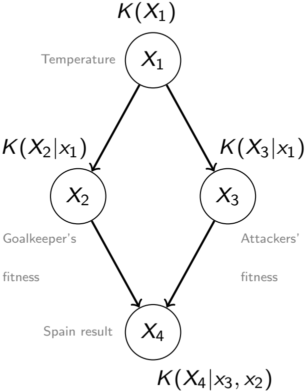
  $\{P(X_i|\mathrm{pa}(X_i))\} \Rightarrow \{K(X_i|\mathrm{pa}(X_i))\}$



E.g., $K(X_1)$ defined by constraint $P(x_1) > .75$, very likely to be warm

Credal networks *(Cozman, 2000)*

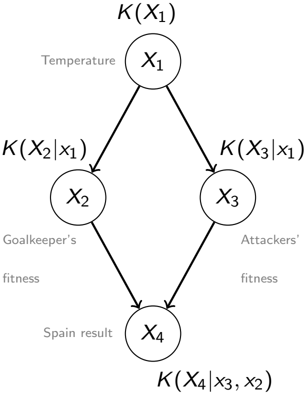
- Generalization of BNs to imprecise probabilities
- Credal sets instead of prob mass functions
  $\{P(X_i|\mathrm{pa}(X_i))\} \Rightarrow \{K(X_i|\mathrm{pa}(X_i))\}$
- Strong (instead of stochastic) independence in the semantics of the Markov condition



E.g., $K(X_1)$ defined by constraint $P(x_1) > .75$, very likely to be warm

Credal networks *(Cozman, 2000)*

- Generalization of BNs to imprecise probabilities
- Credal sets instead of prob mass functions
  $\{P(X_i|\mathrm{pa}(X_i))\} \Rightarrow \{K(X_i|\mathrm{pa}(X_i))\}$
- Strong (instead of stochastic) independence in the semantics of the Markov condition
- Convex set of joint mass functions
  $K(X_1,\ldots,X_n) = \mathrm{CH}\Big\{P(X_1,\ldots,X_n)\Big\}$
  $P(x_1,\ldots,x_n) = \prod_{i=1}^n P(x_i|\mathrm{pa}(X_i))$
  $\forall P(X_i|\mathrm{pa}(X_i)) \in K(X_i|\mathrm{pa}(X_i))$
  $\forall i = 1,\ldots,n \quad \forall \mathrm{pa}(X_i)$



E.g., $K(X_1)$ defined by constraint $P(x_1) > .75$, very likely to be warm

Credal networks (Cozman, 2000)

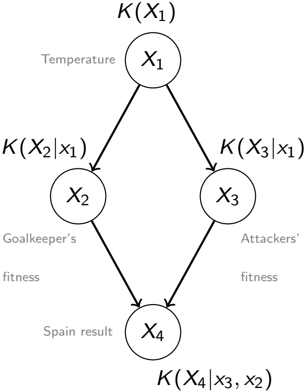
- Generalization of BNs to imprecise probabilities

- Credal sets instead of prob mass functions
  $\{P(X_i|\mathrm{pa}(X_i))\} \Rightarrow \{K(X_i|\mathrm{pa}(X_i))\}$

- Strong (instead of stochastic) independence in the semantics of the Markov condition

- Convex set of joint mass functions
  $$K(X_1,\ldots,X_n) = \mathrm{CH}\Big\{P(X_1,\ldots,X_n)\Big\}$$
  $$P(x_1,\ldots,x_n) = \prod_{i=1}^{n} P(x_i|\mathrm{pa}(X_i))$$
  $\forall P(X_i|\mathrm{pa}(X_i)) \in K(X_i|\mathrm{pa}(X_i))$
  $\forall i = 1,\ldots,n \quad \forall \mathrm{pa}(X_i)$

- Every conditional mass function takes values in its credal set independently of the others
  CN $\equiv$ (exponential) number of BNs



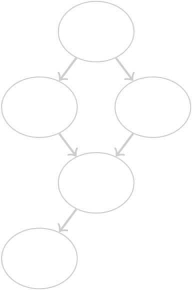
E.g., $K(X_1)$ defined by constraint $P(x_1) > .75$, very likely to be warm

Non-separately specified CNs

- Constraints among different conditional mass functions of a CN
- Explicit enumeration of the relative BNs
  - Auxiliary parent selecting the conditional probabilities *(Cano, Cano, Moral, 1994)* with a vacuous prior
- "Extensive" specification
  - Constraints among conditional mass functions of the same variable
  - Each CPT takes values from a set of tables an auxiliary parent selecting the tables
- An unconstrained (i.e., separated) specification is always possible *(Antonucci & Zaffalon, IJAR, 2008)*

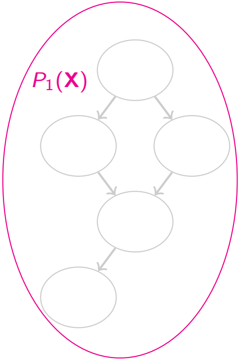
Non-separately specified CNs

- Constraints among different conditional mass functions of a CN
- Explicit enumeration of the relative BNs
  - Auxiliary parent selecting the conditional probabilities *(Cano, Cano, Moral, 1994)* with a vacuous prior
- "Extensive" specification
  - Constraints among conditional mass functions of the same variable
  - Each CPT takes values from a set of tables an auxiliary parent selecting the tables
- An unconstrained (i.e., separated) specification is always possible *(Antonucci & Zaffalon, IJAR, 2008)*



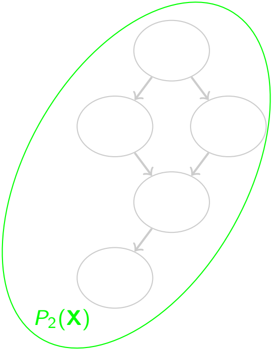
Non-separately specified CNs

- Constraints among different conditional mass functions of a CN
- Explicit enumeration of the relative BNs
    - Auxiliary parent selecting the conditional probabilities *(Cano, Cano, Moral, 1994)* with a vacuous prior
- "Extensive" specification
    - Constraints among conditional mass functions of the same variable
    - Each CPT takes values from a set of tables an auxiliary parent selecting the tables
- An unconstrained (i.e., separated) specification is always possible *(Antonucci & Zaffalon, IJAR, 2008)*



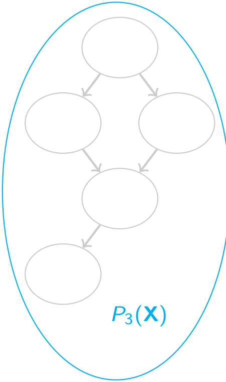
Non-separately specified CNs

- Constraints among different conditional mass functions of a CN
- Explicit enumeration of the relative BNs
    - Auxiliary parent selecting the conditional probabilities *(Cano, Cano, Moral, 1994)* with a vacuous prior
- "Extensive" specification
    - Constraints among conditional mass functions of the same variable
    - Each CPT takes values from a set of tables an auxiliary parent selecting the tables
- An unconstrained (i.e., separated) specification is always possible *(Antonucci & Zaffalon, IJAR, 2008)*



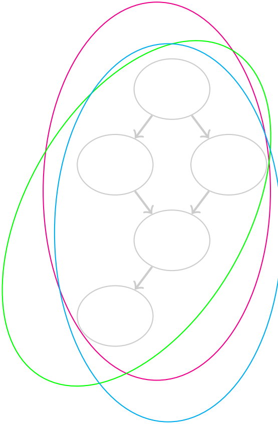
Non-separately specified CNs

- Constraints among different conditional mass functions of a CN
- Explicit enumeration of the relative BNs
  - Auxiliary parent selecting the conditional probabilities *(Cano, Cano, Moral, 1994)* with a vacuous prior
- "Extensive" specification
  - Constraints among conditional mass functions of the same variable
  - Each CPT takes values from a set of tables an auxiliary parent selecting the tables
- An unconstrained (i.e., separated) specification is always possible *(Antonucci & Zaffalon, IJAR, 2008)*



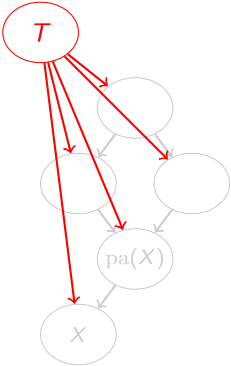
Non-separately specified CNs

- Constraints among different conditional mass functions of a CN
- Explicit enumeration of the relative BNs
  - Auxiliary parent selecting the conditional probabilities *(Cano, Cano, Moral, 1994)* with a vacuous prior
- "Extensive" specification
  - Constraints among conditional mass functions of the same variable
  - Each CPT takes values from a set of tables an auxiliary parent selecting the tables
- An unconstrained (i.e., separated) specification is always possible *(Antonucci & Zaffalon, IJAR, 2008)*



Non-separately specified CNs

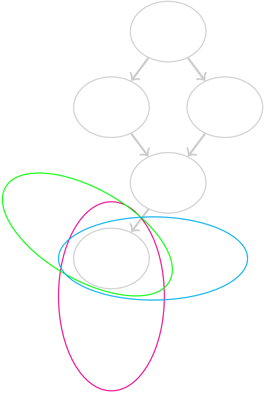
- Constraints among different conditional mass functions of a CN
- Explicit enumeration of the relative BNs
  - Auxiliary parent selecting the conditional probabilities *(Cano, Cano, Moral, 1994)* with a vacuous prior
- "Extensive" specification
  - Constraints among conditional mass functions of the same variable
  - Each CPT takes values from a set of tables an auxiliary parent selecting the tables
- An unconstrained (i.e., separated) specification is always possible *(Antonucci & Zaffalon, IJAR, 2008)*



$$P(X|\mathrm{pa}(X),\ T = t_j\ ) = P_j(X|\mathrm{pa}(X))$$

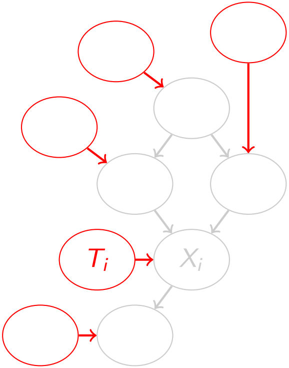
Non-separately specified CNs

- Constraints among different conditional mass functions of a CN
- Explicit enumeration of the relative BNs
  - Auxiliary parent selecting the conditional probabilities *(Cano, Cano, Moral, 1994)* with a vacuous prior
- "Extensive" specification
  - Constraints among conditional mass functions of the same variable
  - Each CPT takes values from a set of tables an auxiliary parent selecting the tables
- An unconstrained (i.e., separated) specification is always possible *(Antonucci & Zaffalon, IJAR, 2008)*



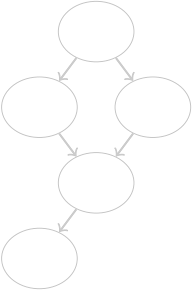
Non-separately specified CNs

- Constraints among different conditional mass functions of a CN
- Explicit enumeration of the relative BNs
  - Auxiliary parent selecting the conditional probabilities *(Cano, Cano, Moral, 1994)* with a vacuous prior
- "Extensive" specification
  - Constraints among conditional mass functions of the same variable
  - Each CPT takes values from a set of tables an auxiliary parent selecting the tables
- An unconstrained (i.e., separated) specification is always possible *(Antonucci & Zaffalon, IJAR, 2008)*

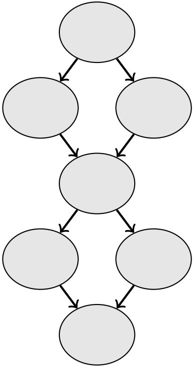


Non-separately specified CNs

- Constraints among different conditional mass functions of a CN
- Explicit enumeration of the relative BNs
  - Auxiliary parent selecting the conditional probabilities *(Cano, Cano, Moral, 1994)* with a vacuous prior
- "Extensive" specification
  - Constraints among conditional mass functions of the same variable
  - Each CPT takes values from a set of tables an auxiliary parent selecting the tables
- An unconstrained (i.e., separated) specification is always possible *(Antonucci & Zaffalon, IJAR, 2008)*



Updating credal networks



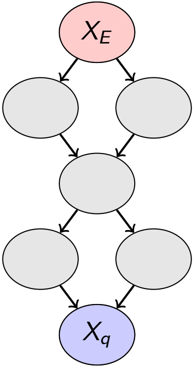
- Conditional probs for a variable of interest X_q given observations $X_E = x_E$
- Updating Bayesian nets is NP-hard
  (fast algorithms for polytrees)

$$P(x_q|x_E) = \frac{P(x_q, x_E)}{P(x_E)} = \frac{\sum_{\mathbf{x}\setminus\{x_q,x_E\}} \prod_{i=1}^n P(x_i|\pi_i)}{\sum_{\mathbf{x}\setminus\{x_E\}} \prod_{i=1}^n P(x_i|\pi_i)}$$

- Updating credal nets is $\mathrm{NP}^{\mathrm{PP}}$-hard,
  NP-hard on polytrees *(Mauá et al., 2013)*

$$\underline{P}(x_q|x_E) = \min_{\substack{P(X_i|\pi_i)\in K(X_i|\pi_i)\\ i=1,\dots,n}} \frac{\sum_{\mathbf{x}\setminus\{x_Q,x_E\}} \prod_{i=1}^n P(x_i|\pi_i)}{\sum_{\mathbf{x}\setminus\{x_Q\}} \prod_{i=1}^n P(x_i|\pi_i)}$$

Updating credal networks



- Conditional probs for a variable of interest X_q given observations $X_E = x_E$
- Updating Bayesian nets is NP-hard
  (fast algorithms for polytrees)

$$P(x_q|x_E) = \frac{P(x_q, x_E)}{P(x_E)} = \frac{\sum_{\mathbf{x}\setminus\{x_q,x_E\}} \prod_{i=1}^{n} P(x_i|\pi_i)}{\sum_{\mathbf{x}\setminus\{x_E\}} \prod_{i=1}^{n} P(x_i|\pi_i)}$$

- Updating credal nets is $\mathrm{NP}^{\mathrm{PP}}$-hard,
  NP-hard on polytrees *(Mauá et al., 2013)*

$$\underline{P}(x_q|x_E) = \min_{\substack{P(X_i|\pi_i)\in K(X_i|\pi_i)\\ i=1,\ldots,n}} \frac{\sum_{\mathbf{x}\setminus\{x_Q,x_E\}} \prod_{i=1}^{n} P(x_i|\pi_i)}{\sum_{\mathbf{x}\setminus\{x_Q\}} \prod_{i=1}^{n} P(x_i|\pi_i)}$$

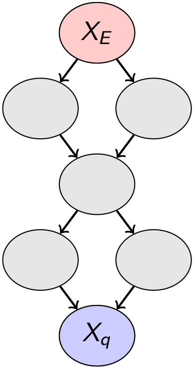
Updating credal networks

- Conditional probs for a variable of interest X_q given observations $X_E = x_E$
- Updating Bayesian nets is NP-hard
  (fast algorithms for polytrees)

$$P(x_q|x_E) = \frac{P(x_q, x_E)}{P(x_E)} = \frac{\sum_{\mathbf{x}\setminus\{x_q,x_E\}} \prod_{i=1}^{n} P(x_i|\pi_i)}{\sum_{\mathbf{x}\setminus\{x_E\}} \prod_{i=1}^{n} P(x_i|\pi_i)}$$

- Updating credal nets is NP^{PP}-hard,
  NP-hard on polytrees *(Mauá et al., 2013)*

$$\underline{P}(x_q|x_E) = \min_{\substack{P(X_i|\pi_i)\in K(X_i|\pi_i)\\ i=1,\ldots,n}} \frac{\sum_{\mathbf{x}\setminus\{x_Q,x_E\}} \prod_{i=1}^{n} P(x_i|\pi_i)}{\sum_{\mathbf{x}\setminus\{x_Q\}} \prod_{i=1}^{n} P(x_i|\pi_i)}$$



$P(x_q|x_E) = .38$

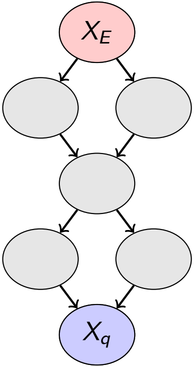
Updating credal networks

- Conditional probs for a variable of interest X_q given observations $X_E = x_E$
- Updating Bayesian nets is NP-hard
  (fast algorithms for polytrees)

$$P(x_q|x_E) = \frac{P(x_q, x_E)}{P(x_E)} = \frac{\sum_{\mathbf{x}\setminus\{x_q,x_E\}} \prod_{i=1}^{n} P(x_i|\pi_i)}{\sum_{\mathbf{x}\setminus\{x_E\}} \prod_{i=1}^{n} P(x_i|\pi_i)}$$

- Updating credal nets is $\mathrm{NP}^{\mathrm{PP}}$-hard,
  NP-hard on polytrees *(Mauá et al., 2013)*

$$\underline{P}(x_q|x_E) = \min_{\substack{P(X_i|\pi_i)\in K(X_i|\pi_i)\\ i=1,\ldots,n}} \frac{\sum_{\mathbf{x}\setminus\{x_Q,x_E\}} \prod_{i=1}^{n} P(x_i|\pi_i)}{\sum_{\mathbf{x}\setminus\{x_Q\}} \prod_{i=1}^{n} P(x_i|\pi_i)}$$



$.21 \leq P(X_q|x_E) \leq .46$

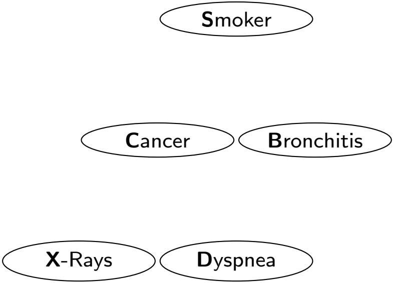
Medical diagnosis by CNs (a simple example of)

- Five Boolean vars
- Conditional independence relations by a DAG
- Elicitation of the local (conditional) CSs
- This is a CN specification
- The strong extension $K(S,C,B,X,D) =$

$$\mathrm{CH}\left\{ P(S,C,B,X,D) \left| \begin{array}{l} P(s,c,b,x,d){=}P(s)P(c|s)P(b|s)P(x|c)P(d|c,b) \\ P(S) \in K(S) \\ P(C|s) \in K(C|s), P(C|\neg s) \in K(C|\neg s) \\ \ldots \end{array} \right. \right\}$$

Medical diagnosis by CNs (a simple example of)

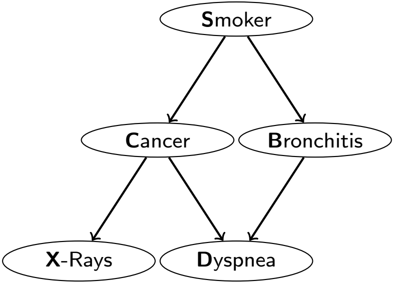
- Five Boolean vars
- Conditional independence relations by a DAG
- Elicitation of the local (conditional) CSs
- This is a CN specification
- The strong extension $K(S, C, B, X, D) =$



$$\mathrm{CH}\left\{ P(S,C,B,X,D) \;\middle|\; \begin{array}{l} P(s,c,b,x,d){=}P(s)P(c|s)P(b|s)P(x|c)P(d|c,b) \\ P(S) \in K(S) \\ P(C|s) \in K(C|s), P(C|\neg s) \in K(C|\neg s) \\ \ldots \end{array} \right\}$$

Medical diagnosis by CNs (a simple example of)

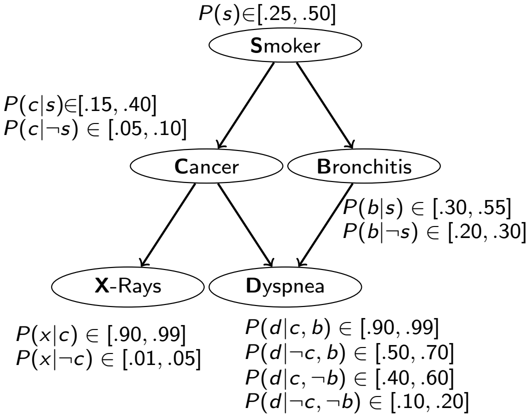
- Five Boolean vars
- Conditional independence relations by a DAG
- Elicitation of the local (conditional) CSs
- This is a CN specification
- The strong extension $K(S, C, B, X, D) =$



$$\mathrm{CH}\left\{P(S,C,B,X,D)\;\middle|\;\begin{array}{l} P(s,c,b,x,d){=}P(s)P(c|s)P(b|s)P(x|c)P(d|c,b)\\ P(S)\in K(S)\\ P(C|s)\in K(C|s), P(C|\neg s)\in K(C|\neg s)\\ \ldots \end{array}\right\}$$

Medical diagnosis by CNs (a simple example of)

- Five Boolean vars
- Conditional independence relations by a DAG
- Elicitation of the local (conditional) CSs
- This is a CN specification
- The strong extension $K(S, C, B, X, D) =$



$P(s) \in [.25, .50]$

$P(c|s) \in [.15, .40]$
$P(c|\neg s) \in [.05, .10]$

$P(b|s) \in [.30, .55]$
$P(b|\neg s) \in [.20, .30]$

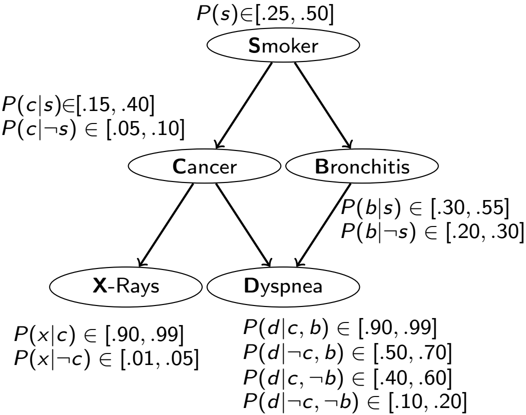
$P(x|c) \in [.90, .99]$
$P(x|\neg c) \in [.01, .05]$

