# IMPRECISE IMMEDIATE PREDICTIONS

## GETTING IP TO WORK, AND FAST!

Gert de Cooman
Erik Quaeghebeur
Filip Hermans

SYSTeMS Research Group
Ghent University
gert.decooman@UGent.be
http://users.UGent.be/~gdcooma

SIPTA School 2008, Montpellier
7 July 2008

# Precise probability models

## Mass functions and expectations

Assume we are uncertain about:

- the value or a variable $X$
- in a set of possible values $\mathscr{X}$.

This is usually modelled by a probability mass function $p$ on $\mathscr{X}$:

$$p(x) \geq 0 \text{ and } \sum_{x \in \mathscr{X}} p(x) = 1;$$

With $p$ we can associate an expectation operator $E_p$:

$$E_p(f) := \sum_{x \in \mathscr{X}} p(x) f(x) \text{ where } f \colon \mathscr{X} \to \mathbb{R}.$$

If $A \subseteq \mathscr{X}$ is an event, then its probability is given by

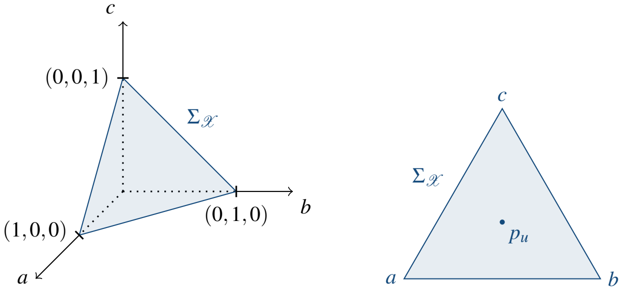
$$P_p(A) = \sum_{x \in A} p(x) = E_p(I_A).$$

# Precise probability models

## The simplex of all probability mass functions

Consider the simplex $\Sigma_{\mathscr{X}}$ of all mass functions on $\mathscr{X}$:

$$\Sigma_{\mathscr{X}} := \left\{ p \in \mathbb{R}_{+}^{\mathscr{X}} : \sum_{x \in \mathscr{X}} p(x) = 1 \right\}.$$



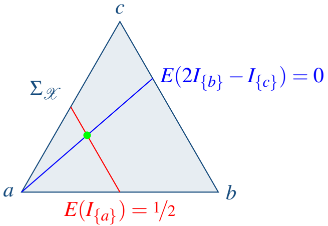
## ASSESSMENTS LEAD TO CONSTRAINTS

Specifying an expectation $E(f)$ for a map $f\colon \mathscr{X} \to \mathbb{R}$

$$\sum_{x\in\mathscr{X}} p(x)f(x) = E(f)$$

imposes a linear constraint on the possible values for $p$ in $\Sigma_{\mathscr{X}}$.

It corresponds to intersecting the simplex $\Sigma_{\mathscr{X}}$ with a hyperplane, whose direction depends on $f$:



# Imprecise probability models

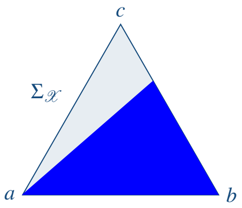
## Linear inequality constraints

### More flexible assessments

Impose linear inequality constraints on $p$ in $\Sigma_{\mathscr{X}}$:

$$\underline{E}(f) \leq \sum_{x \in \mathscr{X}} p(x) f(x) \quad \text{or} \quad \sum_{x \in \mathscr{X}} p(x) f(x) \leq \overline{E}(f).$$

Corresponds to intersecting $\Sigma_{\mathscr{X}}$ with affine semi-spaces:

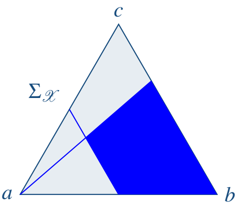


## MORE FLEXIBLE ASSESSMENTS

Impose linear inequality constraints on $p$ in $\Sigma_{\mathscr{X}}$:

$$\underline{E}(f) \leq \sum_{x \in \mathscr{X}} p(x) f(x) \quad \text{or} \quad \sum_{x \in \mathscr{X}} p(x) f(x) \leq \overline{E}(f).$$