$P(d|c, b) \in [.90, .99]$
$P(d|\neg c, b) \in [.50, .70]$
$P(d|c, \neg b) \in [.40, .60]$
$P(d|\neg c, \neg b) \in [.10, .20]$

$$\mathrm{CH}\left\{ P(S, C, B, X, D) \;\middle|\; \begin{array}{l} P(s, c, b, x, d) = P(s)P(c|s)P(b|s)P(x|c)P(d|c, b) \\ P(S) \in K(S) \\ P(C|s) \in K(C|s), P(C|\neg s) \in K(C|\neg s) \\ \ldots \end{array} \right\}$$

Medical diagnosis by CNs (a simple example of)

- Five Boolean vars
- Conditional independence relations by a DAG
- Elicitation of the local (conditional) CSs
- This is a CN specification
- The strong extension $K(S, C, B, X, D) =$



$P(s) \in [.25, .50]$

Smoker

$P(c|s) \in [.15, .40]$
$P(c|\neg s) \in [.05, .10]$

Cancer

Bronchitis

$P(b|s) \in [.30, .55]$
$P(b|\neg s) \in [.20, .30]$

X-Rays

Dyspnea

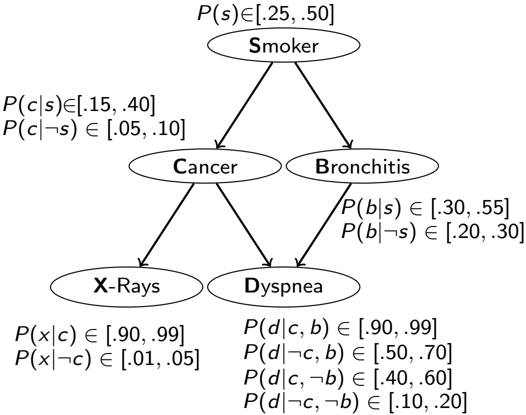
$P(x|c) \in [.90, .99]$
$P(x|\neg c) \in [.01, .05]$

$P(d|c, b) \in [.90, .99]$
$P(d|\neg c, b) \in [.50, .70]$
$P(d|c, \neg b) \in [.40, .60]$
$P(d|\neg c, \neg b) \in [.10, .20]$

$$\mathrm{CH}\left\{ P(S, C, B, X, D) \;\middle|\; \begin{array}{l} P(s, c, b, x, d) = P(s)P(c|s)P(b|s)P(x|c)P(d|c, b) \\ P(S) \in K(S) \\ P(C|s) \in K(C|s), P(C|\neg s) \in K(C|\neg s) \\ \ldots \end{array} \right\}$$

Medical diagnosis by CNs (a simple example of)

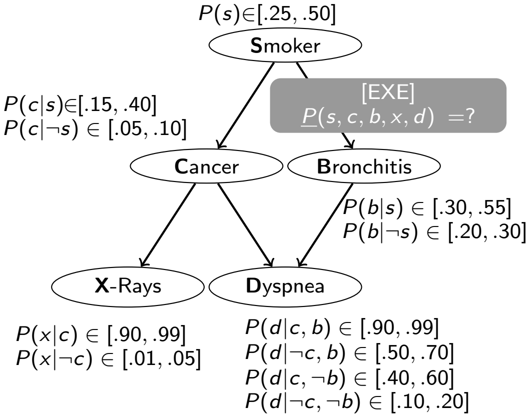
- Five Boolean vars
- Conditional independence relations by a DAG
- Elicitation of the local (conditional) CSs
- This is a CN specification
- The strong extension $K(S, C, B, X, D) =$



$$\mathrm{CH}\left\{P(S, C, B, X, D)\;\middle|\;\begin{array}{l} P(s, c, b, x, d)=P(s)P(c|s)P(b|s)P(x|c)P(d|c, b)\\ P(S) \in K(S)\\ P(C|s) \in K(C|s), P(C|\neg s) \in K(C|\neg s)\\ \ldots \end{array}\right\}$$

Medical diagnosis by CNs (a simple example of)

- Five Boolean vars
- Conditional independence relations by a DAG
- Elicitation of the local (conditional) CSs
- This is a CN specification
- The strong extension $K(S, C, B, X, D) =$



$P(s) \in [.25, .50]$

[EXE]
$\underline{P}(s, c, b, x, d) \; =?$

$P(c|s) \in [.15, .40]$
$P(c|\neg s) \in [.05, .10]$

$P(b|s) \in [.30, .55]$
$P(b|\neg s) \in [.20, .30]$

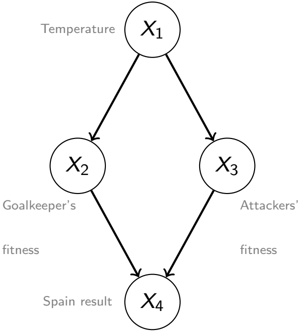
$P(x|c) \in [.90, .99]$
$P(x|\neg c) \in [.01, .05]$

$P(d|c, b) \in [.90, .99]$
$P(d|\neg c, b) \in [.50, .70]$
$P(d|c, \neg b) \in [.40, .60]$
$P(d|\neg c, \neg b) \in [.10, .20]$

$$\mathrm{CH}\left\{ P(S, C, B, X, D) \;\middle|\; \begin{array}{l} P(s, c, b, x, d) = P(s)P(c|s)P(b|s)P(x|c)P(d|c, b) \\ P(S) \in K(S) \\ P(C|s) \in K(C|s), P(C|\neg s) \in K(C|\neg s) \\ \ldots \end{array} \right\}$$

Updating with incomplete observations

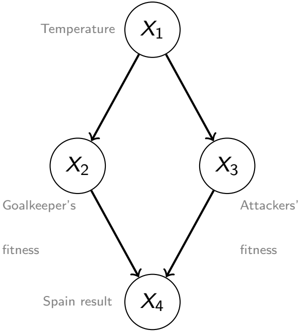
- $\underline{P}(X_q = x_q | X_E = x_E, X_M = *)$
  $= \underline{P}(X_q = x_q | X_E = x_E)$
  right only if *missing at random* assumption holds
- *Conservative inference rule* (CIR)
  $\underline{P}(x_q | x_E, *) = \min_{x_M \in \Omega_{X_M}} P(x_q | x_E, x_M)$
  near-ignorance about the process preventing some variable from being observed *(de Cooman & Zaffalon, 2004)*
- CIR on CNs?
  - Add a (dummy) binary child for each missing, with vacuous quantification
  - Use standard updating algorithms



$\underline{P}(x_4 | X_1 = warm, X_2 = *)$

Updating with incomplete observations

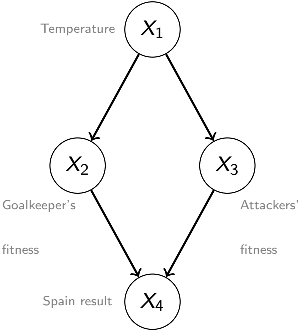
- $\underline{P}(X_q = x_q | X_E = x_E, X_M = *)$ $= \underline{P}(X_q = x_q | X_E = x_E)$ right only if *missing at random* assumption holds
- *Conservative inference rule* (CIR) $\underline{P}(x_q|x_E,*) = \min_{x_M \in \Omega_{X_M}} \underline{P}(x_q|x_E,x_M)$ near-ignorance about the process preventing some variable from being observed *(de Cooman & Zaffalon, 2004)*
- CIR on CNs?
  - Add a (dummy) binary child for each missing, with vacuous quantification
  - Use standard updating algorithms



$\underline{P}(x_4|X_1 = warm, X_2 = *)$

Updating with incomplete observations

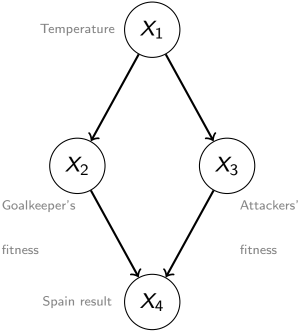
- $\underline{P}(X_q = x_q | X_E = x_E, X_M = *)$ $= \underline{P}(X_q = x_q | X_E = x_E)$ right only if *missing at random* assumption holds
- *Conservative inference rule* (CIR) $\underline{P}(x_q | x_E, *) = \min_{x_M \in \Omega_{X_M}} P(x_q | x_E, x_M)$ near-ignorance about the process preventing some variable from being observed *(de Cooman & Zaffalon, 2004)*
- CIR on CNs?
  - Add a (dummy) binary child for each missing, with vacuous quantification
  - Use standard updating algorithms



$$\underline{P}(x_4 | X_1 = warm, X_2 = *)$$
$$= \min\{\underline{P}(x_4 | warm, good), \underline{P}(x_4 | warm, bad)\}$$

Updating with incomplete observations

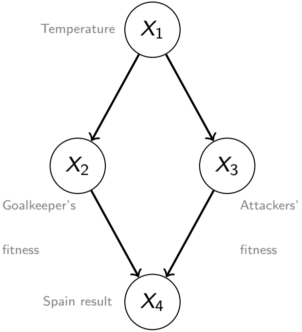
- $\underline{P}(X_q = x_q | X_E = x_E, X_M = *)$ $= \underline{P}(X_q = x_q | X_E = x_E)$ right only if *missing at random* assumption holds
- *Conservative inference rule* (CIR) $\underline{P}(x_q | x_E, *) = \min_{x_M \in \Omega_{X_M}} P(x_q | x_E, x_M)$ near-ignorance about the process preventing some variable from being observed *(de Cooman & Zaffalon, 2004)*
- CIR on CNs?
  - Add a (dummy) binary child for each missing, with vacuous quantification
  - Use standard updating algorithms



$$\underline{P}(x_4 | X_1 = warm, X_2 = *)$$
$$= \min\{\underline{P}(x_4 | warm, good), \underline{P}(x_4 | warm, bad)\}$$

Updating with incomplete observations

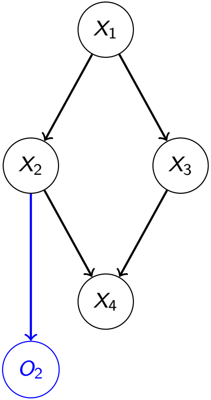
- $\underline{P}(X_q = x_q | X_E = x_E, X_M = *)$
  $= \underline{P}(X_q = x_q | X_E = x_E)$
  right only if *missing at random* assumption holds
- *Conservative inference rule* (CIR)
  $\underline{P}(x_q | x_E, *) = \min_{x_M \in \Omega_{X_M}} P(x_q | x_E, x_M)$
  near-ignorance about the process preventing some variable from being observed *(de Cooman & Zaffalon, 2004)*
- CIR on CNs?
  - Add a (dummy) binary child for each missing, with vacuous quantification
  - Use standard updating algorithms



$$\underline{P}(x_4 | X_1 = warm, X_2 = *)$$
$$= \min\{\underline{P}(x_4 | warm, good), \underline{P}(x_4 | warm, bad)\}$$

Updating with incomplete observations

- $\underline{P}(X_q = x_q | X_E = x_E, X_M = *)$
  $= \underline{P}(X_q = x_q | X_E = x_E)$
  right only if *missing at random* assumption holds
- *Conservative inference rule* (CIR)
  $\underline{P}(x_q|x_E, *) = \min_{x_M \in \Omega_{X_M}} \underline{P}(x_q|x_E, x_M)$
  near-ignorance about the process preventing some variable from being observed *(de Cooman & Zaffalon, 2004)*
- CIR on CNs?
  - Add a (dummy) binary child for each missing, with vacuous quantification
  - Use standard updating algorithms



$\underline{K}(O_2|x_2)$ vacuous CS

$$\underline{P}(x_4|X_1 = warm, X_2 = *)$$
$$= \min\{\underline{P}(x_4|warm, good), \underline{P}(x_4|warm, bad)\}$$

Updating with incomplete observations

- $\underline{P}(X_q = x_q | X_E = x_E, X_M = *)$

  $= \underline{P}(X_q = x_q | X_E = x_E)$

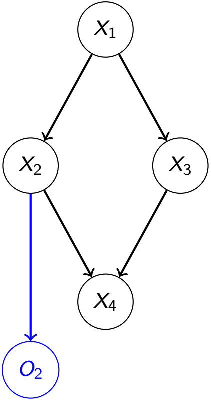
  right only if *missing at random* assumption holds

- *Conservative inference rule* (CIR)

  $\underline{P}(x_q | x_E, *) = \min_{x_M \in \Omega_{X_M}} P(x_q | x_E, x_M)$

  near-ignorance about the process preventing some variable from being observed *(de Cooman & Zaffalon, 2004)*

- CIR on CNs?
  - Add a (dummy) binary child for each missing, with vacuous quantification
  - Use standard updating algorithms

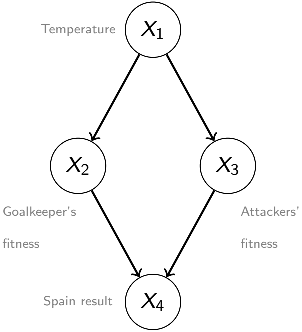


$\underline{K}(O_2|x_2)$ vacuous CS

$$\underline{P}(x_4|X_1 = warm, X_2 = *)$$
$$= \min\{\underline{P}(x_4|warm, good), \underline{P}(x_4|warm, bad)\}$$
$$= \underline{P}(x_4|warm, O_2 = 1)$$

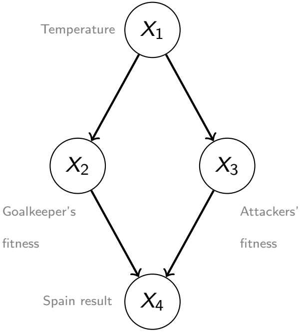
Modelling the observational process

- Each X as a *latent* variable
- For each X a *manifest* variable O_X modelling the observation $\Omega_O = \Omega_X \cup \{*\}$
- Conditional independence, given X between O and the other variables (or weaker conditions)
- Quantifying link between O and X (observational process)
- A CS $K(O|x)$ might a realistic model! (better than $P(O|X)$)
- Standard updating problem $\underline{P}(X_q|O_E = x_E)$



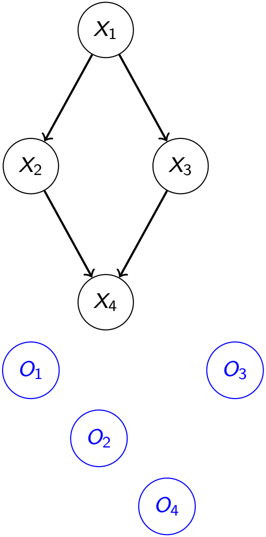
Modelling the observational process

- Each X as a *latent* variable
- For each X a *manifest* variable O_X modelling the observation $\Omega_O = \Omega_X \cup \{*\}$
- Conditional independence, given X between O and the other variables (or weaker conditions)
- Quantifying link between O and X (observational process)
- A CS $K(O|x)$ might a realistic model! (better than $P(O|X)$)
- Standard updating problem $\underline{P}(X_q|O_E = x_E)$



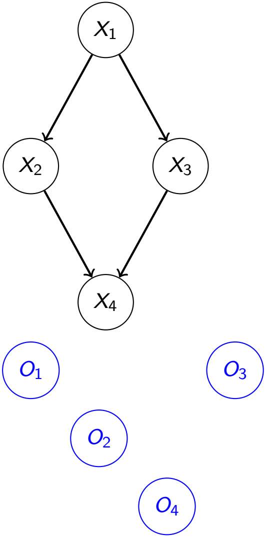
Modelling the observational process

- Each X as a *latent* variable
- For each X a *manifest* variable O_X modelling the observation $\Omega_O = \Omega_X \cup \{*\}$
- Conditional independence, given X between O and the other variables (or weaker conditions)
- Quantifying link between O and X (observational process)
- A CS $K(O|x)$ might a realistic model! (better than $P(O|X)$)
- Standard updating problem $\underline{P}(X_q|O_E = x_E)$



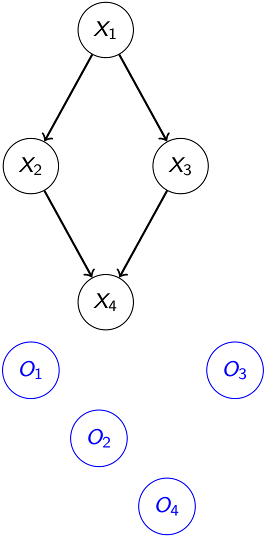
Modelling the observational process

- Each X as a *latent* variable
- For each X a *manifest* variable O_X modelling the observation $\Omega_O = \Omega_X \cup \{*\}$
- Conditional independence, given X between O and the other variables (or weaker conditions)
- Quantifying link between O and X (observational process)
- A CS $K(O|x)$ might a realistic model! (better than $P(O|X)$)
- Standard updating problem $\underline{P}(X_q|O_E = x_E)$



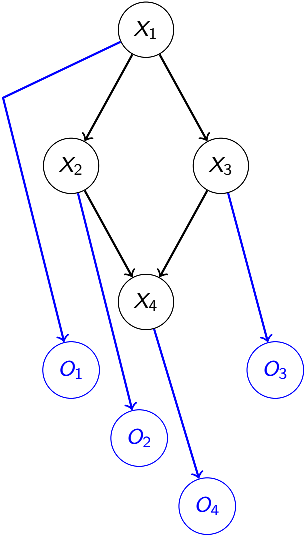
Modelling the observational process

- Each X as a *latent* variable
- For each X a *manifest* variable O_X modelling the observation $\Omega_O = \Omega_X \cup \{*\}$
- Conditional independence, given X between O and the other variables (or weaker conditions)
- Quantifying link between O and X (observational process)
- A CS $K(O|x)$ might a realistic model! (better than $P(O|X)$)
- Standard updating problem $\underline{P}(X_q|O_E = x_E)$



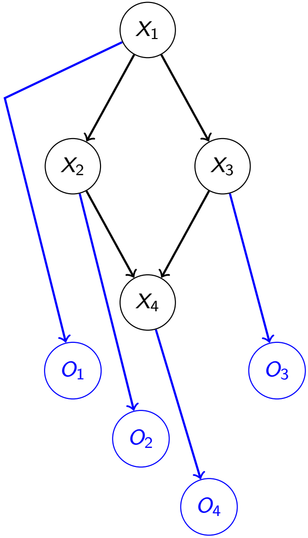
Modelling the observational process

- Each X as a *latent* variable
- For each X a *manifest* variable O_X modelling the observation $\Omega_O = \Omega_X \cup \{*\}$
- Conditional independence, given X between O and the other variables (or weaker conditions)
- Quantifying link between O and X (observational process)
- A CS $K(O|x)$ might a realistic model! (better than $P(O|X)$)
- Standard updating problem $\underline{P}(X_q|O_E = x_E)$



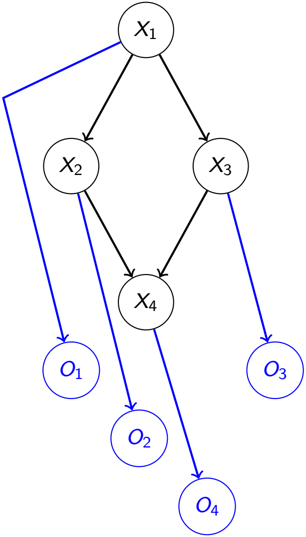
Modelling the observational process

- Each X as a *latent* variable
- For each X a *manifest* variable O_X modelling the observation $\Omega_O = \Omega_X \cup \{*\}$
- Conditional independence, given X between O and the other variables (or weaker conditions)
- Quantifying link between O and X (observational process)
- A CS $K(O|x)$ might a realistic model! (better than $P(O|X)$)
- Standard updating problem $\underline{P}(X_q|O_E = x_E)$



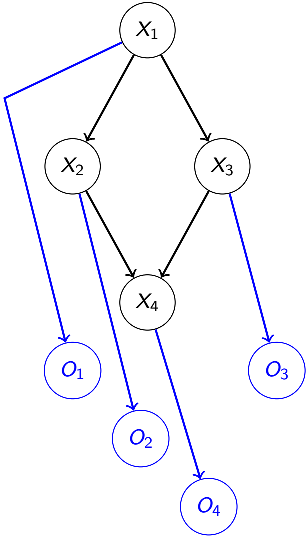
Modelling the observational process

- Each X as a *latent* variable
- For each X a *manifest* variable O_X modelling the observation $\Omega_O = \Omega_X \cup \{*\}$
- Conditional independence, given X between O and the other variables (or weaker conditions)
- Quantifying link between O and X (observational process)
- A CS $K(O|x)$ might a realistic model! (better than $P(O|X)$)
- Standard updating problem $\underline{P}(X_q|O_E = x_E)$



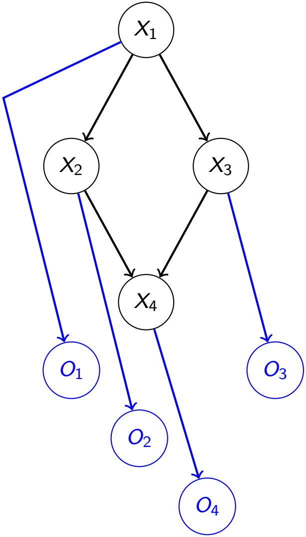
Modelling the observational process

- Each X as a *latent* variable
- For each X a *manifest* variable O_X modelling the observation $\Omega_O = \Omega_X \cup \{*\}$
- Conditional independence, given X between O and the other variables (or weaker conditions)
- Quantifying link between O and X (observational process)
- A CS $K(O|x)$ might a realistic model! (better than $P(O|X)$)
- Standard updating problem $\underline{P}(X_q|O_E = x_E)$



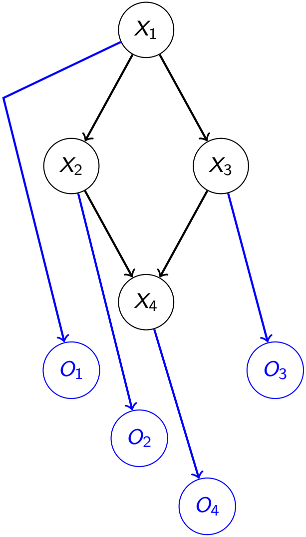
Modelling the observational process (ii)

- Manifest variables reduced to binary variables (coarsen to $\{o, \neg o\}$)
- Elicit only lower/upper likelihoods of observation given the latent $\underline{P}(o|x) \leq P(o|x) \leq \overline{P}(o|x)$
  - Perfect observation: $\underline{P}(o|x) = \overline{P}(o|x) = \delta_{o,x}$
  - MAR: $\underline{P}(o|x) = \overline{P}(o|x) = k$
  - CIR: $\underline{P}(o|x) = 0, \overline{P}(o|x) = 1$
  - Imprecise likelihood ratio (and Jeffrey's rule)
- Hard evidence? Drop leaving arcs!
- Only the subnet connected to the query node



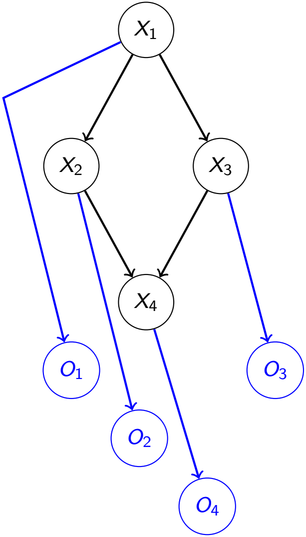
Modelling the observational process (ii)

- Manifest variables reduced to binary variables (coarsen to $\{o, \neg o\}$)
- Elicit only lower/upper likelihoods of observation given the latent $\underline{P}(o|x) \leq P(o|x) \leq \overline{P}(o|x)$
  - Perfect observation: $\underline{P}(o|x) = \overline{P}(o|x) = \delta_{o,x}$
  - MAR: $\underline{P}(o|x) = \overline{P}(o|x) = k$
  - CIR: $\underline{P}(o|x) = 0, \overline{P}(o|x) = 1$
  - Imprecise likelihood ratio (and Jeffrey's rule)
- Hard evidence? Drop leaving arcs!
- Only the subnet connected to the query node



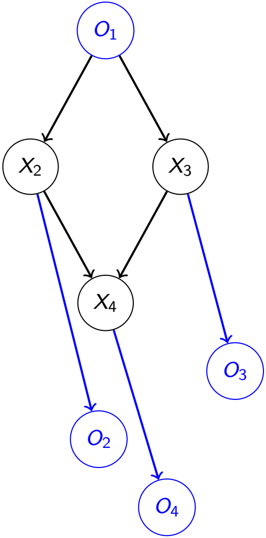
Modelling the observational process (ii)

- Manifest variables reduced to binary variables (coarsen to $\{o, \neg o\}$)
- Elicit only lower/upper likelihoods of observation given the latent $\underline{P}(o|x) \leq P(o|x) \leq \overline{P}(o|x)$
  - Perfect observation: $\underline{P}(o|x) = \overline{P}(o|x) = \delta_{o,x}$
  - MAR: $\underline{P}(o|x) = \overline{P}(o|x) = k$
  - CIR: $\underline{P}(o|x) = 0, \overline{P}(o|x) = 1$
  - Imprecise likelihood ratio (and Jeffrey's rule)
- Hard evidence? Drop leaving arcs!
- Only the subnet connected to the query node



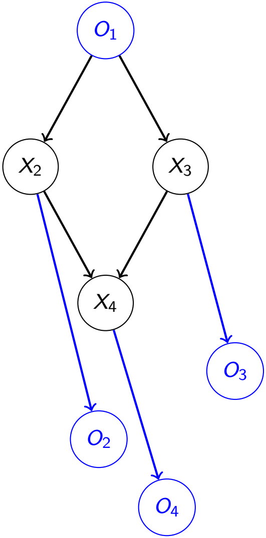
Modelling the observational process (ii)

- Manifest variables reduced to binary variables (coarsen to $\{o, \neg o\}$)
- Elicit only lower/upper likelihoods of observation given the latent $\underline{P}(o|x) \leq P(o|x) \leq \overline{P}(o|x)$
  - Perfect observation: $\underline{P}(o|x) = \overline{P}(o|x) = \delta_{o,x}$
  - MAR: $\underline{P}(o|x) = \overline{P}(o|x) = k$
  - CIR: $\underline{P}(o|x) = 0, \overline{P}(o|x) = 1$
  - Imprecise likelihood ratio (and Jeffrey's rule)
- Hard evidence? Drop leaving arcs!
- Only the subnet connected to the query node



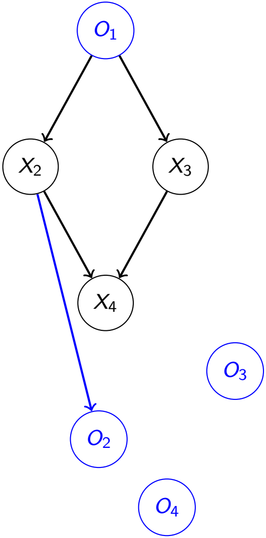
Modelling the observational process (ii)

- Manifest variables reduced to binary variables (coarsen to $\{o, \neg o\}$)
- Elicit only lower/upper likelihoods of observation given the latent $\underline{P}(o|x) \leq P(o|x) \leq \overline{P}(o|x)$
  - Perfect observation: $\underline{P}(o|x) = \overline{P}(o|x) = \delta_{o,x}$
  - MAR: $\underline{P}(o|x) = \overline{P}(o|x) = k$
  - CIR: $\underline{P}(o|x) = 0, \overline{P}(o|x) = 1$
  - Imprecise likelihood ratio (and Jeffrey's rule)
- Hard evidence? Drop leaving arcs!
- Only the subnet connected to the query node



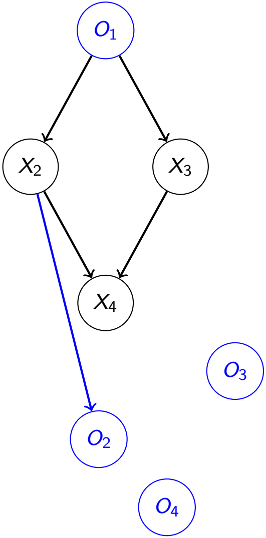
Modelling the observational process (ii)

- Manifest variables reduced to binary variables (coarsen to $\{o, \neg o\}$)
- Elicit only lower/upper likelihoods of observation given the latent $\underline{P}(o|x) \leq P(o|x) \leq \overline{P}(o|x)$
  - Perfect observation: $\underline{P}(o|x) = \overline{P}(o|x) = \delta_{o,x}$
  - MAR: $\underline{P}(o|x) = \overline{P}(o|x) = k$
  - CIR: $\underline{P}(o|x) = 0, \overline{P}(o|x) = 1$
  - Imprecise likelihood ratio (and Jeffrey's rule)
- Hard evidence? Drop leaving arcs!
- Only the subnet connected to the query node



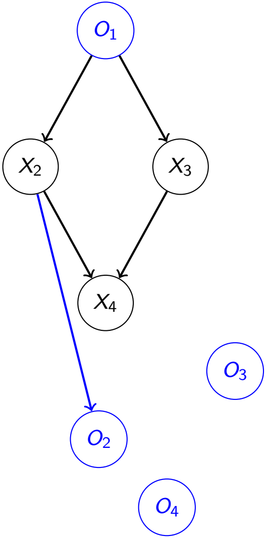
Modelling the observational process (ii)

- Manifest variables reduced to binary variables (coarsen to $\{o, \neg o\}$)
- Elicit only lower/upper likelihoods of observation given the latent $\underline{P}(o|x) \leq P(o|x) \leq \overline{P}(o|x)$
  - Perfect observation: $\underline{P}(o|x) = \overline{P}(o|x) = \delta_{o,x}$
  - MAR: $\underline{P}(o|x) = \overline{P}(o|x) = k$
  - CIR: $\underline{P}(o|x) = 0, \overline{P}(o|x) = 1$
  - Imprecise likelihood ratio (and Jeffrey's rule)
- Hard evidence? Drop leaving arcs!
- Only the subnet connected to the query node



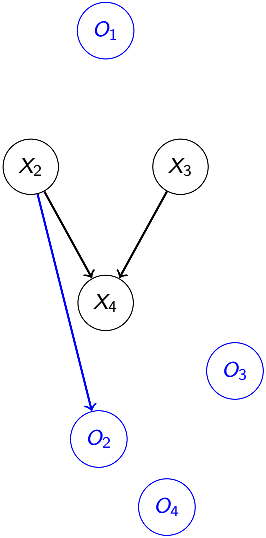
Modelling the observational process (ii)

- Manifest variables reduced to binary variables (coarsen to $\{o, \neg o\}$)
- Elicit only lower/upper likelihoods of observation given the latent $\underline{P}(o|x) \leq P(o|x) \leq \overline{P}(o|x)$
    - Perfect observation: $\underline{P}(o|x) = \overline{P}(o|x) = \delta_{o,x}$
    - MAR: $\underline{P}(o|x) = \overline{P}(o|x) = k$
    - CIR: $\underline{P}(o|x) = 0, \overline{P}(o|x) = 1$
    - Imprecise likelihood ratio (and Jeffrey's rule)
- Hard evidence? Drop leaving arcs!
- Only the subnet connected to the query node



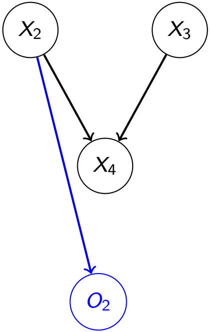
Modelling the observational process (ii)

- Manifest variables reduced to binary variables (coarsen to $\{o, \neg o\}$)
- Elicit only lower/upper likelihoods of observation given the latent $\underline{P}(o|x) \leq P(o|x) \leq \overline{P}(o|x)$
  - Perfect observation: $\underline{P}(o|x) = \overline{P}(o|x) = \delta_{o,x}$
  - MAR: $\underline{P}(o|x) = \overline{P}(o|x) = k$
  - CIR: $\underline{P}(o|x) = 0, \overline{P}(o|x) = 1$
  - Imprecise likelihood ratio (and Jeffrey's rule)
- Hard evidence? Drop leaving arcs!
- Only the subnet connected to the query node



Modelling the observational process (ii)

- Manifest variables reduced to binary variables (coarsen to $\{o, \neg o\}$)
- Elicit only lower/upper likelihoods of observation given the latent $\underline{P}(o|x) \leq P(o|x) \leq \overline{P}(o|x)$
  - Perfect observation: $\underline{P}(o|x) = \overline{P}(o|x) = \delta_{o,x}$
  - MAR: $\underline{P}(o|x) = \overline{P}(o|x) = k$
  - CIR: $\underline{P}(o|x) = 0, \overline{P}(o|x) = 1$
  - Imprecise likelihood ratio (and Jeffrey's rule)
- Hard evidence? Drop leaving arcs!
- Only the subnet connected to the query node



Averaging expert opinions

- Experts A and B give (precise knowledge) $P_A(X)$ and $P_B(X)$
- Arithmetic average $P_+(x) \propto P_A(x) \cdot P_B(x)$
- Geometric average $P_\times(x) \propto P_A(x) \cdot P_B(x)$
- Averages can be computed within the BN formalism
- This makes easy to extend the averages to credal sets
  by using credal networks [i.e., experts give $K_A(X)$ and $K_B(X)$]

Averaging expert opinions

- Experts A and B give (precise knowledge) $P_A(X)$ and $P_B(X)$
- Arithmetic average $P_+(x) \propto P_A(x) \cdot P_B(x)$
- Geometric average $P_\times(x) \propto P_A(x) \cdot P_B(x)$
- Averages can be computed within the BN formalism
- This makes easy to extend the averages to credal sets

  by using credal networks [i.e., experts give $K_A(X)$ and $K_B(X)$]

Averaging expert opinions

- Experts A and B give (precise knowledge) $P_A(X)$ and $P_B(X)$
- Arithmetic average $P_+(x) \propto P_A(x) \cdot P_B(x)$
- Geometric average $P_\times(x) \propto P_A(x) \cdot P_B(x)$
- Averages can be computed within the BN formalism
- This makes easy to extend the averages to credal sets

  by using credal networks [i.e., experts give $K_A(X)$ and $K_B(X)$]

Averaging expert opinions

- Experts A and B give (precise knowledge) $P_A(X)$ and $P_B(X)$
- Arithmetic average $P_+(x) \propto P_A(x) \cdot P_B(x)$
- Geometric average $P_\times(x) \propto P_A(x) \cdot P_B(x)$
- Averages can be computed within the BN formalism
- This makes easy to extend the averages to credal sets

  by using credal networks [i.e., experts give $K_A(X)$ and $K_B(X)$]

Averaging expert opinions

- Experts A and B give (precise knowledge) $P_A(X)$ and $P_B(X)$
- Arithmetic average $P_+(x) \propto P_A(x) \cdot P_B(x)$
- Geometric average $P_\times(x) \propto P_A(x) \cdot P_B(x)$
- Averages can be computed within the BN formalism
- This makes easy to extend the averages to credal sets

  by using credal networks [i.e., experts give $K_A(X)$ and $K_B(X)$]

Arithmetic average (Bayesian case)

Input $P_A(X)$, $P_B(X)$

Define $X_A, X_B : \Omega_{X_A} = \Omega_{X_B} = \Omega_X$

Two (independent) expert opinions

X is the averaged knowledge

D auxiliary var indexing experts: $\Omega_D = \{A, B\}$

$$P(x|x_A, x_B, d) = \begin{cases} \delta(x_a, x) & \text{if} \quad D = a \\ \delta(x_B, x) & \text{if} \quad D = b \end{cases}$$

Uniform prior: $P(D = A) = P(D = B)$

$$P(X = x) = \sum_{x_A, x_B, d} P(x|x_A, x_B, d) P(x_A) P(x_B) P(d) = P_+(x)$$

Arithmetic average (Bayesian case)

Input $P_A(X)$, $P_B(X)$

Define $X_A, X_B : \Omega_{X_A} = \Omega_{X_B} = \Omega_X$

Two (independent) expert opinions

X is the averaged knowledge

D auxiliary var indexing experts: $\Omega_D = \{A, B\}$

$$P(x|x_A, x_B, d) = \begin{cases} \delta(x_a, x) & \text{if} \quad D = a \\ \delta(x_B, x) & \text{if} \quad D = b \end{cases}$$

Uniform prior: $P(D = A) = P(D = B)$

$$P(X = x) = \sum_{x_A, x_B, d} P(x|x_A, x_B, d)P(x_A)P(x_B)P(d) = P_+(x)$$

Arithmetic average (Bayesian case)

Input $P_A(X)$, $P_B(X)$

Define $X_A, X_B : \Omega_{X_A} = \Omega_{X_B} = \Omega_X$

Two (independent) expert opinions

X is the averaged knowledge

D auxiliary var indexing experts: $\Omega_D = \{A, B\}$

$$P(x|x_A, x_B, d) = \begin{cases} \delta(x_a, x) & \text{if} \quad D = a \\ \delta(x_B, x) & \text{if} \quad D = b \end{cases}$$

Uniform prior: $P(D = A) = P(D = B)$

$$P(X = x) = \sum_{x_A, x_B, d} P(x|x_A, x_B, d)P(x_A)P(x_B)P(d) = P_+(x)$$

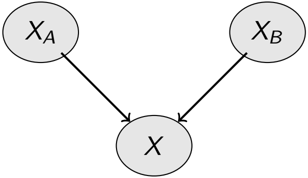
Arithmetic average (Bayesian case)

Input $P_A(X)$, $P_B(X)$

Define $X_A, X_B : \Omega_{X_A} = \Omega_{X_B} = \Omega_X$

Two (independent) expert opinions

X is the averaged knowledge



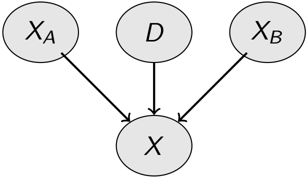
D auxiliary var indexing experts: $\Omega_D = \{A, B\}$

$$P(x|x_A, x_B, d) = \begin{cases} \delta(x_a, x) & \text{if} \quad D = a \\ \delta(x_B, x) & \text{if} \quad D = b \end{cases}$$

Uniform prior: $P(D = A) = P(D = B)$

$$P(X = x) = \sum_{x_A, x_B, d} P(x|x_A, x_B, d)P(x_A)P(x_B)P(d) = P_+(x)$$

Arithmetic average (Bayesian case)



Input $P_A(X)$, $P_B(X)$

Define $X_A, X_B : \Omega_{X_A} = \Omega_{X_B} = \Omega_X$

Two (independent) expert opinions

X is the averaged knowledge

D auxiliary var indexing experts: $\Omega_D = \{A, B\}$

$$P(x|x_A, x_B, d) = \begin{cases} \delta(x_a, x) & \text{if} \quad D = a \\ \delta(x_B, x) & \text{if} \quad D = b \end{cases}$$

Uniform prior: $P(D = A) = P(D = B)$

$$P(X = x) = \sum_{x_A, x_B, d} P(x|x_A, x_B, d) P(x_A) P(x_B) P(d) = P_+(x)$$

Arithmetic average (Bayesian case)

Input $P_A(X)$, $P_B(X)$

Define $X_A, X_B : \Omega_{X_A} = \Omega_{X_B} = \Omega_X$

Two (independent) expert opinions

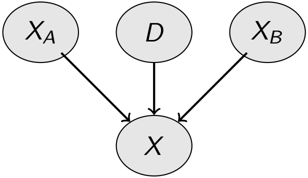
X is the averaged knowledge

D auxiliary var indexing experts: $\Omega_D = \{A, B\}$

$$P(x|x_A, x_B, d) = \begin{cases} \delta(x_a, x) & \text{if} \quad D = a \\ \delta(x_B, x) & \text{if} \quad D = b \end{cases}$$

Uniform prior: $P(D = A) = P(D = B)$

$P(X = x) = \sum_{x_A, x_B, d} P(x|x_A, x_B, d) P(x_A) P(x_B) P(d) = P_+(x)$



Arithmetic average (Bayesian case)

Input $P_A(X)$, $P_B(X)$

Define $X_A, X_B : \Omega_{X_A} = \Omega_{X_B} = \Omega_X$

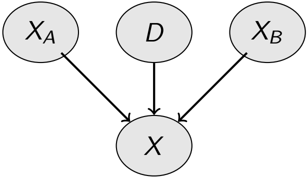
Two (independent) expert opinions

X is the averaged knowledge

D auxiliary var indexing experts: $\Omega_D = \{A, B\}$

$$P(x|x_A, x_B, d) = \begin{cases} \delta(x_a, x) & \text{if} \quad D = a \\ \delta(x_B, x) & \text{if} \quad D = b \end{cases}$$

Uniform prior: $P(D = A) = P(D = B)$



$P(X = x) = \sum_{x_A, x_B, d} P(x|x_A, x_B, d) P(x_A) P(x_B) P(d) = P_+(x)$

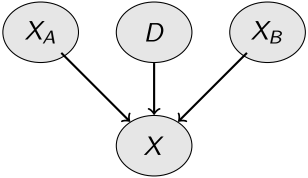
Arithmetic average (Bayesian case)

Input $P_A(X)$, $P_B(X)$

Define $X_A, X_B : \Omega_{X_A} = \Omega_{X_B} = \Omega_X$

Two (independent) expert opinions

X is the averaged knowledge



D auxiliary var indexing experts: $\Omega_D = \{A, B\}$

$$P(x|x_A, x_B, d) = \begin{cases} \delta(x_a, x) & \text{if} \quad D = a \\ \delta(x_B, x) & \text{if} \quad D = b \end{cases}$$

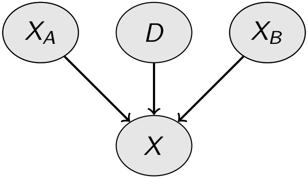
Uniform prior: $P(D = A) = P(D = B)$

$$P(X = x) = \sum_{x_A, x_B, d} P(x|x_A, x_B, d) P(x_A) P(x_B) P(d) = P_+(x)$$

Arithmetic average (credal case)

Same setup as before but non-uniform $P(D)$

$P(D = A)$ and $P(D = B)$ as experts' reliabilities



$P(D = A) = 1 \Rightarrow P_+(X) = P_A(X)$

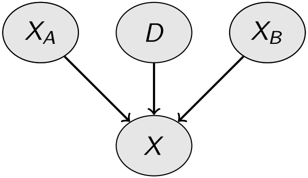
Prior ignorance $P(D = A) \in [0, 1]$ (vacuous)
$K_+(X) =$ convex closure of $\{P_A(X), P_B(X)\}$