Corresponds to intersecting $\Sigma_{\mathscr{X}}$ with affine semi-spaces:

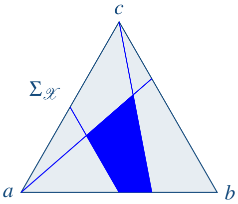


## MORE FLEXIBLE ASSESSMENTS

Impose linear inequality constraints on $p$ in $\Sigma_{\mathscr{X}}$:

$$\underline{E}(f) \leq \sum_{x\in\mathscr{X}} p(x)f(x) \quad \text{or} \quad \sum_{x\in\mathscr{X}} p(x)f(x) \leq \overline{E}(f).$$

Corresponds to intersecting $\Sigma_{\mathscr{X}}$ with affine semi-spaces:

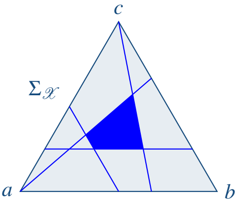


## More flexible assessments

Impose linear inequality constraints on $p$ in $\Sigma_{\mathscr{X}}$:

$$\underline{E}(f) \leq \sum_{x\in\mathscr{X}} p(x)f(x) \quad \text{or} \quad \sum_{x\in\mathscr{X}} p(x)f(x) \leq \overline{E}(f).$$

Corresponds to intersecting $\Sigma_{\mathscr{X}}$ with affine semi-spaces:

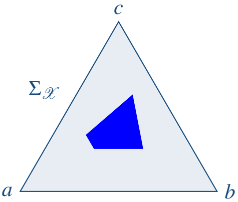


## MORE FLEXIBLE ASSESSMENTS

Impose linear inequality constraints on $p$ in $\Sigma_{\mathscr{X}}$:

$$\underline{E}(f) \leq \sum_{x\in\mathscr{X}} p(x)f(x) \quad \text{or} \quad \sum_{x\in\mathscr{X}} p(x)f(x) \leq \overline{E}(f).$$

Corresponds to intersecting $\Sigma_{\mathscr{X}}$ with affine semi-spaces:



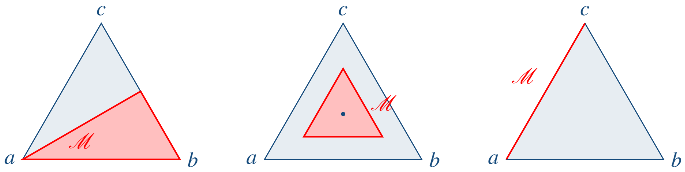
# Imprecise probability models

## Credal sets

Any such number of assessments leads to a credal set $\mathscr{M}$.

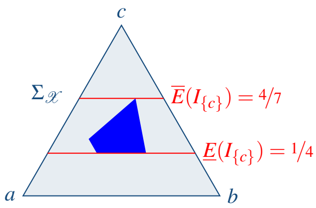
### Definition

A credal set $\mathscr{M}$ is a convex closed subset of $\Sigma_{\mathscr{X}}$.



# IMPRECISE PROBABILITY MODELS

## LOWER AND UPPER EXPECTATIONS




## EQUIVALENT MODEL

Consider the set $\mathscr{L}(\mathscr{X}) = \mathbb{R}^{\mathscr{X}}$ of all real-valued maps on $\mathscr{X}$.
We define two real functionals on $\mathscr{L}(\mathscr{X})$: for all $f\colon \mathscr{X} \to \mathbb{R}$

$$\underline{E}_{\mathscr{M}}(f) = \min\left\{E_p(f) : p \in \mathscr{M}\right\} \text{ lower expectation}$$
$$\overline{E}_{\mathscr{M}}(f) = \max\left\{E_p(f) : p \in \mathscr{M}\right\} \text{ upper expectation.}$$

Observe that [conjugacy]

$$\overline{E}_{\mathscr{M}}(f) = -\underline{E}_{\mathscr{M}}(-f).$$

## DEFINITION

We call a real functional $\overline{E}$ on $\mathscr{L}(\mathscr{X})$ an upper expectation if it satisfies the following properties:
For all $f$ and $g$ in $\mathscr{L}(\mathscr{X})$ and all real $\lambda \geq 0$:

1. $\overline{E}(f) \leq \max f$ [boundedness];
2. $\overline{E}(f+g) \leq \overline{E}(f)+\overline{E}(g)$ [sub-additivity];
3. $\overline{E}(\lambda f) = \lambda \overline{E}(f)$ [non-negative homogeneity].