Averaging imprecise expert knowledge,
i.e., $K_A(X)$ $K_B(X)$, vacuous case
$K_+(X) = K_A(X) \cup K_B(X)$

Arithmetic average (credal case)

Same setup as before but non-uniform $P(D)$

$P(D = A)$ and $P(D = B)$ as experts' reliabilities



$P(D = A) = 1 \Rightarrow P_+(X) = P_A(X)$

Prior ignorance $P(D = A) \in [0, 1]$ (vacuous)
$K_+(X) =$ convex closure of $\{P_A(X), P_B(X)\}$

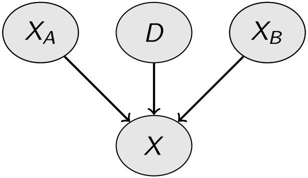
Averaging imprecise expert knowledge,
i.e., $K_A(X)$ $K_B(X)$, vacuous case
$K_+(X) = K_A(X) \cup K_B(X)$

Arithmetic average (credal case)

Same setup as before but non-uniform $P(D)$

$P(D = A)$ and $P(D = B)$ as experts' reliabilities

$P(D = A) = 1 \Rightarrow P_+(X) = P_A(X)$



Prior ignorance $P(D = A) \in [0, 1]$ (vacuous)
$K_+(X) =$ convex closure of $\{P_A(X), P_B(X)\}$

Averaging imprecise expert knowledge,
i.e., $K_A(X)$ $K_B(X)$, vacuous case
$K_+(X) = K_A(X) \cup K_B(X)$

Arithmetic average (credal case)

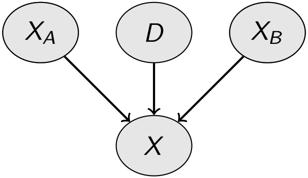
Same setup as before but non-uniform $P(D)$

$P(D = A)$ and $P(D = B)$ as experts' reliabilities

$P(D = A) = 1 \Rightarrow P_+(X) = P_A(X)$

Prior ignorance $P(D = A) \in [0, 1]$ (vacuous)
$K_+(X) =$ convex closure of $\{P_A(X), P_B(X)\}$

Averaging imprecise expert knowledge,
i.e., $K_A(X)$ $K_B(X)$, vacuous case
$K_+(X) = K_A(X) \cup K_B(X)$

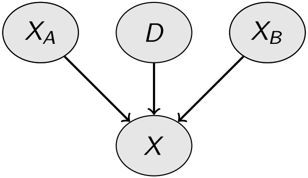


Arithmetic average (credal case)

Same setup as before but non-uniform $P(D)$

$P(D = A)$ and $P(D = B)$ as experts' reliabilities

$P(D = A) = 1 \Rightarrow P_+(X) = P_A(X)$



Prior ignorance $P(D = A) \in [0, 1]$ (vacuous)
$K_+(X) =$ convex closure of $\{P_A(X), P_B(X)\}$

Averaging imprecise expert knowledge,
i.e., $K_A(X)$ $K_B(X)$, vacuous case
$K_+(X) = K_A(X) \cup K_B(X)$

Geometric average

Again two experts opinions $P(X_A) := P_A(X)$ and $P(X_B) := P_B(X)$

Modeling that X_a and X_B are the same variable

Logical constraints in BNs:
a dummy Boolean child D
true IFF $X_A = X_B$

$$P(X_A = x|D = T) = \frac{\sum_{x_B} P(D{=}T|X_A{=}x,X_B{=}x')P(X_A{=}x)P(X_B{=}x')}{\sum_{x_A,x_B}\cdots} = P_\times(x)$$

We can geometrically average credal sets!

Geometric average

Again two experts opinions $P(X_A) := P_A(X)$ and $P(X_B) := P_B(X)$

Modeling that X_a and X_B are the same variable

Logical constraints in BNs:
a dummy Boolean child D
true IFF $X_A = X_B$

$$P(X_A = x|D = T) = \frac{\sum_{x_B} P(D=T|X_A=x,X_B=x')P(X_A=x)P(X_B=x')}{\sum_{x_A,x_B}\cdots} = P_\times(x)$$

We can geometrically average credal sets!

Geometric average

Again two experts opinions $P(X_A) := P_A(X)$ and $P(X_B) := P_B(X)$

Modeling that X_a and X_B are the same variable

Logical constraints in BNs:
a dummy Boolean child D
true IFF $X_A = X_B$

$$P(X_A = x|D = T) = \frac{\sum_{x_B} P(D{=}T|X_A{=}x,X_B{=}x')P(X_A{=}x)P(X_B{=}x')}{\sum_{x_A,x_B}\cdots} = P_\times(x)$$

We can geometrically average credal sets!

Geometric average

Again two experts opinions $P(X_A) := P_A(X)$ and $P(X_B) := P_B(X)$

Modeling that X_a and X_B are the same variable

Logical constraints in BNs:
a dummy Boolean child D
true IFF $X_A = X_B$

$$P(X_A = x|D = T) = \frac{\sum_{x_B} P(D{=}T|X_A{=}x,X_B{=}x')P(X_A{=}x)P(X_B{=}x')}{\sum_{x_A,x_B}\cdots} = P_\times(x)$$

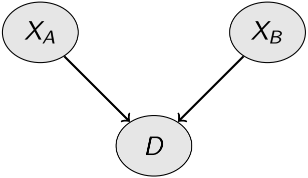
We can geometrically average credal sets!

Geometric average

Again two experts opinions $P(X_A) := P_A(X)$ and $P(X_B) := P_B(X)$

Modeling that X_a and X_B are the same variable

Logical constraints in BNs:
a dummy Boolean child D
true IFF $X_A = X_B$



$$P(X_A = x|D = T) = \frac{\sum_{x_B} P(D{=}T|X_A{=}x, X_B{=}x')P(X_A{=}x)P(X_B{=}x')}{\sum_{x_A, x_B} \cdots} = P_\times(x)$$

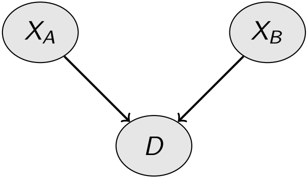
We can geometrically average credal sets!

Geometric average

Again two experts opinions $P(X_A) := P_A(X)$ and $P(X_B) := P_B(X)$

Modeling that X_a and X_B are the same variable

Logical constraints in BNs:
a dummy Boolean child D
true IFF $X_A = X_B$



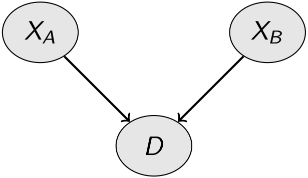
$$P(X_A = x|D = T) = \frac{\sum_{x_B} P(D{=}T|X_A{=}x, X_B{=}x')P(X_A{=}x)P(X_B{=}x')}{\sum_{x_A,x_B} \cdots} = P_\times(x)$$

We can geometrically average credal sets!

Geometric average

Again two experts opinions $P(X_A) := P_A(X)$ and $P(X_B) := P_B(X)$

Modeling that X_a and X_B are the same variable



Logical constraints in BNs:
a dummy Boolean child D
true IFF $X_A = X_B$

$$P(X_A = x|D = T) = \frac{\sum_{x_B} P(D{=}T|X_A{=}x, X_B{=}x')P(X_A{=}x)P(X_B{=}x')}{\sum_{x_A,x_B} \cdots} = P_\times(x)$$

We can geometrically average credal sets!

Some experiments

- Tests with Boolean X
- Credal sets over Booleans are equivalent to intervals
  $P(X = T) \in [l, u] \Rightarrow K(X) = \left\{[l, 1 - l]^T, [u, 1 - u]^T\right\}$

$P(X_a = x)$	$P(X_b = x)$	$P(D = a)$	$P_+(x)$	$P_\times(x)$
.20	.40	.50	.30	.14
.20	.80	.50	.50	.50
[.10, .30]	[.30, .50]	.50	[.20, .40]	[.05, .30]
[.10, .30]	[.70, .90]	.50	[.40, .60]	[.21, .79]
[.10, .30]	[.70, .90]	[0, 1]	[.10, .90]	[.21, .79]

- Bayesian results included in the credal ones
- More cautious approach to (high-level) information fusion

Some experiments

- Tests with Boolean X
- Credal sets over Booleans are equivalent to intervals
  $P(X = T) \in [l, u] \Rightarrow K(X) = \left\{[l, 1-l]^T, [u, 1-u]^T\right\}$

$P(X_a = x)$	$P(X_b = x)$	$P(D = a)$	$P_+(x)$	$P_\times(x)$
.20	.40	.50	.30	.14
.20	.80	.50	.50	.50
[.10, .30]	[.30, .50]	.50	[.20, .40]	[.05, .30]
[.10, .30]	[.70, .90]	.50	[.40, .60]	[.21, .79]
[.10, .30]	[.70, .90]	[0, 1]	[.10, .90]	[.21, .79]

- Bayesian results included in the credal ones
- More cautious approach to (high-level) information fusion

Some experiments

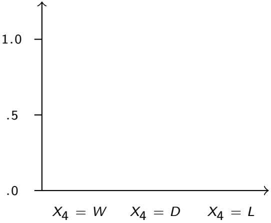
- Tests with Boolean X
- Credal sets over Booleans are equivalent to intervals $P(X = T) \in [l, u] \Rightarrow K(X) = \left\{[l, 1-l]^T, [u, 1-u]^T\right\}$

$P(X_a = x)$	$P(X_b = x)$	$P(D = a)$	$P_+(x)$	$P_\times(x)$
.20	.40	.50	.30	.14
.20	.80	.50	.50	.50
[.10, .30]	[.30, .50]	.50	[.20, .40]	[.05, .30]
[.10, .30]	[.70, .90]	.50	[.40, .60]	[.21, .79]
[.10, .30]	[.70, .90]	[0, 1]	[.10, .90]	[.21, .79]

- Bayesian results included in the credal ones
- More cautious approach to (high-level) information fusion

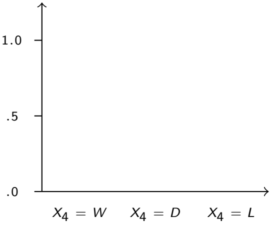
Decision Making on CNs

- Update beliefs about X_q (query)
  after the observation of x_E (evidence)
  What about the state of X_q?
- BN updating: compute $P(X_q|x_E)$
  State of X_q: $x_q^* = \arg\max_{x_q \in \Omega_{X_q}} P(x_q|x_E)$
- CN updating: compute $K(X_q|x_E)$?
  Algorithms only compute $\underline{P}(X_q|x_E)$
- State(s) of X_q by interval dominance
  $\Omega^*_{X_q} = \left\{ x_q \middle| \nexists x'_q \text{ s.t. } \underline{P}(x'_q|x_E) > \overline{P}(x_q|x_E) \right\}$
- More informative criterion: maximality
  $\left\{ x_q \middle| \nexists x'_q \text{ s.t. } P(x'_q|x_E) > P(x_q|x_E) \forall P(X_q|x_E) \in K(X_q|x_E) \right\}$
- Maximality can be computed with an auxiliary dummy child and standard updating (Antonucci & de Campos



Decision Making on CNs

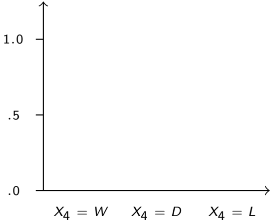
- Update beliefs about X_q (query)
  after the observation of x_E (evidence)
  What about the state of X_q?
- BN updating: compute $P(X_q|x_E)$
  State of X_q: $x_q^* = \arg\max_{x_q \in \Omega_{X_q}} P(x_q|x_E)$
- CN updating: compute $K(X_q|x_E)$?
  Algorithms only compute $\underline{P}(X_q|x_E)$
- State(s) of X_q by interval dominance
  $\Omega^*_{X_q} = \left\{ x_q \middle| \nexists x'_q \text{ s.t. } \underline{P}(x'_q|x_E) > \overline{P}(x_q|x_E) \right\}$
- More informative criterion: maximality
  $\left\{ x_q \middle| \nexists x'_q \text{ s.t. } P(x'_q|x_E) > P(x_q|x_E) \forall P(X_q|x_E) \in K(X_q|x_E) \right\}$
- Maximality can be computed with an auxiliary dummy child and standard updating *(Antonucci & de Campos*



Spain wins with high temperature?

Decision Making on CNs

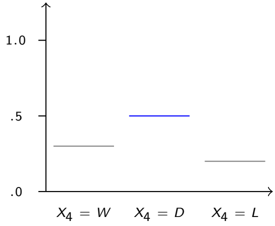
- Update beliefs about X_q (query)
  after the observation of x_E (evidence)
  What about the state of X_q?
- BN updating: compute $P(X_q|x_E)$
  State of X_q: $x_q^* = \arg\max_{x_q \in \Omega_{X_q}} P(x_q|x_E)$
- CN updating: compute $K(X_q|x_E)$?
  Algorithms only compute $\underline{P}(X_q|x_E)$
- State(s) of X_q by interval dominance
  $\Omega^*_{X_q} = \left\{ x_q \middle| \, \nexists x'_q \, \text{s.t.} \, \underline{P}(x'_q|x_E) > \overline{P}(x_q|x_E) \right\}$
- More informative criterion: maximality
  $\left\{ x_q \middle| \, \nexists x'_q \, \text{s.t.} \, P(x'_q|x_E) > P(x_q|x_E) \forall P(X_q|x_E) \in K(X_q|x_E) \right\}$
- Maximality can be computed with an auxiliary dummy child and standard updating *(Antonucci & de Campos*



Spain wins with high temperature?

Decision Making on CNs

- Update beliefs about X_q (query) after the observation of x_E (evidence) What about the state of X_q?
- BN updating: compute $P(X_q|x_E)$ State of X_q: $x_q^* = \arg\max_{x_q \in \Omega_{X_q}} P(x_q|x_E)$
- CN updating: compute $K(X_q|x_E)$? Algorithms only compute $\underline{P}(X_q|x_E)$
- State(s) of X_q by interval dominance $\Omega^*_{X_q} = \left\{ x_q \middle| \nexists x'_q \text{ s.t. } \underline{P}(x'_q|x_E) > \overline{P}(x_q|x_E) \right\}$
- More informative criterion: maximality $\left\{ x_q \middle| \nexists x'_q \text{ s.t. } P(x'_q|x_E) > P(x_q|x_E) \forall P(X_q|x_E) \in K(X_q|x_E) \right\}$
- Maximality can be computed with an auxiliary dummy child and standard updating *(Antonucci & de Campos*

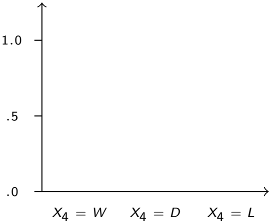


Spain wins with high temperature?

$$P(X_4|x_1) = \begin{bmatrix} .3 \\ .5 \\ .2 \end{bmatrix}$$

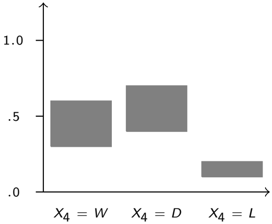
Decision Making on CNs

- Update beliefs about X_q (query) after the observation of x_E (evidence)
  What about the state of X_q?
- BN updating: compute $P(X_q|x_E)$
  State of X_q: $x_q^* = \arg\max_{x_q \in \Omega_{X_q}} P(x_q|x_E)$
- CN updating: compute $K(X_q|x_E)$?
  Algorithms only compute $\underline{P}(X_q|x_E)$
- State(s) of X_q by interval dominance
  $\Omega^*_{X_q} = \left\{ x_q \middle| \nexists x'_q \text{ s.t. } \underline{P}(x'_q|x_E) > \overline{P}(x_q|x_E) \right\}$
- More informative criterion: maximality
  $\left\{ x_q \middle| \nexists x'_q \text{ s.t. } P(x'_q|x_E) > P(x_q|x_E) \forall P(X_q|x_E) \in K(X_q|x_E) \right\}$
- Maximality can be computed with an auxiliary dummy child and standard updating *(Antonucci & de Campos*



Decision Making on CNs

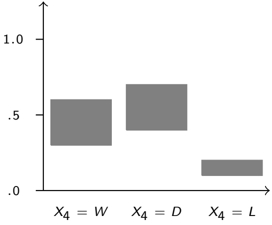
- Update beliefs about X_q (query)
  after the observation of x_E (evidence)
  What about the state of X_q?
- BN updating: compute $P(X_q|x_E)$
  State of X_q: $x_q^* = \arg\max_{x_q \in \Omega_{X_q}} P(x_q|x_E)$
- CN updating: compute $K(X_q|x_E)$?
  Algorithms only compute $\underline{P}(X_q|x_E)$
- State(s) of X_q by interval dominance
  $\Omega^*_{X_q} = \left\{ x_q \middle| \nexists x'_q \text{ s.t. } \underline{P}(x'_q|x_E) > \overline{P}(x_q|x_E) \right\}$
- More informative criterion: maximality
  $\left\{ x_q \middle| \nexists x'_q \text{ s.t. } P(x'_q|x_E) > P(x_q|x_E) \forall P(X_q|x_E) \in K(X_q|x_E) \right\}$
- Maximality can be computed with an auxiliary dummy child and standard updating *(Antonucci & de Campos*



$$P(X_4|x_1) \in \begin{array}{c} [.3, .6] \\ [.4, .7] \\ [.1, .2] \end{array}$$

Decision Making on CNs

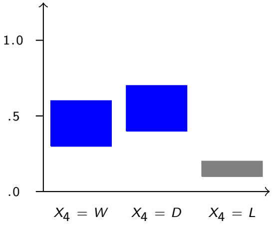
- Update beliefs about X_q (query)
  after the observation of x_E (evidence)
  What about the state of X_q?
- BN updating: compute $P(X_q|x_E)$
  State of X_q: $x_q^* = \arg\max_{x_q \in \Omega_{X_q}} P(x_q|x_E)$
- CN updating: compute $K(X_q|x_E)$?
  Algorithms only compute $\underline{P}(X_q|x_E)$
- State(s) of X_q by interval dominance
  $\Omega^*_{X_q} = \left\{ x_q \middle| \nexists x'_q \text{ s.t. } \underline{P}(x'_q|x_E) > \overline{P}(x_q|x_E) \right\}$
- More informative criterion: maximality
  $\left\{ x_q \middle| \nexists x'_q \text{ s.t. } P(x'_q|x_E) > P(x_q|x_E) \forall P(X_q|x_E) \in K(X_q|x_E) \right\}$
- Maximality can be computed with an auxiliary dummy child and standard updating *(Antonucci & de Campos*



$$P(X_4|x_1) \in \begin{array}{c} [.3,\ .6] \\ [.4,\ .7] \\ [.1,\ .2] \end{array}$$

Decision Making on CNs

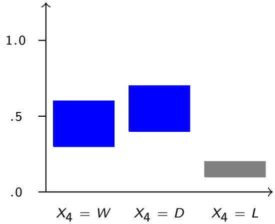
- Update beliefs about X_q (query) after the observation of x_E (evidence) What about the state of X_q?
- BN updating: compute $P(X_q|x_E)$ State of X_q: $x_q^* = \arg\max_{x_q \in \Omega_{X_q}} P(x_q|x_E)$
- CN updating: compute $K(X_q|x_E)$? Algorithms only compute $\underline{P}(X_q|x_E)$
- State(s) of X_q by interval dominance $\Omega^*_{X_q} = \left\{ x_q \middle| \nexists x'_q \,\text{s.t.}\, \underline{P}(x'_q|x_E) > \overline{P}(x_q|x_E) \right\}$
- More informative criterion: maximality $\left\{ x_q \middle| \nexists x'_q \,\text{s.t.}\, P(x'_q|x_E) > P(x_q|x_E) \forall P(X_q|x_E) \in K(X_q|x_E) \right\}$
- Maximality can be computed with an auxiliary dummy child and standard updating (*Antonucci & de Campos*



$$P(X_4|x_1) \in \begin{matrix} [.3, .6] \\ [.4, .7] \\ [.1, .2] \end{matrix}$$

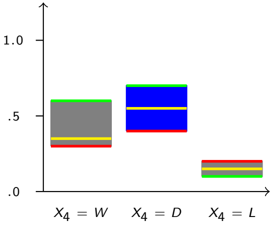
Decision Making on CNs

- Update beliefs about X_q (query)
  after the observation of x_E (evidence)
  What about the state of X_q?
- BN updating: compute $P(X_q|x_E)$
  State of X_q: $x_q^* = \arg\max_{x_q \in \Omega_{X_q}} P(x_q|x_E)$
- CN updating: compute $K(X_q|x_E)$?
  Algorithms only compute $\underline{P}(X_q|x_E)$
- State(s) of X_q by interval dominance
  $\Omega^*_{X_q} = \left\{ x_q \middle| \; \nexists x'_q \, \text{s.t.} \, \underline{P}(x'_q|x_E) > \overline{P}(x_q|x_E) \right\}$
- More informative criterion: maximality
  $\left\{ x_q \middle| \; \nexists x'_q \, \text{s.t.} \, P(x'_q|x_E) > P(x_q|x_E) \forall P(X_q|x_E) \in K(X_q|x_E) \right\}$
- Maximality can be computed with an
  auxiliary dummy child and standard
  updating (Antonucci & de Campos



$$P(X_4|x_1) \in \begin{matrix}[.3, .6]\\ [.4, .7]\\ [.1, .2]\end{matrix}$$

Decision Making on CNs



$$P(X_4|x_1) \in \begin{array}{c}[.3, .6]\\ [.4, .7]\\ [.1, .2]\end{array}$$

- Update beliefs about X_q (query)

  after the observation of x_E (evidence)

  What about the state of X_q?