## THEOREM (OTHER PROPERTIES)

*Let $\overline{E}$ be an upper expectation, with conjugate lower expectation $\underline{E}$. Then for all real numbers $\mu$ and all $f$ and $g$ in $\mathscr{L}(\mathscr{X})$:*

1. $\underline{E}(f) \leq \overline{E}(f)$;
2. $\underline{E}(f)+\underline{E}(g) \leq \underline{E}(f+g) \leq \underline{E}(f)+\overline{E}(g) \leq \overline{E}(f+g) \leq \overline{E}(f)+\overline{E}(g)$;
3. $\overline{E}(f+\mu) = \overline{E}(f)+\mu$;
4. $\overline{E}(|f|) \geq |\underline{E}(f)|$ *and* $\overline{E}(|f|) \geq |\overline{E}(f)|$.

# IMPRECISE PROBABILITY MODELS

## LOWER ENVELOPE THEOREM

### THEOREM (LOWER ENVELOPE THEOREM)

*A real functional $\overline{E}$ is an upper expectation if and only if it is the upper envelope of some credal set $\mathscr{M}$.*

### PROOF.

Use $\mathscr{M} = \{p \in \Sigma_{\mathscr{X}} : (\forall f \in \mathscr{L}(\mathscr{X}))(E_p(f) \leq \overline{E}(f))\}$. □

# DISCRETE-TIME UNCERTAIN PROCESSES

## PRECISE PROBABILITY TREES

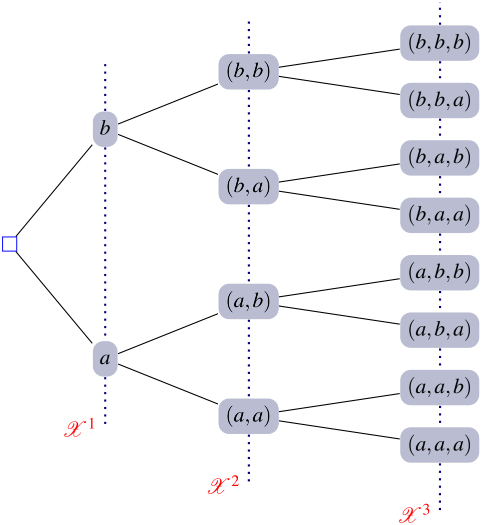
We consider an uncertain process with variables $X_1$, $X_2$, ..., $X_n$, ... assuming values in a finite set of states $\mathscr{X}$.

This leads to a standard event tree with nodes

$$s = (x_1, x_2, \dots, x_n), \quad x_k \in \mathscr{X}, \quad n \geq 0.$$

# Discrete-time uncertain processes

## Precise probability trees



# DISCRETE-TIME UNCERTAIN PROCESSES

## PRECISE PROBABILITY TREES

We consider an uncertain process with variables $X_1$, $X_2$, ..., $X_n$, ... assuming values in a finite set of states $\mathscr{X}$.

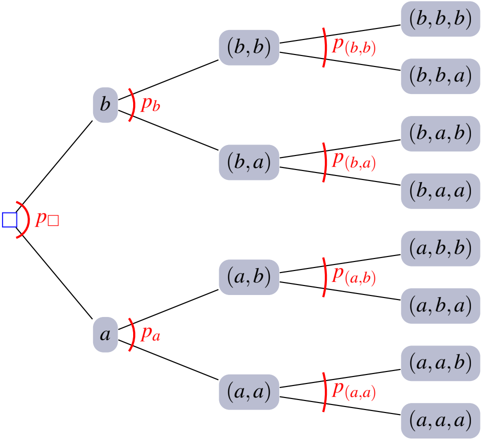
This leads to a standard event tree with nodes

$$s = (x_1, x_2, \dots, x_n), \quad x_k \in \mathscr{X}, \quad n \geq 0.$$

The standard event tree becomes a probability tree by attaching to each node $s$ a local probability mass function $p_s$ on $\mathscr{X}$ with associated expectation operator $E_s$.

# Discrete-time uncertain processes

## Precise probability trees



# PRECISE PROBABILITY TREES

## CALCULATING GLOBAL EXPECTATIONS FROM LOCAL ONES

Consider a function $g\colon \mathscr{X}^n \to \mathbb{R}$ of the first $n$ variables:

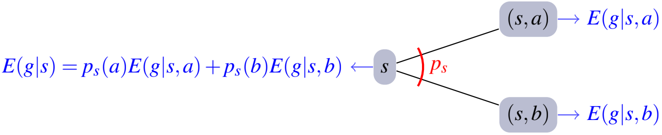
$$g = g(X_1, X_2, \dots, X_n)$$

We want to calculate its expectation $E(g|s)$ in $s = (x_1, \dots, x_k)$.

### THEOREM (LAW OF ITERATED EXPECTATION)

*Suppose we know $E(g|s,x)$ for all $x \in \mathscr{X}$, then we can calculate $E(g|s)$ by backwards recursion using the local model $p_s$:*

$$E(g|s) = \underbrace{E_s}_{\text{local}}(E(g|s,\cdot)) = \sum_{x\in\mathscr{X}} p_s(x)E(g|s,x).$$



# Precise probability trees

## Calculating global expectations from local ones

All expectations $E(g|x_1,\dots,x_k)$ in the tree can be calculated from the local models as follows:

1. start in the final cut $\mathscr{X}^n$ and let:

$$E(g|x_1,x_2,\dots,x_n) = g(x_1,x_2,\dots,x_n);$$

2. do backwards recursion using the Law of Iterated Expectation:

$$E(g|x_1,\dots,x_k) = \underbrace{E_{(x_1,\dots,x_k)}}_{\text{local}}(E(g|x_1,\dots,x_k,\cdot))$$

3. go on until you get to the root node $\square$, where:

$$E(g|\square) = E(g).$$

# EXERCISE

## EVENT TREES

1. Draw the event tree corresponding to three successive flips of a coin (the possible outcomes are heads and tails). Label all situations unambiguously. Differentiate between the root, terminal situations, and intermediate situations.
2. Would you draw a different tree for the successive flips of three different coins?
3. Draw, on the event tree, the cuts corresponding to the following stopping rules:
   - Stop after one flip.
   - Stop after two flips or as soon as heads has come up.
   - Stop when both faces have come up or after the last of the three coin flips.
4. Identify the following events on the event tree (i.e., indicate the corresponding terminal nodes):
   - The result of the first flip is heads.
   - There are two consecutive identical flips.
   - The first two flips are identical.

   Which of these events can be identified with a unique situation (i.e., a not necessarily terminal situation)?

# Homework Problems

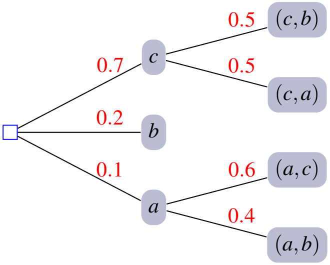
## Event trees

1. Draw the event tree corresponding to
   - A throwing a six faced die (outcomes $1$ to $6$),
   - B followed by again throwing a six-faced die when the outcome is $1$ and a four-faced die (outcomes $1$ to $4$) when the outcome is $5$,
   - C and finally flipping a coin when the sum of the first two outcomes is $7$ or more.
2. Identify the terminal situations. Do they form a cut (of the root)?
3. How many and which cuts are there of the situation '$1$'?
4. For each non-terminal situation, write down the number of children, and then—by using this information—find the number of descendants per node in an efficient manner.