- BN updating: compute $P(X_q|x_E)$

  State of X_q: $x_q^* = \arg\max_{x_q \in \Omega_{X_q}} P(x_q|x_E)$

- CN updating: compute $K(X_q|x_E)$?

  Algorithms only compute $\underline{P}(X_q|x_E)$

- State(s) of X_q by interval dominance

  $\Omega^*_{X_q} = \left\{ x_q \middle| \nexists x'_q \text{ s.t. } \underline{P}(x'_q|x_E) > \overline{P}(x_q|x_E) \right\}$

- More informative criterion: maximality

  $\left\{ x_q \middle| \nexists x'_q \text{ s.t. } P(x'_q|x_E) > P(x_q|x_E) \forall P(X_q|x_E) \in K(X_q|x_E) \right\}$

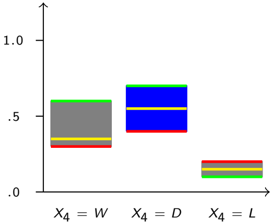
- Maximality can be computed with an

  auxiliary dummy child and standard

  updating *(Antonucci & de Campos*

Decision Making on CNs

- Update beliefs about X_q (query)
  after the observation of x_E (evidence)
  What about the state of X_q?
- BN updating: compute $P(X_q|x_E)$
  State of X_q: $x_q^* = \arg\max_{x_q \in \Omega_{X_q}} P(x_q|x_E)$
- CN updating: compute $K(X_q|x_E)$?
  Algorithms only compute $\underline{P}(X_q|x_E)$
- State(s) of X_q by interval dominance
  $\Omega^*_{X_q} = \left\{ x_q \middle| \; \nexists x'_q \text{ s.t. } \underline{P}(x'_q|x_E) > \overline{P}(x_q|x_E) \right\}$
- More informative criterion: maximality
  $\left\{ x_q \middle| \; \nexists x'_q \text{ s.t. } P(x'_q|x_E) > P(x_q|x_E) \forall P(X_q|x_E) \in K(X_q|x_E) \right\}$
- Maximality can be computed with an auxiliary dummy child and standard updating (Antonucci & de Campos



$$P(X_4|x_1) \in \begin{array}{c} [.3,\ .6] \\ [.4,\ .7] \\ [.1,\ .2] \end{array}$$

A military application: no-fly zones surveillance

- Around important potential targets
  (eg. WEF, dams, nuke plants)
- Twofold circle wraps the target
  - External no-fly zone (sensors)
  - Internal no-fly zone (anti-air units)
- An aircraft entering the zone
  (to be called **intruder**)
- Its presence, speed, height, and other features revealed by the sensors
- A team of military experts decides:
  - what the intruder intends to do
    (external zone / credal level)
  - what to do with the intruder
    (internal zone / pignistic level)

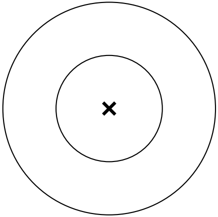
davos

dam

npp

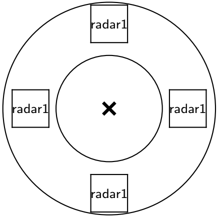
A military application: no-fly zones surveillance

- Around important potential targets
  (eg. WEF, dams, nuke plants)
- Twofold circle wraps the target
  - External no-fly zone (sensors)
  - Internal no-fly zone (anti-air units)
- An aircraft entering the zone
  (to be called **intruder**)
- Its presence, speed, height, and other features revealed by the sensors
- A team of military experts decides:
  - what the intruder intends to do
    (external zone / credal level)
  - what to do with the intruder
    (internal zone / pignistic level)



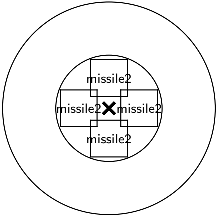
A military application: no-fly zones surveillance

- Around important potential targets (eg. WEF, dams, nuke plants)
- Twofold circle wraps the target
    - External no-fly zone (sensors)
    - Internal no-fly zone (anti-air units)
- An aircraft entering the zone (to be called **intruder**)
- Its presence, speed, height, and other features revealed by the sensors
- A team of military experts decides:
    - what the intruder intends to do (external zone / credal level)
    - what to do with the intruder (internal zone / pignistic level)



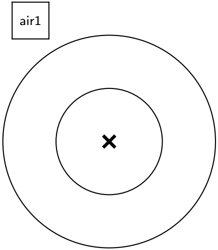
A military application: no-fly zones surveillance

- Around important potential targets
  (eg. WEF, dams, nuke plants)
- Twofold circle wraps the target
  - External no-fly zone (sensors)
  - Internal no-fly zone (anti-air units)
- An aircraft entering the zone
  (to be called **intruder**)
- Its presence, speed, height, and other features revealed by the sensors
- A team of military experts decides:
  - what the intruder intends to do
    (external zone / credal level)
  - what to do with the intruder
    (internal zone / pignistic level)



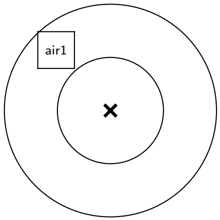
A military application: no-fly zones surveillance

- Around important potential targets
  (eg. WEF, dams, nuke plants)
- Twofold circle wraps the target
  - External no-fly zone (sensors)
  - Internal no-fly zone (anti-air units)
- An aircraft entering the zone
  (to be called **intruder**)
- Its presence, speed, height, and other features revealed by the sensors
- A team of military experts decides:
  - what the intruder intends to do
    (external zone / credal level)
  - what to do with the intruder
    (internal zone / pignistic level)



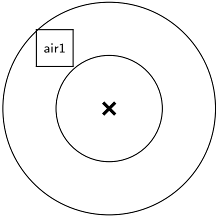
A military application: no-fly zones surveillance

- Around important potential targets (eg. WEF, dams, nuke plants)
- Twofold circle wraps the target
    - External no-fly zone (sensors)
    - Internal no-fly zone (anti-air units)
- An aircraft entering the zone (to be called **intruder**)
- Its presence, speed, height, and other features revealed by the sensors
- A team of military experts decides:
    - what the intruder intends to do (external zone / credal level)
    - what to do with the intruder (internal zone / pignistic level)



A military application: no-fly zones surveillance

- Around important potential targets
  (eg. WEF, dams, nuke plants)
- Twofold circle wraps the target
    - External no-fly zone (sensors)
    - Internal no-fly zone (anti-air units)
- An aircraft entering the zone
  (to be called **intruder**)
- Its presence, speed, height, and other features revealed by the sensors
- A team of military experts decides:
    - what the intruder intends to do
      (external zone / credal level)
    - what to do with the intruder
      (internal zone / pignistic level)



Identifying intruder's goal

- Four (exclusive and exhaustive) options for intruder's goal:

plane	rust	damaged	erroneous
renegade	provocateur	damaged	erroneous

- This identification is difficult
  - Sensors reliabilities are affected by geo/wheather conditions
  - Information fusion from several sensors

Identifying intruder's goal

- Four (exclusive and exhaustive) options for intruder's goal:

plane rust damaged erroneous

renegade provocateur damaged erroneous

- This identification is difficult
  - Sensors reliabilities are affected by geo/wheather conditions
  - Information fusion from several sensors

Identifying intruder's goal

- Four (exclusive and exhaustive) options for intruder's goal:

plane rust damaged erroneous

renegade provocateur damaged erroneous

- This identification is difficult
  - Sensors reliabilities are affected by geo/wheather conditions
  - Information fusion from several sensors

Identifying intruder's goal

- Four (exclusive and exhaustive) options for intruder's goal:

plane	rust	damaged	erroneous
renegade	provocateur	damaged	erroneous

- This identification is difficult
  - Sensors reliabilities are affected by geo/wheather conditions
  - Information fusion from several sensors

Why credal networks?

- Why a probabilistic model?
  - No deterministic relations between the different variables
  - Pervasive uncertainty in the observations
- Why a graphical model?
  - Many independence relations among the different variables
- Why an imprecise (probabilistic) model?
  - Expert evaluations are mostly based on qualitative judgements
  - The model should be (over)cautious

Why credal networks?

- Why a probabilistic model?
  - No deterministic relations between the different variables
  - Pervasive uncertainty in the observations
- Why a graphical model?
  - Many independence relations among the different variables
- Why an imprecise (probabilistic) model?
  - Expert evaluations are mostly based on qualitative judgements
  - The model should be (over)cautious

Why credal networks?

- Why a probabilistic model?
  - No deterministic relations between the different variables
  - Pervasive uncertainty in the observations
- Why a graphical model?
  - Many independence relations among the different variables
- Why an imprecise (probabilistic) model?
  - Expert evaluations are mostly based on qualitative judgements
  - The model should be (over)cautious

Why credal networks?

- Why a probabilistic model?
    - No deterministic relations between the different variables
    - Pervasive uncertainty in the observations
- Why a graphical model?
    - Many independence relations among the different variables
- Why an imprecise (probabilistic) model?
    - Expert evaluations are mostly based on qualitative judgements
    - The model should be (over)cautious

Why credal networks?

- Why a probabilistic model?
    - No deterministic relations between the different variables
    - Pervasive uncertainty in the observations
- Why a graphical model?
    - Many independence relations among the different variables
- Why an imprecise (probabilistic) model?
    - Expert evaluations are mostly based on qualitative judgements
    - The model should be (over)cautious

Why credal networks?

- Why a probabilistic model?
  - No deterministic relations between the different variables
  - Pervasive uncertainty in the observations
- Why a graphical model?
  - Many independence relations among the different variables
- Why an imprecise (probabilistic) model?
  - Expert evaluations are mostly based on qualitative judgements
  - The model should be (over)cautious

Why credal networks?

- Why a probabilistic model?
  - No deterministic relations between the different variables
  - Pervasive uncertainty in the observations
- Why a graphical model?
  - Many independence relations among the different variables
- Why an imprecise (probabilistic) model?
  - Expert evaluations are mostly based on qualitative judgements
  - The model should be (over)cautious

Why credal networks?

- Why a probabilistic model?
  - No deterministic relations between the different variables
  - Pervasive uncertainty in the observations
- Why a graphical model?
  - Many independence relations among the different variables
- Why an imprecise (probabilistic) model?
  - Expert evaluations are mostly based on qualitative judgements
  - The model should be (over)cautious

Why credal networks?

- Why a probabilistic model?
  - No deterministic relations between the different variables
  - Pervasive uncertainty in the observations
- Why a graphical model?
  - Many independence relations among the different variables
- Why an imprecise (probabilistic) model?
  - Expert evaluations are mostly based on qualitative judgements
  - The model should be (over)cautious

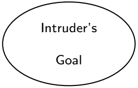
Network core

Network core

- Intruder's goal and features as categorical variables

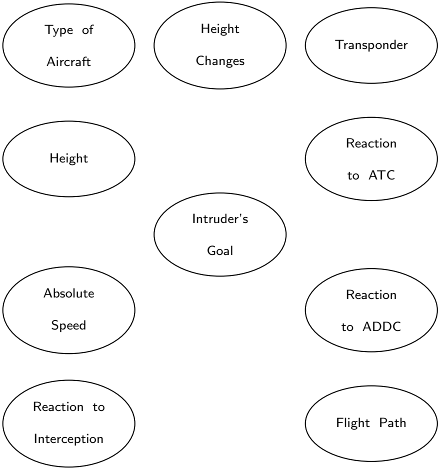
Network core

- Intruder's goal and features as categorical variables



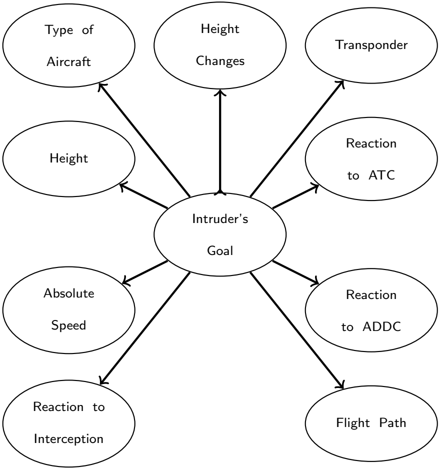
Network core

- Intruder's goal and features as categorical variables



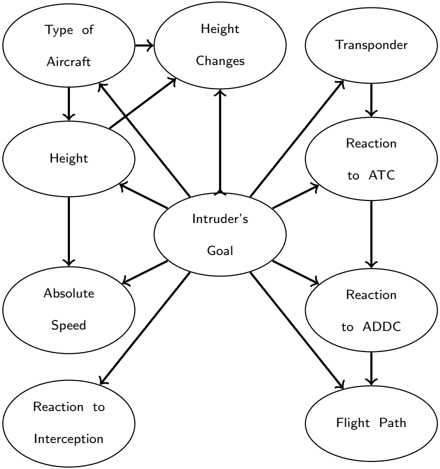
Network core

- Intruder's goal and features as categorical variables
- Independencies depicted by a directed graph (acyclic)



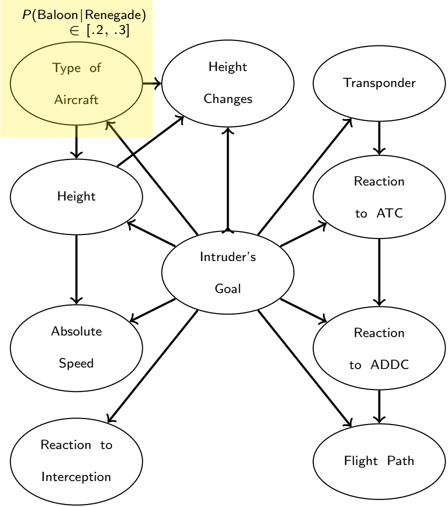
Network core

- Intruder's goal and features as categorical variables
- Independencies depicted by a directed graph (acyclic)



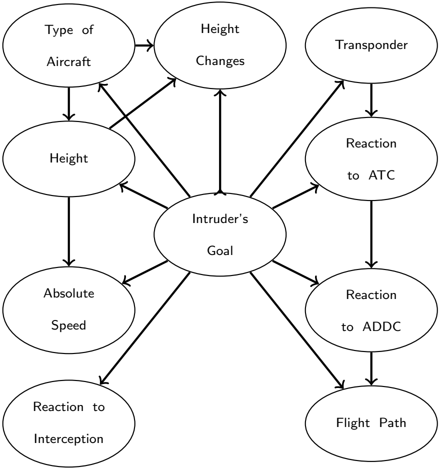
Network core

- Intruder's goal and features as categorical variables
- Independencies depicted by a directed graph (acyclic)
- Experts provide interval-valued probabilistic assessments, we compute credal sets



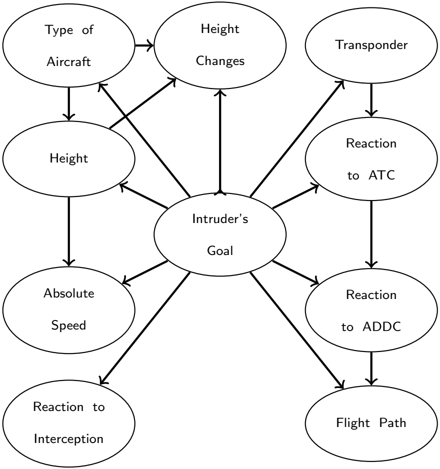
Network core

- Intruder's goal and features as categorical variables
- Independencies depicted by a directed graph (acyclic)
- Experts provide interval-valued probabilistic assessments, we compute credal sets
- A (small) credal network



Network core

- Intruder's goal and features as categorical variables
- Independencies depicted by a directed graph (acyclic)
- Experts provide interval-valued probabilistic assessments, we compute credal sets
- A (small) credal network
- Complex observation process!



Observations modelling and fusion

- Each sensor modeled by an auxiliary child of the (ideal) variable to be observed
- $P(\text{sensor}|\text{ideal})$ models sensor reliability

  (eg. identity matrix = perfectly reliable sensor)
- Many sensors? Many children!

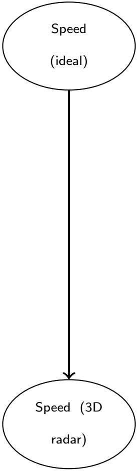
  (conditional independence between sensors given the ideal)

Observations modelling and fusion

- Each sensor modeled by an auxiliary child of the (ideal) variable to be observed
- $P(\text{sensor}|\text{ideal})$ models sensor reliability

  (eg. identity matrix = perfectly reliable sensor)
- Many sensors? Many children!

  (conditional independence between sensors given the ideal)

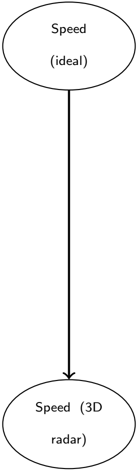


Observations modelling and fusion

- Each sensor modeled by an auxiliary child of the (ideal) variable to be observed
- $P(\text{sensor}|\text{ideal})$ models sensor reliability

  (eg. identity matrix = perfectly reliable sensor)
- Many sensors? Many children!

  (conditional independence between sensors given the ideal)

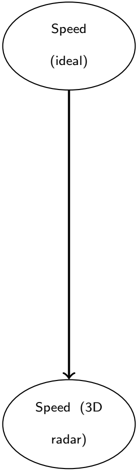


Observations modelling and fusion

- Each sensor modeled by an auxiliary child of the (ideal) variable to be observed
- $P(\text{sensor}|\text{ideal})$ models sensor reliability

  (eg. identity matrix = perfectly reliable sensor)
- Many sensors? Many children!

  (conditional independence between sensors given the ideal)

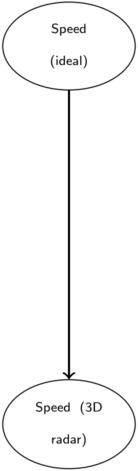


Observations modelling and fusion

- Each sensor modeled by an auxiliary child of the (ideal) variable to be observed
- $P(\text{sensor}|\text{ideal})$ models sensor reliability

  (eg. identity matrix = perfectly reliable sensor)
- Many sensors? Many children!

  (conditional independence between sensors given the ideal)

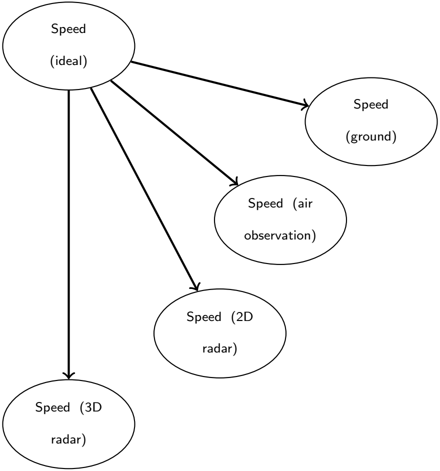


Observations modelling and fusion

- Each sensor modeled by an auxiliary child of the (ideal) variable to be observed
- $P(\text{sensor}|\text{ideal})$ models sensor reliability

  (eg. identity matrix = perfectly reliable sensor)
- Many sensors? Many children!

  (conditional independence between sensors given the ideal)

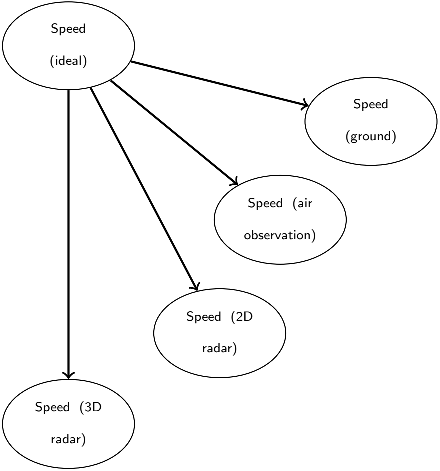


Observations modelling and fusion

- Each sensor modeled by an auxiliary child of the (ideal) variable to be observed
- $P(\text{sensor}|\text{ideal})$ models sensor reliability

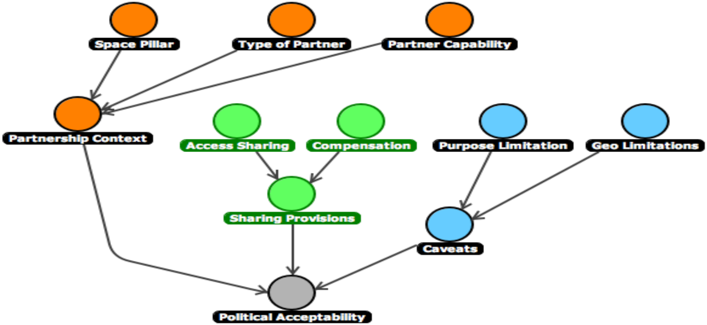
  (eg. identity matrix = perfectly reliable sensor)
- Many sensors? Many children!