# Exercise

## Probability trees

1. Check that the following is a probability tree:

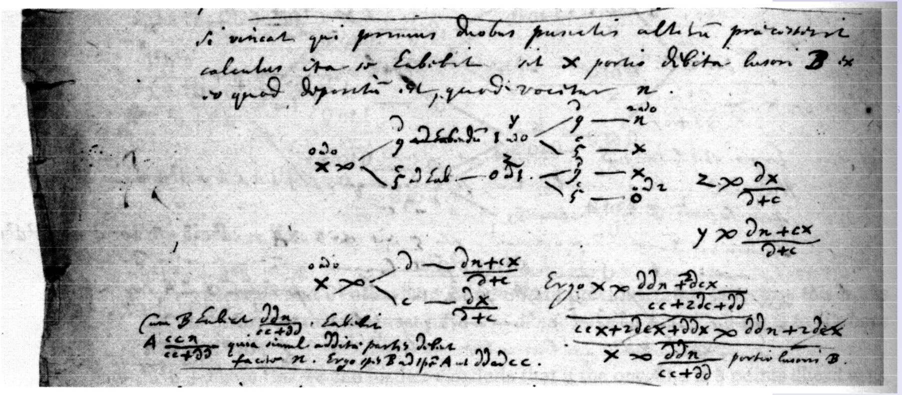



2. Terminal situations containing a vowel yield $1$, all others $-1$. Calculate the expected return in two ways:
   - by forward propagation of probabilities, i.e., using the product rule to calculate the probabilities for each of the terminal situations;
   - by backward-propagation of expectations; write these expectations down in the tree.

# THE FIRST PROBABILITY TREE?

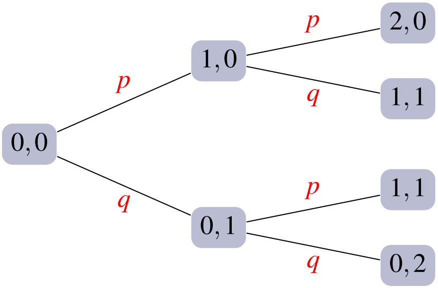
CHRISTIAAN HUYGENS, *Van Rekeningh in Spelen van Geluck (1656–1657)*





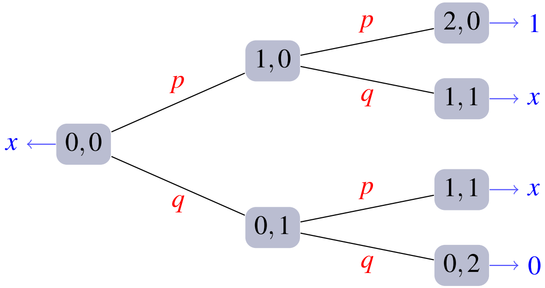
# Huygens's problem

## A more modern version of Huygens's probability tree



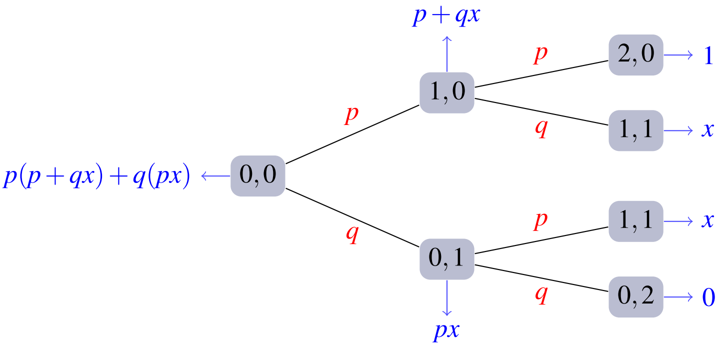
# Huygens's solution

## Adding the probabilities to the picture



# Huygens's solution

## Expectations are calculated backward



# Huygens's solution

## An elegant solution

So we get

$$x = p(p + qx) + q(px)$$

and this leads to:

$$\boxed{x = \frac{p^2}{p^2 + q^2}.}$$

The general solution when the score difference is $n$:

$$\boxed{x = \frac{p^n}{p^n + q^n}.}$$

# Homework Problems

## Probability trees

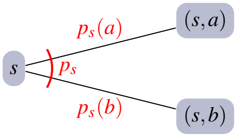
1. Draw the probability tree for the three-step problem of points and calculate, as was done for the two-step case, by identifying equivalent situations and solving for the root expectation.
2. Do the same for the four-step problem of points, but now exploit the solution found for the two-step problem of points.
3. Find the solution to the problem of points for any number of steps $m$.
   *Hint:* Use the Law of Iterated Expectation to find the (second order) difference equation that expresses the relationship between the expectations in the tree as a function of the difference of points for each player. Identify the border conditions to be imposed, and then solve the difference equation.

# IMPRECISE PROBABILITY MODELS

## SETS OF MASS FUNCTIONS

### MAJOR RESTRICTIVE ASSUMPTION

Until now, we have assumed that we have sufficient information in order to specify, in each node $s$, a probability mass function $p_s$ on the set $\mathscr{X}$ of possible values for the next state.




### MORE GENERAL UNCERTAINTY MODELS

We consider credal sets as more general uncertainty models: closed convex subsets of $\Sigma_{\mathscr{X}}$.

# Imprecise probability trees

## Definition and interpretation





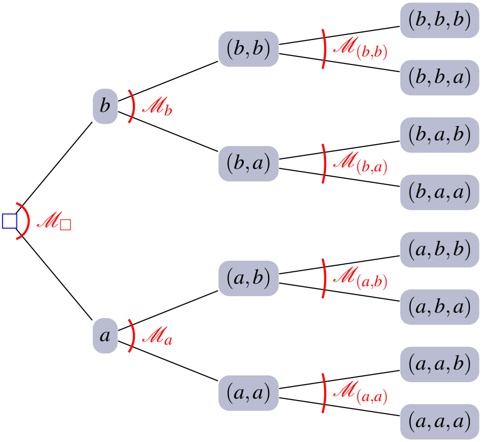
### Definition

An imprecise probability tree is a probability tree where in each node $s$ the local uncertainty model is an imprecise probability model $\mathscr{M}_s$, or equivalently, its associated upper expectation $\overline{E}_s$:

$$\overline{E}_s(f) = \max\left\{E_p(f) \colon p \in \mathscr{M}_s\right\} \text{ for all real maps } f \text{ on } \mathscr{X}.$$

# Imprecise probability trees

## Definition and interpretation



# Imprecise probability trees

## Definition and interpretation

### Definition

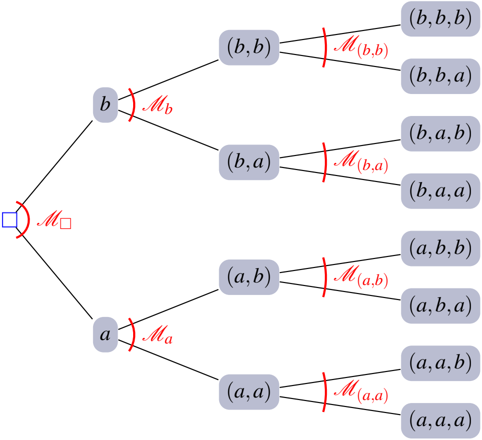
An imprecise probability tree is a probability tree where in each node $s$ the local uncertainty model is an imprecise probability model $\mathscr{M}_s$, or equivalently, its associated upper expectation $\overline{E}_s$:

$$\overline{E}_s(f) = \max\left\{E_p(f) \colon p \in \mathscr{M}_s\right\} \text{ for all real maps } f \text{ on } \mathscr{X}.$$

An imprecise probability tree can be seen as an infinity of compatible precise probability trees: choose in each node $s$ a probability mass function $p_s$ from the set $\mathscr{M}_s$.

# Imprecise probability trees

## Definition and interpretation



# IMPRECISE PROBABILITY TREES

## ASSOCIATED LOWER AND UPPER EXPECTATIONS



For each real map $g = g(X_1, \dots, X_n)$, each node $s = (x_1, \dots, x_k)$, and each such compatible precise probability tree, we can calculate the expectation

$$E(g|x_1, \dots, x_k)$$

using the backwards recursion method described before.