  (conditional independence between sensors given the ideal)



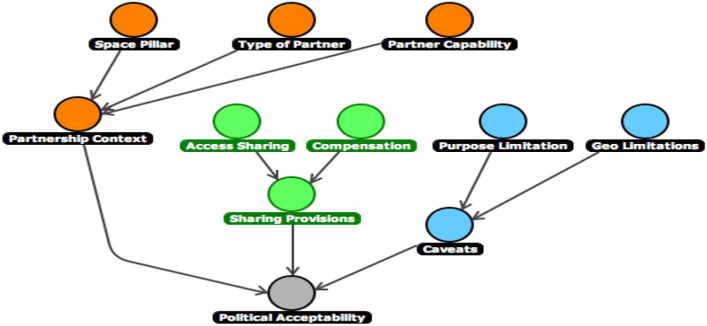
The whole network

- A huge multiply-connected credal network
- Efficient (approximate) updating with GL2U



The whole network

- A huge multiply-connected credal network
- Efficient (approximate) updating with GL2U



Simulations

- Simulating a dam in the Swiss Alps, with no interceptors, relatively good coverage for other sensors, discontinuous low clouds and daylight
- Sensors return:
    - Height = very low / very low / very low / low
    - Type = helicopter / helicopter
    - Flight Path = U-path / U-path / U-path / U-path / U-path / missing
    - Height Changes = descent / descent / descent / descent / missing
    - Speed = slow / slow / slow / slow / slow
    - ADDC reaction = positive / positive / positive / positive / positive / positive
- We reject renegade and damaged, but indecision between provocateur and erroneous
- Assuming higher levels of reliability
- The aircraft is a provocateur!

Simulations

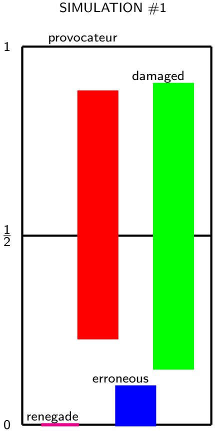
- Simulating a dam in the Swiss Alps, with no interceptors, relatively good coverage for other sensors, discontinuous low clouds and daylight
- Sensors return:
  - Height = very low / very low / very low / low
  - Type = helicopter / helicopter
  - Flight Path = U-path / U-path / U-path / U-path / U-path / missing
  - Height Changes = descent / descent / descent / descent / missing
  - Speed = slow / slow / slow / slow / slow
  - ADDC reaction = positive / positive / positive / positive / positive / positive
- We reject renegade and damaged, but indecision between provocateur and erroneous
- Assuming higher levels of reliability
- The aircraft is a provocateur!

Simulations

- Simulating a dam in the Swiss Alps, with no interceptors, relatively good coverage for other sensors, discontinuous low clouds and daylight
- Sensors return:
  - Height = very low / very low / very low / low
  - Type = helicopter / helicopter
  - Flight Path = U-path / U-path / U-path / U-path / U-path / missing
  - Height Changes = descent / descent / descent / descent / missing
  - Speed = slow / slow / slow / slow / slow
  - ADDC reaction = positive / positive / positive / positive / positive / positive
- We reject renegade and damaged, but indecision between provocateur and erroneous
- Assuming higher levels of reliability
- The aircraft is a provocateur!

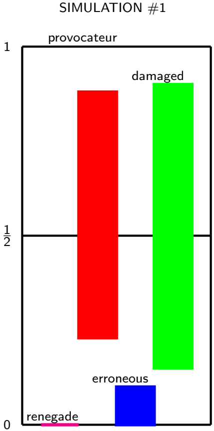
Simulations

- Simulating a dam in the Swiss Alps, with no interceptors, relatively good coverage for other sensors, discontinuous low clouds and daylight
- Sensors return:
  - Height = very low / very low / very low / low
  - Type = helicopter / helicopter
  - Flight Path = U-path / U-path / U-path / U-path / U-path / missing
  - Height Changes = descent / descent / descent / descent / missing
  - Speed = slow / slow / slow / slow / slow
  - ADDC reaction = positive / positive / positive / positive / positive / positive
- We reject renegade and damaged, but indecision between provocateur and erroneous
- Assuming higher levels of reliability
- The aircraft is a provocateur!



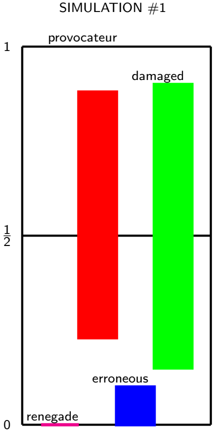
Simulations

- Simulating a dam in the Swiss Alps, with no interceptors, relatively good coverage for other sensors, discontinuous low clouds and daylight
- Sensors return:
  - Height = very low / very low / very low / low
  - Type = helicopter / helicopter
  - Flight Path = U-path / U-path / U-path / U-path / U-path / missing
  - Height Changes = descent / descent / descent / descent / missing
  - Speed = slow / slow / slow / slow / slow
  - ADDC reaction = positive / positive / positive / positive / positive / positive
- We reject renegade and damaged, but indecision between provocateur and erroneous
- Assuming higher levels of reliability
- The aircraft is a provocateur!



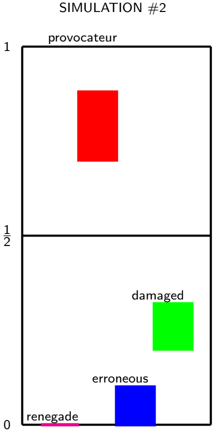
Simulations

- Simulating a dam in the Swiss Alps, with no interceptors, relatively good coverage for other sensors, discontinuous low clouds and daylight
- Sensors return:
  - Height = very low / very low / very low / low
  - Type = helicopter / helicopter
  - Flight Path = U-path / U-path / U-path / U-path / U-path / missing
  - Height Changes = descent / descent / descent / descent / missing
  - Speed = slow / slow / slow / slow / slow
  - ADDC reaction = positive / positive / positive / positive / positive / positive
- We reject renegade and damaged, but indecision between provocateur and erroneous
- Assuming higher levels of reliability
- The aircraft is a provocateur!



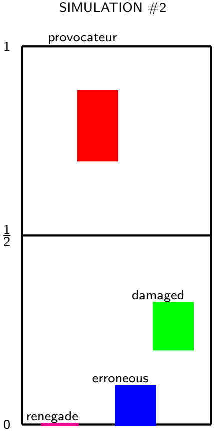
Simulations

- Simulating a dam in the Swiss Alps, with no interceptors, relatively good coverage for other sensors, discontinuous low clouds and daylight
- Sensors return:
  - Height = very low / very low / very low / low
  - Type = helicopter / helicopter
  - Flight Path = U-path / U-path / U-path / U-path / U-path / missing
  - Height Changes = descent / descent / descent / descent / missing
  - Speed = slow / slow / slow / slow / slow
  - ADDC reaction = positive / positive / positive / positive / positive / positive
- We reject renegade and damaged, but indecision between provocateur and erroneous
- Assuming higher levels of reliability
- The aircraft is a provocateur!

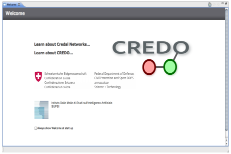


Simulations

- Simulating a dam in the Swiss Alps, with no interceptors, relatively good coverage for other sensors, discontinuous low clouds and daylight
- Sensors return:
  - Height = very low / very low / very low / low
  - Type = helicopter / helicopter
  - Flight Path = U-path / U-path / U-path / U-path / U-path / missing
  - Height Changes = descent / descent / descent / descent / missing
  - Speed = slow / slow / slow / slow / slow
  - ADDC reaction = positive / positive / positive / positive / positive / positive
- We reject renegade and damaged, but indecision between provocateur and erroneous
- Assuming higher levels of reliability
- The aircraft is a provocateur!



The CREDO software



Learn about Credal Networks...

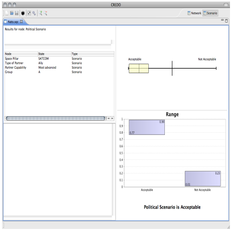
Learn about CREDO...

Schweizerische Eidgenossenschaft
Confédération suisse
Confederazione Svizzera
Confederaziun svizra

Federal Department of Defense,
Civil Protection and Sport DDPS
armasuisse
Science + Technology

Istituto Dalle Molle di Studi sull'Intelligenza Artificiale
SUPSI

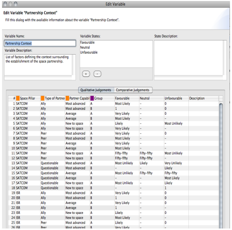
☐ Always show Welcome at start up



Results for node: Political Scenario

Node	State	Type
Space Pillar	SATCOM	Scenario
Type of Partner	Ally	Scenario
Partner Capability	Most advanced	Scenario
Group	A	Scenario

Political Scenario is Acceptable



Edit Variable "Partnership Context"

Fill this dialog with the available information about the variable "Partnership Context".

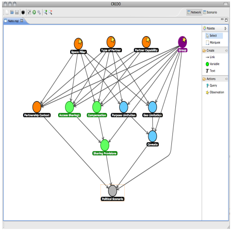
Variable Name: Partnership Context

Variable Description: List of factors defining the context surrounding the establishment of the space partnership.

Variable States:
Favourable
Neutral
Unfavourable

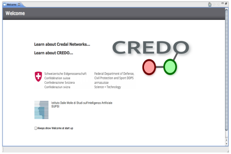
State Description:

#	Space Pillar	Type of Partner	Partner Capabil	Group	Favourable	Neutral	Unfavourable	Description
1	SATCOM	Ally	Most advanced	A	Most Likely	-	0	
2	SATCOM	Ally	Most advanced	B	1	-	0	
3	SATCOM	Ally	Average	A	Very Likely	-	0	
4	SATCOM	Ally	Average	B	Most Likely	-	-	
5	SATCOM	Ally	New to space	A	Likely	-	Most Unlikely	
6	SATCOM	Ally	New to space	B	Very Likely	-	-	
7	SATCOM	Peer	Most advanced	A	Very Likely	-	0	
8	SATCOM	Peer	Most advanced	B	Most Likely	-	-	
9	SATCOM	Peer	Average	A	Very Likely	-	-	
10	SATCOM	Peer	Average	B	Most Likely	-	-	
11	SATCOM	Peer	New to space	A	Fifty-fifty	Fifty-fifty	Most Unlikely	
12	SATCOM	Peer	New to space	B	Fifty-fifty	Fifty-fifty	-	
13	SATCOM	Questionable	Most advanced	A	Most Unlikely	Likely	Very Unlikely	
14	SATCOM	Questionable	Most advanced	B	-	-	Likely	
15	SATCOM	Questionable	Average	A	Most Unlikely	Fifty-fifty	Fifty-fifty	
16	SATCOM	Questionable	Average	B	-	-	Most Likely	
17	SATCOM	Questionable	New to space	A	Most Unlikely	-	Likely	
18	SATCOM	Questionable	New to space	B	-	-	1	
19	ISR	Ally	Most advanced	A	Very Likely	-	0	
20	ISR	Ally	Most advanced	B	1	-	0	
21	ISR	Ally	Average	A	Very Likely	-	0	
22	ISR	Ally	Average	B	1	-	-	
23	ISR	Ally	New to space	A	Likely	-	0	
24	ISR	Ally	New to space	B	Most Likely	-	-	
25	ISR	Peer	Most advanced	A	Likely	-	0	
26	ISR	Peer	Most advanced	B	Very Likely	-	-	



- A GUI software for CNs developed by IDSIA for Armasuisse
- Designed for military decision making
  but an academic version to be released by the end of 2013

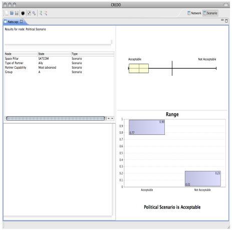
The CREDO software



Learn about Credal Networks...

Learn about CREDO...

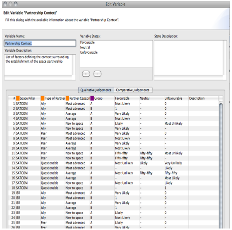
☐ Always show Welcome at start up



Results for node: Political Scenario

Node	State	Type
Space Pillar	SATCOM	Scenario
Type of Partner	Ally	Scenario
Partner Capability	Most advanced	Scenario
Group	A	Scenario

Political Scenario is Acceptable



Edit Variable "Partnership Context"

Fill this dialog with the available information about the variable "Partnership Context".

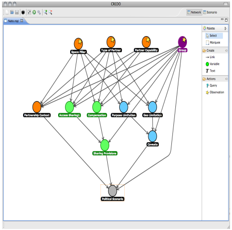
Variable Name: Partnership Context

Variable Description: List of factors defining the context surrounding the establishment of the space partnership.

Variable States:
Favourable
Neutral
Unfavourable

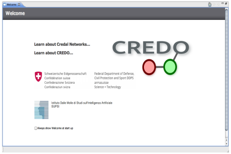
State Description:

#	Space Pillar	Type of Partner	Partner Capabil	Group	Favourable	Neutral	Unfavourable	Description
1	SATCOM	Ally	Most advanced	A	Most Likely	-	0	
2	SATCOM	Ally	Most advanced	B	1	-	0	
3	SATCOM	Ally	Average	A	Very Likely	-	0	
4	SATCOM	Ally	Average	B	Most Likely	-	-	
5	SATCOM	Ally	New to space	A	Likely	-	Most Unlikely	
6	SATCOM	Ally	New to space	B	Very Likely	-	-	
7	SATCOM	Peer	Most advanced	A	Very Likely	-	0	
8	SATCOM	Peer	Most advanced	B	Most Likely	-	-	
9	SATCOM	Peer	Average	A	Very Likely	-	-	
10	SATCOM	Peer	Average	B	Most Likely	-	-	
11	SATCOM	Peer	New to space	A	Fifty-fifty	Fifty-fifty	Most Unlikely	
12	SATCOM	Peer	New to space	B	Fifty-fifty	Fifty-fifty	-	
13	SATCOM	Questionable	Most advanced	A	Most Unlikely	Likely	Very Unlikely	
14	SATCOM	Questionable	Most advanced	B	-	-	Likely	
15	SATCOM	Questionable	Average	A	Most Unlikely	Fifty-fifty	Fifty-fifty	
16	SATCOM	Questionable	Average	B	-	-	Most Likely	
17	SATCOM	Questionable	New to space	A	Most Unlikely	-	Likely	
18	SATCOM	Questionable	New to space	B	-	-	1	
19	ISR	Ally	Most advanced	A	Very Likely	-	0	
20	ISR	Ally	Most advanced	B	1	-	0	
21	ISR	Ally	Average	A	Very Likely	-	0	
22	ISR	Ally	Average	B	1	-	-	
23	ISR	Ally	New to space	A	Likely	-	0	
24	ISR	Ally	New to space	B	Most Likely	-	-	
25	ISR	Peer	Most advanced	A	Likely	-	0	
26	ISR	Peer	Most advanced	B	Very Likely	-	-	



- A GUI software for CNs developed by IDSIA for Armasuisse
- Designed for military decision making
  but an academic version to be released by the end of 2013

The CREDO software



Welcome

Learn about Credal Networks...

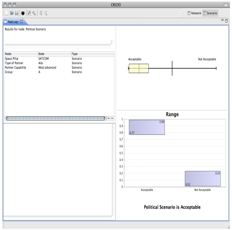
Learn about CREDO...

Schweizerische Eidgenossenschaft
Confédération suisse
Confederazione Svizzera
Confederaziun svizra

Federal Department of Defense,
Civil Protection and Sport DDPS
armasuisse
Science + Technology

Istituto Dalle Molle di Studi sull'Intelligenza Artificiale
SUPSI

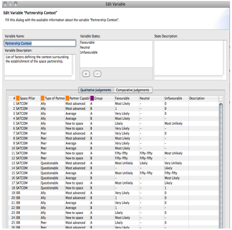
☐ Always show Welcome at start up



Results for node: Political Scenario

Node	State	Type
Space Pillar	SATCOM	Scenario
Type of Partner	Ally	Scenario
Partner Capability	Most advanced	Scenario
Group	A	Scenario

Political Scenario is Acceptable



Edit Variable "Partnership Context"

Fill this dialog with the available information about the variable "Partnership Context".

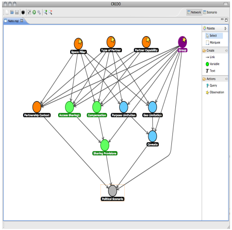
Variable Name: Partnership Context

Variable Description: List of factors defining the context surrounding the establishment of the space partnership.

Variable States:
Favourable
Neutral
Unfavourable

State Description:

#	Space Pillar	Type of Partner	Partner Capabil	Group	Favourable	Neutral	Unfavourable	Description
1	SATCOM	Ally	Most advanced	A	Most Likely	-	0	
2	SATCOM	Ally	Most advanced	B	1	-	0	
3	SATCOM	Ally	Average	A	Very Likely	-	0	
4	SATCOM	Ally	Average	B	Most Likely	-	-	
5	SATCOM	Ally	New to space	A	Likely	-	Most Unlikely	
6	SATCOM	Ally	New to space	B	Very Likely	-	-	
7	SATCOM	Peer	Most advanced	A	Very Likely	-	0	
8	SATCOM	Peer	Most advanced	B	Most Likely	-	-	
9	SATCOM	Peer	Average	A	Very Likely	-	-	
10	SATCOM	Peer	Average	B	Most Likely	-	-	
11	SATCOM	Peer	New to space	A	Fifty-fifty	Fifty-fifty	Most Unlikely	
12	SATCOM	Peer	New to space	B	Fifty-fifty	Fifty-fifty	-	
13	SATCOM	Questionable	Most advanced	A	Most Unlikely	Likely	Very Unlikely	
14	SATCOM	Questionable	Most advanced	B	-	-	Likely	
15	SATCOM	Questionable	Average	A	Most Unlikely	Fifty-fifty	Fifty-fifty	
16	SATCOM	Questionable	Average	B	-	-	Most Likely	
17	SATCOM	Questionable	New to space	A	Most Unlikely	-	Likely	
18	SATCOM	Questionable	New to space	B	-	-	1	
19	ISR	Ally	Most advanced	A	Very Likely	-	0	
20	ISR	Ally	Most advanced	B	1	-	0	
21	ISR	Ally	Average	A	Very Likely	-	0	
22	ISR	Ally	Average	B	1	-	-	
23	ISR	Ally	New to space	A	Likely	-	0	
24	ISR	Ally	New to space	B	Most Likely	-	-	
25	ISR	Peer	Most advanced	A	Likely	-	0	
26	ISR	Peer	Most advanced	B	Very Likely	-	-	



- A GUI software for CNs developed by IDSIA for Armasuisse
- Designed for military decision making
  but an academic version to be released by the end of 2013

Decision-Support System for Space Security

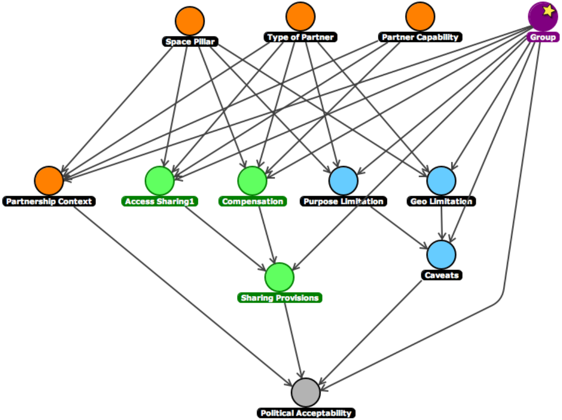
Variable of interest: *Political acceptability* (*acceptable* / *unacceptable*)

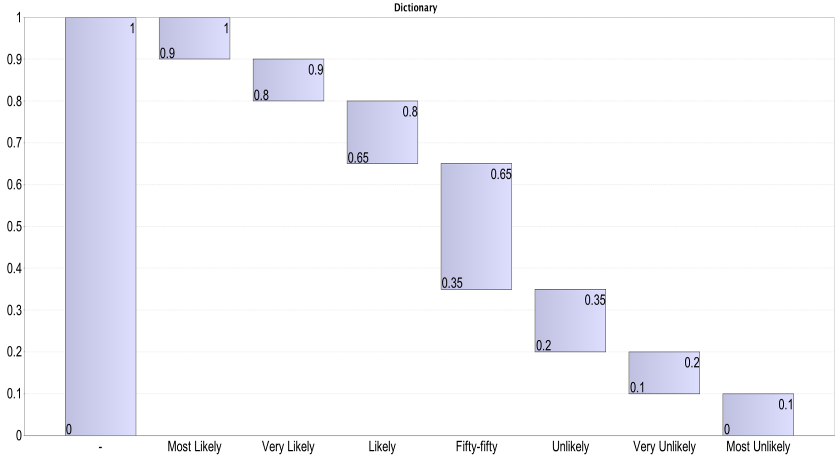
Observed features

- *Space pillar*: possible states *SATCOM* (command, control, communication and computer systems dependent on satellites communication), *ISR* (synchronized and integrated planning) and *SSA* (ability to obtain information and knowledge about the space beyond the Earth atmosphere).
- *Type of partner*: *ally*, *peer* and *questionable*.
- *Partner capability*: *most advanced*, *average* and *new to space*.

Intermediate variables

- *Access sharing*: *C2 payload and raw data* (ability to directly manage the beam), *raw data only* and *no direct access*.
- *Compensation*: *in-kind*, *small*, *medium* and *large compensation*.
- *Purpose limitation*: *peaceful and non-economic*, *peaceful only* and *no limitations*.
- *Geographical limitation*: *peace-keeping exclusion*, *partner exclusion* and *no limitations*.