By varying over each compatible probability tree, we get a closed real interval:

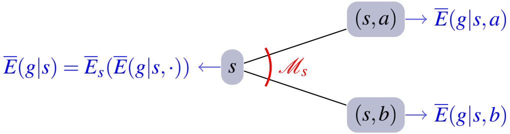
$$[\underline{E}(g|x_1, \dots, x_k), \overline{E}(g|x_1, \dots, x_k)]$$

We want a better, more efficient method to calculate these lower and upper expectations $\underline{E}(g|x_1, \dots x_k)$ and $\overline{E}(g|x_1, \dots, x_k)$.

## THEOREM (LAW OF ITERATED EXPECTATION)

*Suppose we know* $\overline{E}(g|s,x)$ *for all* $x \in \mathscr{X}$, *then we can calculate* $\overline{E}(g|s)$ *by backwards recursion using the local model* $\overline{E}_s$*:*

$$\overline{E}(g|s) = \underbrace{\overline{E}_s}_{\text{local}}(\overline{E}(g|s,\cdot)) = \max_{p_s \in \mathscr{M}_s} \sum_{x \in \mathscr{X}} p_s(x)\overline{E}(g|s,x).$$




The complexity of calculating the $\overline{E}(g|s)$, as a function of $n$, is therefore essentially the same as in the precise case!

# Exercise

## Imprecise probability trees

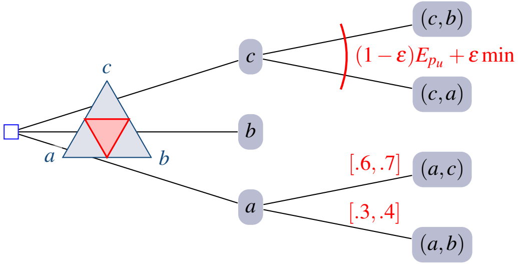
1. Draw the imprecise probability tree corresponding to flipping two coins in succession:
   A. the information available about the first coin flip leads us to assign lower probability $1/4$ to both heads and tails;
   B. the second coin flip is considered to be fair.
2. Calculate the lower and upper probability of getting
   - heads exactly once, and
   - heads at least once.

   *Hint:* First add the 'yields' corresponding to the indicator functions of these events to the terminal nodes and then use backwards recursion.

# Homework problems

## Imprecise probability trees

1. Check that the following is an imprecise probability tree.




Here, $\varepsilon \in [0,1]$ and $p_u = (\frac{1}{2}, \frac{1}{2})$.

2. Again, terminal situations containing a vowel yield $1$, the others $-1$. Calculate the lower and upper expected return using backward recursion. Write these lower and upper expectations down in the tree.

# PRECISE MARKOV CHAINS
## DEFINITION





### DEFINITION

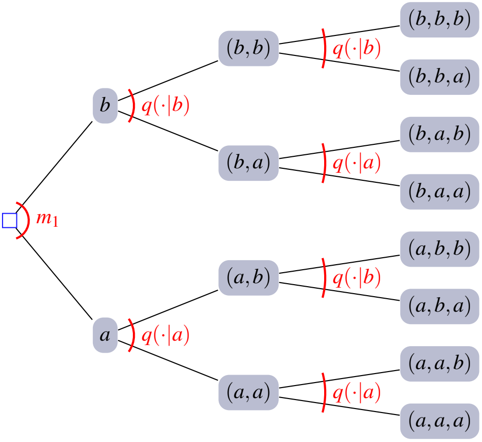
The uncertain process is a stationary precise Markov chain when all $\mathscr{M}_s$ are singletons (precise), and

1. $\mathscr{M}_\Box = \{m_1\}$,
2. the Markov Condition is satisfied:

$$\boxed{\mathscr{M}_{(x_1,\dots,x_n)} = \{q(\cdot|x_n)\}.}$$

# Precise Markov chains

## Definition



# Precise Markov chains

## Definition





### Definition

The uncertain process is a stationary precise Markov chain when all $\mathscr{M}_s$ are singletons (precise), and