Space Pillar
Type of Partner
Partner Capability
Group
Partnership Context
Access Sharing1
Compensation
Purpose Limitation
Geo Limitation
Caveats
Sharing Provisions
Political Acceptability


Dictionary
1
0.9
0.8
0.7
0.6
0.5
0.4
0.3
0.2
0.1
0
-
0
1
Most Likely
0.9
1
Very Likely
0.8
0.9
Likely
0.65
0.8
Fifty-fifty
0.35
0.65
Unlikely
0.2
0.35
Very Unlikely
0.1
0.2
Most Unlikely
0
0.1

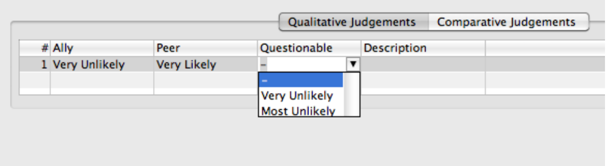
Preventing inconsistent judgements

- E.g., two states of X cannot be both "likely" (as this means $P(x) > .65$, $\sum_x P(x) > 1$).
- Reachability constraints

$$\sum_{x \in \Omega_X \setminus \{x'\}} \underline{P}(x) + \overline{P}(x') \quad \leq 1, \tag{1}$$

$$\sum_{x \in \Omega_X \setminus \{x'\}} \overline{P}(x) + \underline{P}(x') \quad \geq 1. \tag{2}$$

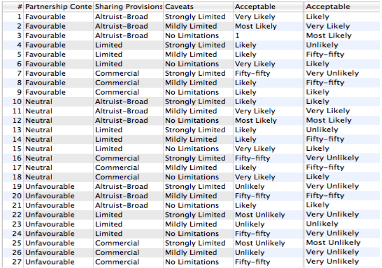
- Judgement specification is sequential, the software displays only consistent options



#	Ally	Peer	Questionable	Description
1	Very Unlikely	Very Likely	-	

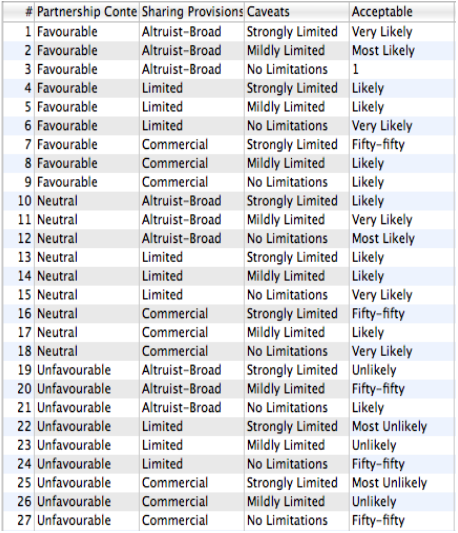
NATO Multinational Experiment 7

- Concerned with protecting our access to the global commons.
- During the final meeting a group of six subject matter experts (divided into two groups) developed its own conclusions about political acceptability for 27 scenarios
- Human experts reasoning vs. (almost) automatic reasoning with credal networks (quantified by expert knowledge)



#	Partnership Conte	Sharing Provisions	Caveats	Acceptable	Acceptable
1	Favourable	Altruist-Broad	Strongly Limited	Very Likely	Likely
2	Favourable	Altruist-Broad	Mildly Limited	Most Likely	Very Likely
3	Favourable	Altruist-Broad	No Limitations	1	Most Likely
4	Favourable	Limited	Strongly Limited	Likely	Unlikely
5	Favourable	Limited	Mildly Limited	Likely	Fifty-fifty
6	Favourable	Limited	No Limitations	Very Likely	Likely
7	Favourable	Commercial	Strongly Limited	Fifty-fifty	Very Unlikely
8	Favourable	Commercial	Mildly Limited	Likely	Fifty-fifty
9	Favourable	Commercial	No Limitations	Likely	Likely
10	Neutral	Altruist-Broad	Strongly Limited	Likely	Likely
11	Neutral	Altruist-Broad	Mildly Limited	Very Likely	Very Likely
12	Neutral	Altruist-Broad	No Limitations	Most Likely	Most Likely
13	Neutral	Limited	Strongly Limited	Likely	Unlikely
14	Neutral	Limited	Mildly Limited	Likely	Fifty-fifty
15	Neutral	Limited	No Limitations	Very Likely	Likely
16	Neutral	Commercial	Strongly Limited	Fifty-fifty	Very Unlikely
17	Neutral	Commercial	Mildly Limited	Likely	Fifty-fifty
18	Neutral	Commercial	No Limitations	Very Likely	Likely
19	Unfavourable	Altruist-Broad	Strongly Limited	Unlikely	Very Unlikely
20	Unfavourable	Altruist-Broad	Mildly Limited	Fifty-fifty	Fifty-fifty
21	Unfavourable	Altruist-Broad	No Limitations	Likely	Likely
22	Unfavourable	Limited	Strongly Limited	Most Unlikely	Very Unlikely
23	Unfavourable	Limited	Mildly Limited	Unlikely	Unlikely
24	Unfavourable	Limited	No Limitations	Fifty-fifty	Very Unlikely
25	Unfavourable	Commercial	Strongly Limited	Most Unlikely	Most Unlikely
26	Unfavourable	Commercial	Mildly Limited	Unlikely	Very Unlikely
27	Unfavourable	Commercial	No Limitations	Fifty-fifty	Very Unlikely

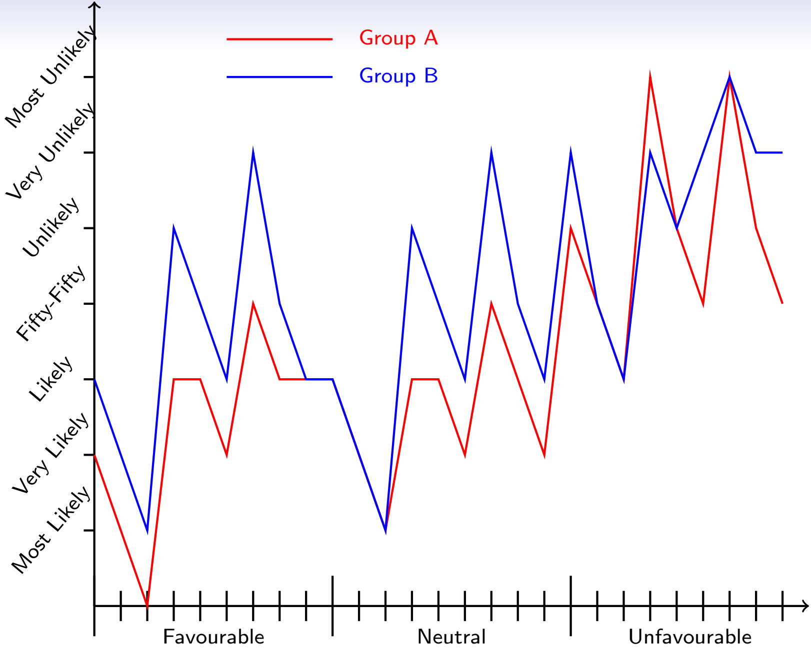
Group A



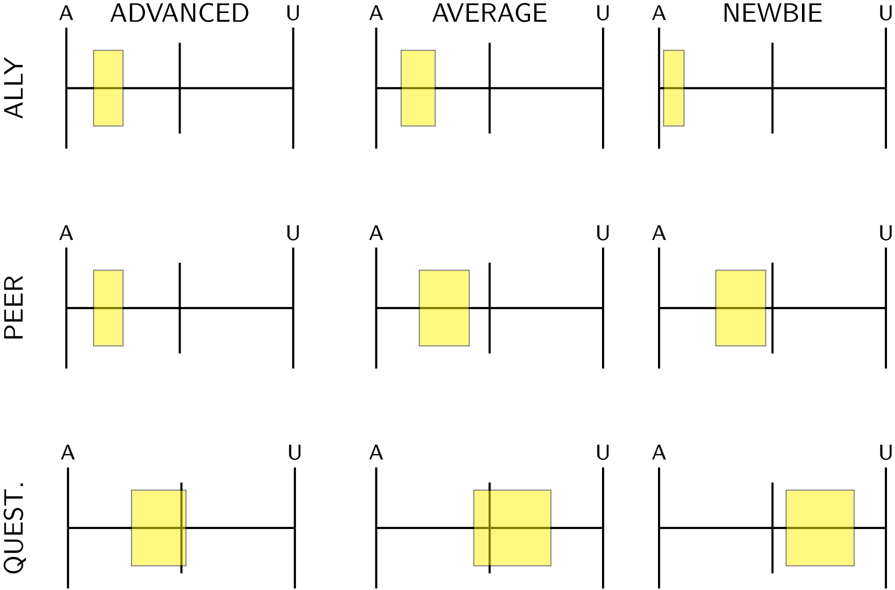
#	Partnership Conte	Sharing Provisions	Caveats	Acceptable
1	Favourable	Altruist-Broad	Strongly Limited	Very Likely
2	Favourable	Altruist-Broad	Mildly Limited	Most Likely
3	Favourable	Altruist-Broad	No Limitations	1
4	Favourable	Limited	Strongly Limited	Likely
5	Favourable	Limited	Mildly Limited	Likely
6	Favourable	Limited	No Limitations	Very Likely
7	Favourable	Commercial	Strongly Limited	Fifty-fifty
8	Favourable	Commercial	Mildly Limited	Likely
9	Favourable	Commercial	No Limitations	Likely
10	Neutral	Altruist-Broad	Strongly Limited	Likely
11	Neutral	Altruist-Broad	Mildly Limited	Very Likely
12	Neutral	Altruist-Broad	No Limitations	Most Likely
13	Neutral	Limited	Strongly Limited	Likely
14	Neutral	Limited	Mildly Limited	Likely
15	Neutral	Limited	No Limitations	Very Likely
16	Neutral	Commercial	Strongly Limited	Fifty-fifty
17	Neutral	Commercial	Mildly Limited	Likely
18	Neutral	Commercial	No Limitations	Very Likely
19	Unfavourable	Altruist-Broad	Strongly Limited	Unlikely
20	Unfavourable	Altruist-Broad	Mildly Limited	Fifty-fifty
21	Unfavourable	Altruist-Broad	No Limitations	Likely
22	Unfavourable	Limited	Strongly Limited	Most Unlikely
23	Unfavourable	Limited	Mildly Limited	Unlikely
24	Unfavourable	Limited	No Limitations	Fifty-fifty
25	Unfavourable	Commercial	Strongly Limited	Most Unlikely
26	Unfavourable	Commercial	Mildly Limited	Unlikely
27	Unfavourable	Commercial	No Limitations	Fifty-fifty

Group B

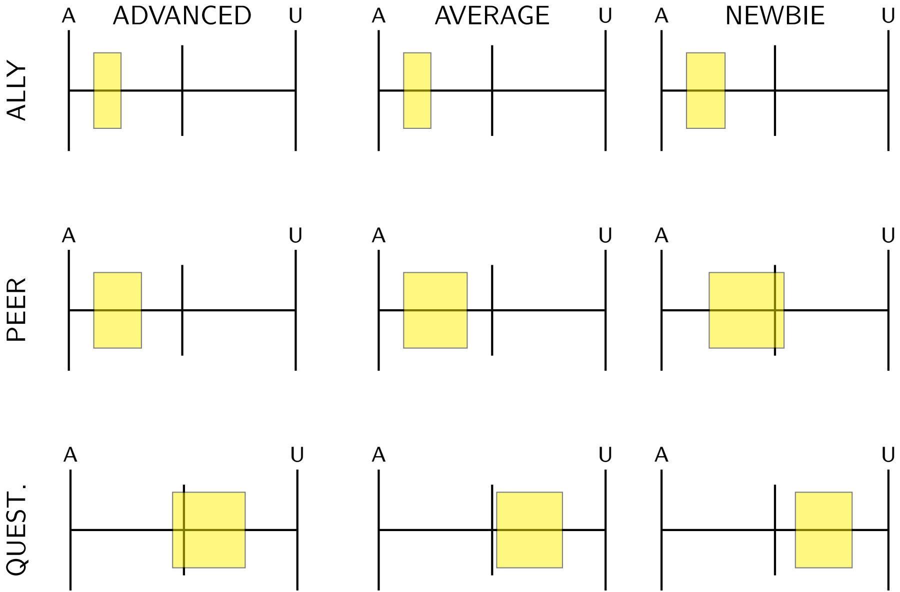
Acceptable
Likely
Very Likely
Most Likely
Unlikely
Fifty-fifty
Likely
Very Unlikely
Fifty-fifty
Likely
Likely
Very Likely
Most Likely
Unlikely
Fifty-fifty
Likely
Very Unlikely
Fifty-fifty
Likely
Very Unlikely
Fifty-fifty
Likely
Very Unlikely
Unlikely
Very Unlikely
Most Unlikely
Very Unlikely
Very Unlikely


Group A
Group B
Most Unlikely
Very Unlikely
Unlikely
Fifty-Fifty
Likely
Very Likely
Most Likely
Favourable
Neutral
Unfavourable

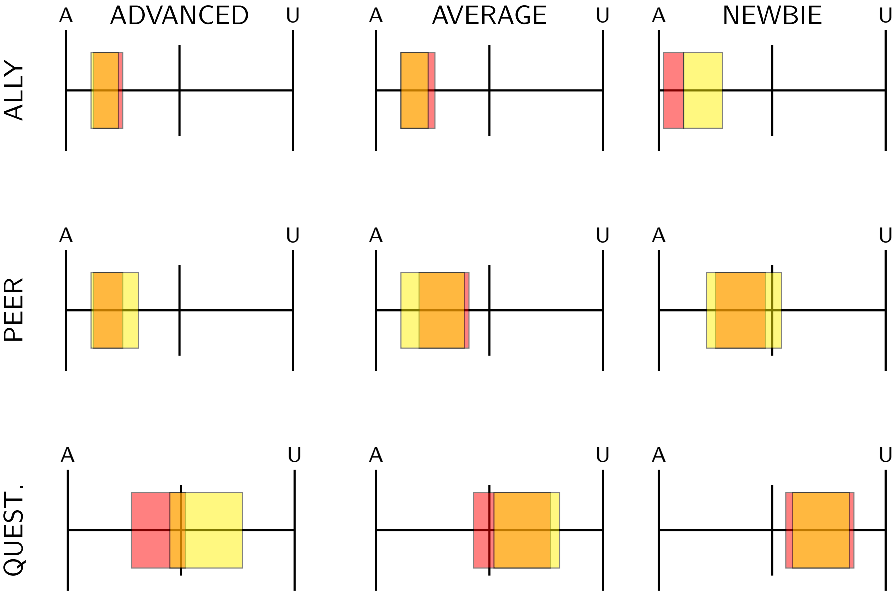
ISR (Group A)



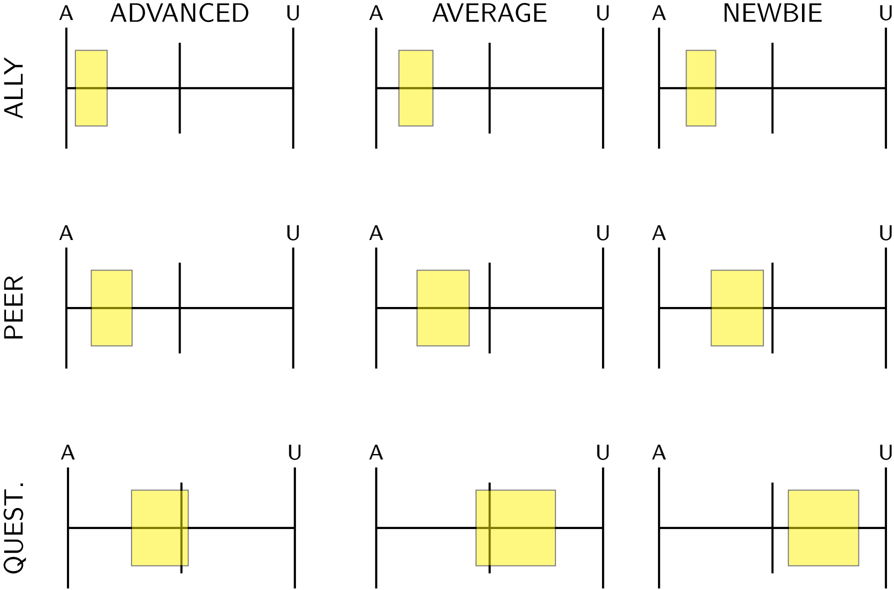
ISR (Group B)


A
ADVANCED
U
A
AVERAGE
U
A
NEWBIE
U
ALLY
PEER
A
U
A
U
A
U
QUEST.
A
U
A
U
A
U

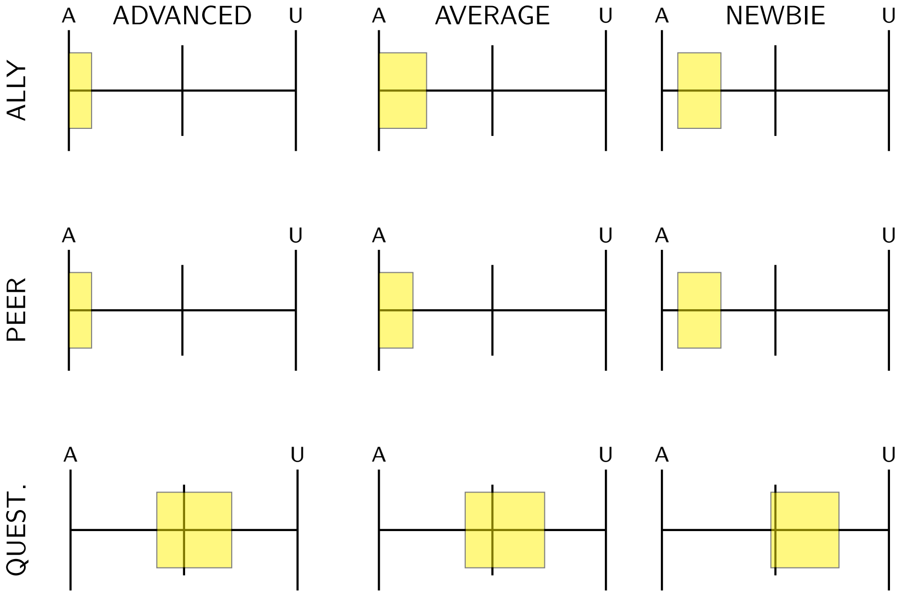
ISR (Group A+B)



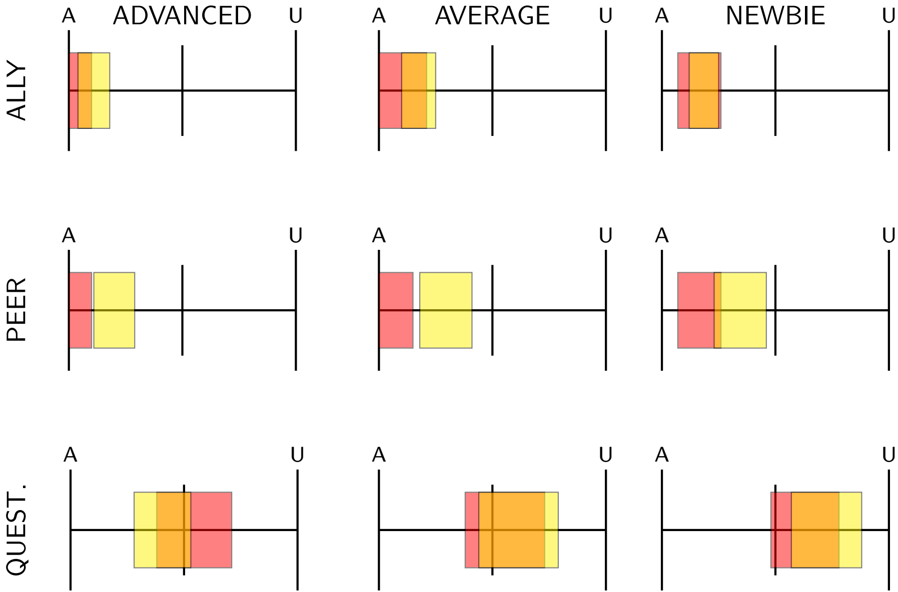
SSA (Group A)



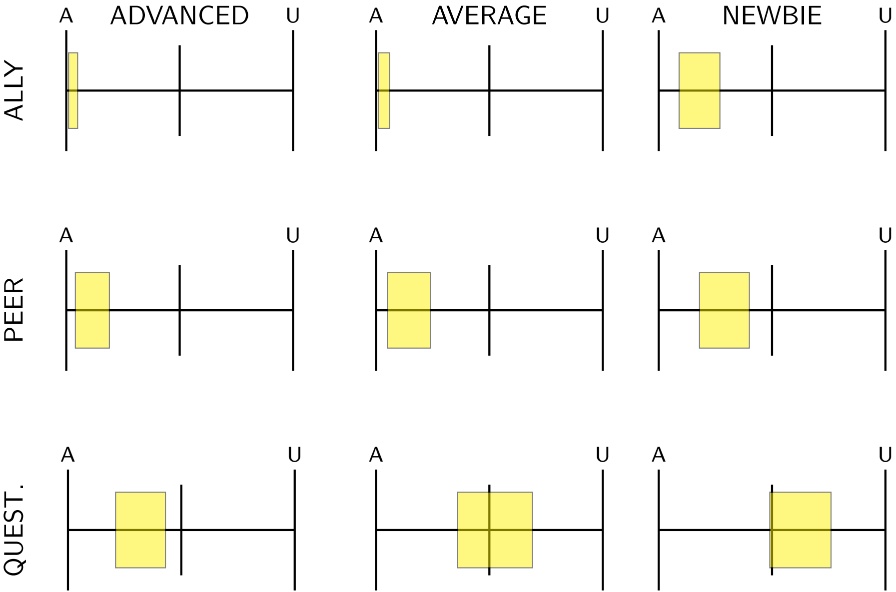
SSA (Group B)



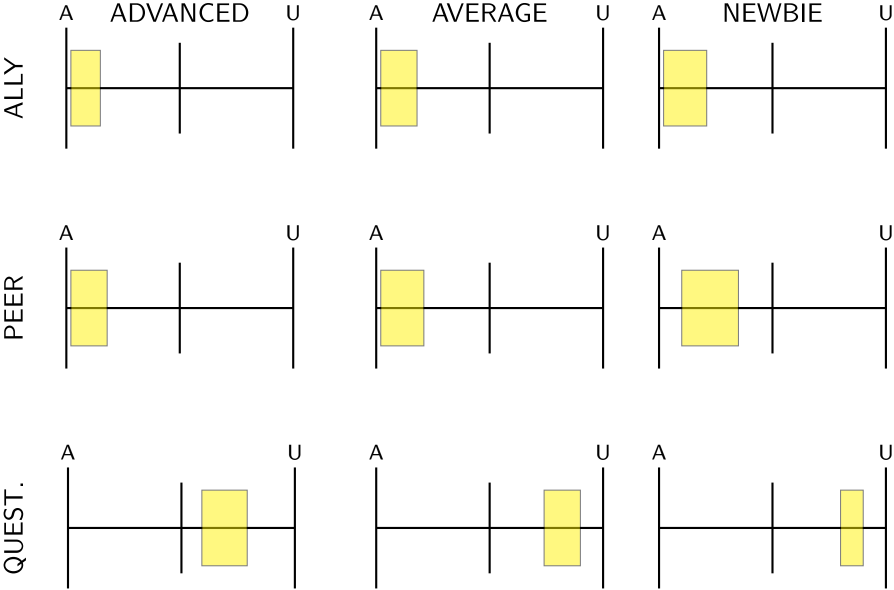
SSA (Group A+B)


A
ADVANCED
U
A
AVERAGE
U
A
NEWBIE
U
ALLY
A
U
A
U
A
U
PEER
A
U
A
U
A
U
QUEST.

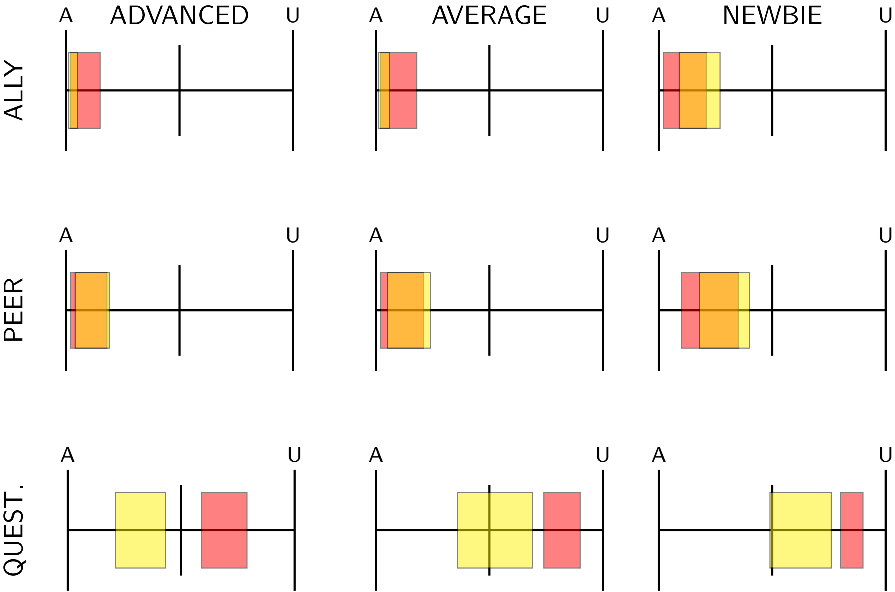
SATCOM (Group A)



SATCOM (Group B)



SATCOM (Group A+B)



Experiment conclusions

- 27 vignettes
- A and B agree on a single answer 18
- A and B agree on suspending judgement 3
- A suspend , B not or vice versa 5
- A and B disagree 1
- Good agreement, not-too-imprecise outputs, results consistent with human conclusions

Environmental example: debris flows hazard assessment

debrisflows

df

- Debris flows are very destructive natural hazards
- Still partially understood
- Human expertize is still fundamental!
- An artificial expert system supporting human experts?

Environmental example: debris flows hazard assessment

debrisflows

df

- Debris flows are very destructive natural hazards
- Still partially understood
- Human expertize is still fundamental!
- An artificial expert system supporting human experts?

Environmental example: debris flows hazard assessment

debrisflows

df

- Debris flows are very destructive natural hazards
- Still partially understood
- Human expertize is still fundamental!
- An artificial expert system supporting human experts?

Environmental example: debris flows hazard assessment

debrisflows

df

- Debris flows are very destructive natural hazards
- Still partially understood
- Human expertize is still fundamental!
- An artificial expert system supporting human experts?

Environmental example: debris flows hazard assessment

debrisflows

df

- Debris flows are very destructive natural hazards
- Still partially understood
- Human expertize is still fundamental!
- An artificial expert system supporting human experts?

Causal modelling

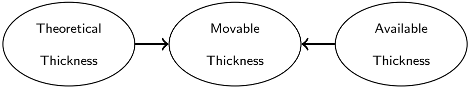
Proxy indicator of the level of risk

Movable

Thickness

Causal modelling

Triggering Factors

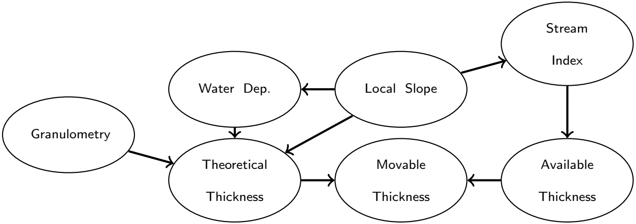


Theoretical Thickness

Movable Thickness

Available Thickness

Causal modelling



Stream Index

Water Dep.

Local Slope

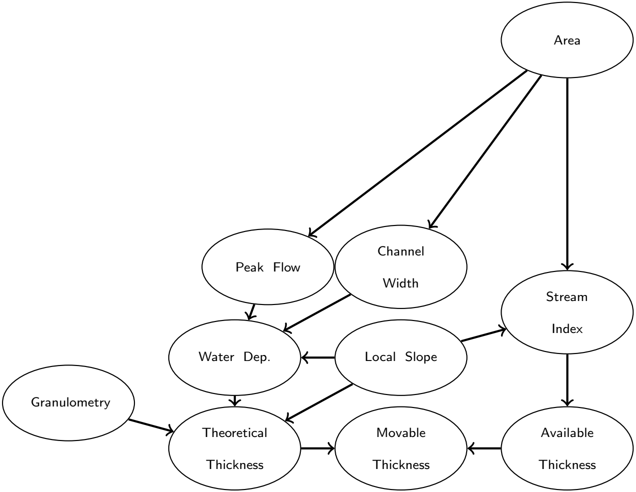
Granulometry

Theoretical Thickness

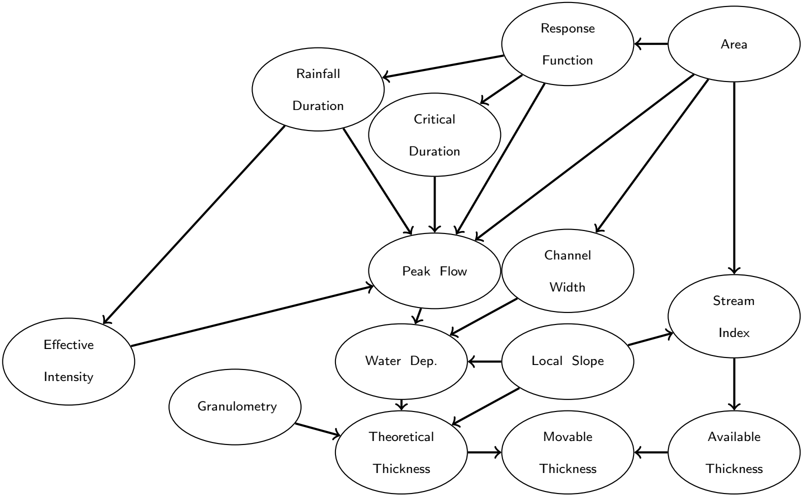
Movable Thickness

Available Thickness

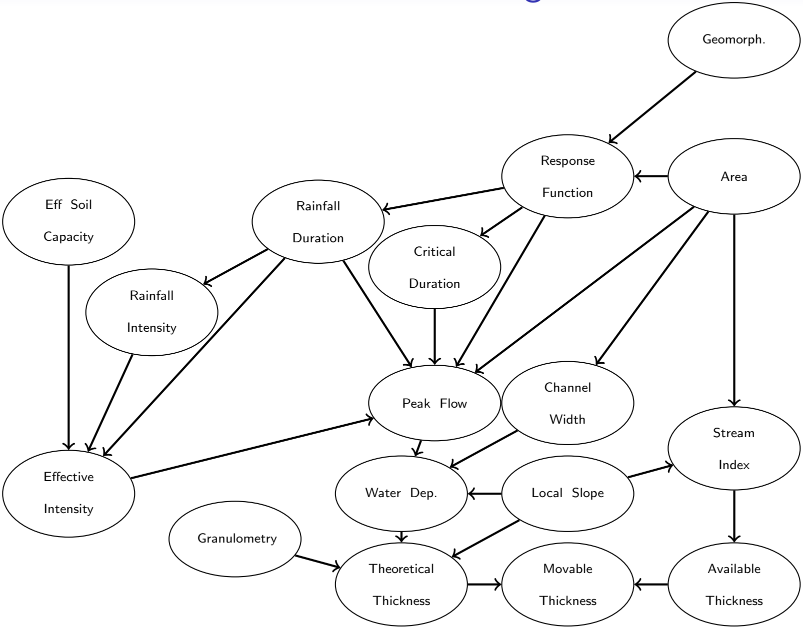
Causal modelling



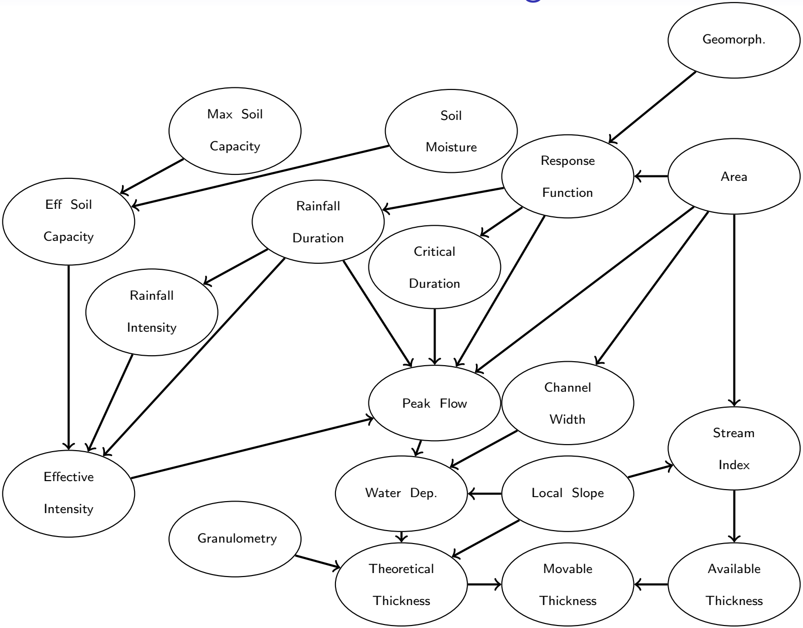
Causal modelling



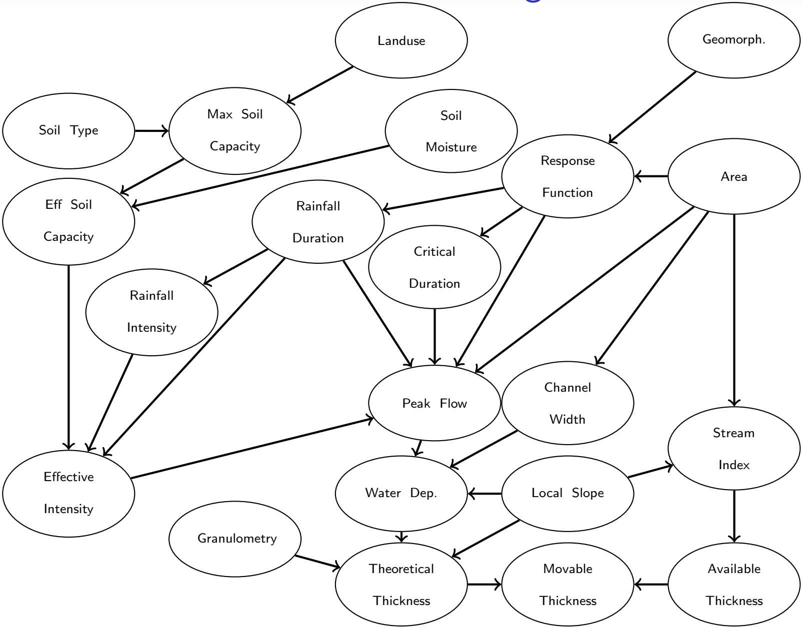
Causal modelling



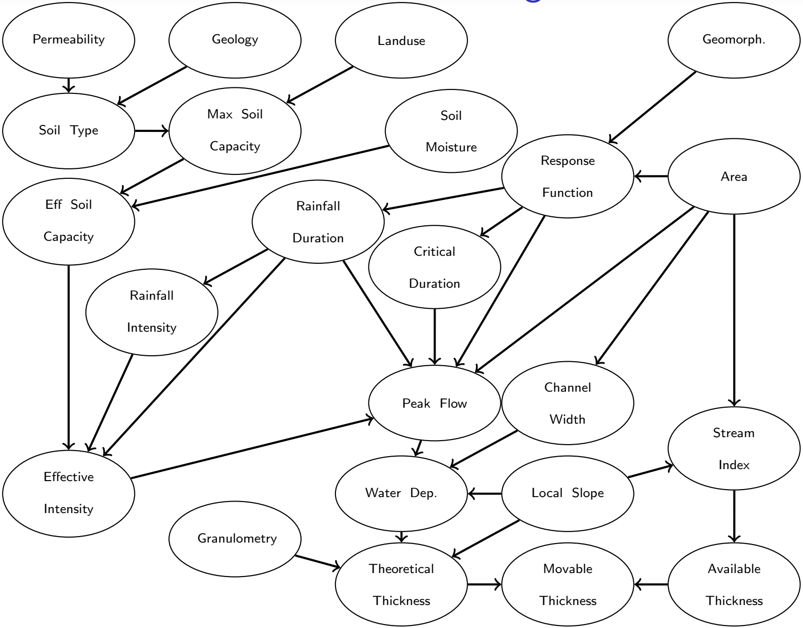
Causal modelling



Causal modelling

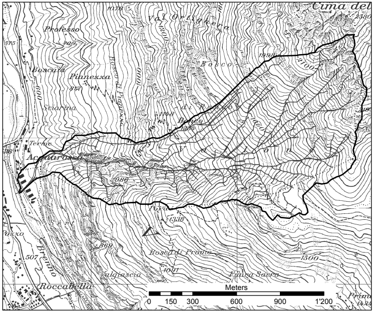


Causal modelling



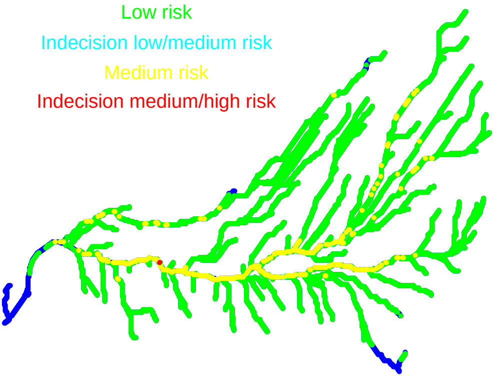
Debris flow hazard assessment by CNs

- Extensive simulations in a debris flow prone watershed Acquarossa Creek Basin (area 1.6 Km^2, length 3.1 Km)



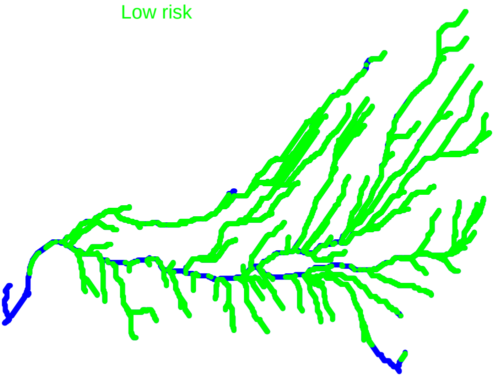
Debris flow hazard assessment by CNs

- Extensive simulations in a debris flow prone watershed Acquarossa Creek Basin (area 1.6 Km^2, length 3.1 Km)



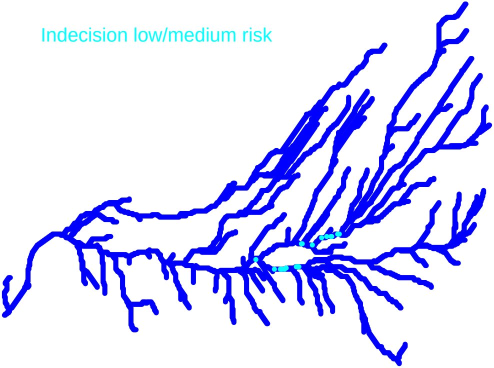
Debris flow hazard assessment by CNs

- Extensive simulations in a debris flow prone watershed Acquarossa Creek Basin (area 1.6 Km^2, length 3.1 Km)



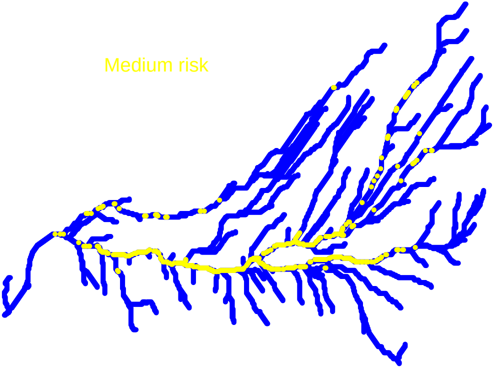
Debris flow hazard assessment by CNs

- Extensive simulations in a debris flow prone watershed Acquarossa Creek Basin (area 1.6 Km^2, length 3.1 Km)



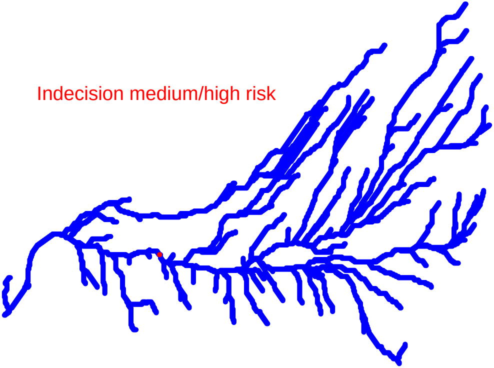
Debris flow hazard assessment by CNs

- Extensive simulations in a debris flow prone watershed Acquarossa Creek Basin (area 1.6 Km^2, length 3.1 Km)



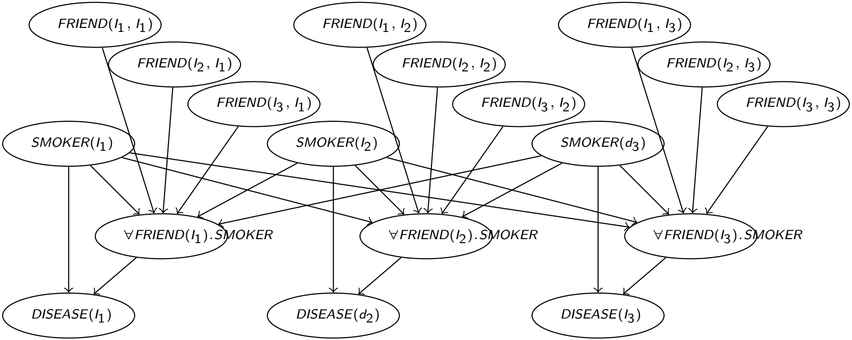
Debris flow hazard assessment by CNs

- Extensive simulations in a debris flow prone watershed Acquarossa Creek Basin (area 1.6 Km^2, length 3.1 Km)



$\mathcal{CRALC}$ probabilistic logic with IPs *(Cozman, 2008)*

- Description logic with interval of probabilities
- N individuals $(I_1, \ldots, I_n)$,
  $P(smoker(I_i)) \in [.3, .5],\ P(friend(I_j, I_i)) \in [.0, .5]$,
  $P(disease(I_i)|smoker(I_i), \forall friend(I_j, I_i).I_i smoker) = \ldots$
- $\underline{P}(disease)$? Inference $\equiv$ updating of a (large) binary CN



References

- Piatti, A., Antonucci, A., Zaffalon, M. (2010). Building knowledge-based expert systems by credal networks: a tutorial. In Baswell, A.R. (Ed), Advances in Mathematics Research 11, Nova Science Publishers, New York.
- Corani, G., Antonucci, A., Zaffalon, M. (2012). Bayesian networks with imprecise probabilities: theory and application to classification. In Holmes, D.E., Jain, L.C. (Eds), Data Mining: Foundations and Intelligent Paradigms, Intelligent Systems Reference Library 23, Springer, Berlin / Heidelberg, pp. 49–93.
- `ipg.idsia.ch`
- `www.sipta.org`