1. $\mathscr{M}_\Box = \{m_1\}$,
2. the Markov Condition is satisfied:

$$\mathscr{M}_{(x_1,\dots,x_n)} = \{q(\cdot|x_n)\}.$$

For each $x \in \mathscr{X}$, the transition mass function $q(\cdot|x)$ corresponds to an expectation operator:

$$E(f|x) = \sum_{z \in \mathscr{X}} q(z|x) f(z).$$

# Precise Markov chains

## Transition operators

### Definition

Consider the linear transformation $\mathrm{T}$ of $\mathscr{L}(\mathscr{X})$, called transition operator:

$$\mathrm{T}\colon \mathscr{L}(\mathscr{X}) \to \mathscr{L}(\mathscr{X})\colon f \mapsto \mathrm{T}f$$

where $\mathrm{T}f$ is the real map given by, for any $x \in \mathscr{X}$:

$$\boxed{\mathrm{T}f(x) := E(f|x) = \sum_{z \in \mathscr{X}} q(z|x)f(z)}$$

$\mathrm{T}$ is the dual of the linear transformation with Markov matrix $M$, with elements $M_{xy} := q(y|x)$.

# Precise Markov chains

## Transition operators

### Definition

Consider the linear transformation T of $\mathscr{L}(\mathscr{X})$, called transition operator:

$$\mathrm{T}\colon \mathscr{L}(\mathscr{X}) \to \mathscr{L}(\mathscr{X}) : f \mapsto \mathrm{T}f$$

where $\mathrm{T}f$ is the real map given by, for any $x \in \mathscr{X}$:

$$\boxed{\mathrm{T}f(x) := E(f|x) = \sum_{z \in \mathscr{X}} q(z|x)f(z)}$$

T is the dual of the linear transformation with Markov matrix $M$, with elements $M_{xy} := q(y|x)$.

Then the Law of Iterated Expectation yields:

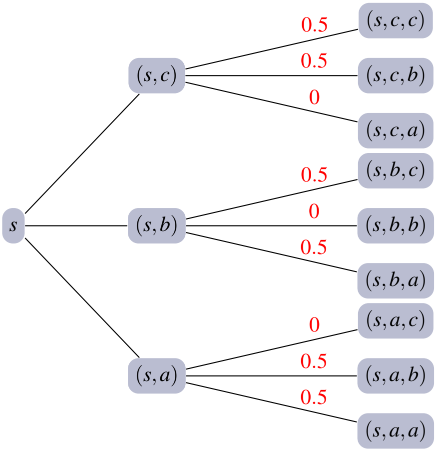
$$\boxed{E_n(f) = E_1(\mathrm{T}^{n-1}f), \text{ and dually, } m_n^T = m_1^T M^{n-1}.}$$

Complexity is linear in the number of time steps $n$. Actually, it is of order $\log_2 n$ using the square-and-multiply algorithm.

# EXERCISE

## PRECISE MARKOV CHAINS

Consider the following partial probability tree characterising a precise Markov chain:



# EXERCISE

## PRECISE MARKOV CHAINS



1. Write down the corresponding Markov matrix $M$.
2. Given the initial mass function described by $m_1 = (0\ 1\ 0)^T$, calculate $m_2$, $m_3$, $m_4$ and $m_5$.
3. Given the gamble $f = (0\ 1\ -1)^T$, calculate $\mathrm{T}f$, $\mathrm{T}^2 f$, $\mathrm{T}^3 f$ and $\mathrm{T}^4 f$.
4. Calculate the expectations $E_1(f)$, $E_2(f)$, $E_3(f)$, $E_4(f)$ and $E_5(f)$ in two ways: using $E_n(f) = m_n^T f$ and $E_n(f) = m_1^T(\mathrm{T}^{n-1} f)$.

# Homework Problems

## Precise Markov chains



Consider your results $m_2$, $m_3$, $m_4$ and $m_5$, and $\mathrm{T}f$, $\mathrm{T}^2f$, $\mathrm{T}^3f$ and $\mathrm{T}^4f$ for the previous exercise.

1. Make an informed guess about what the equilibrium distribution will be on the basis of the observed evolution and the symmetries in $M$. Check your guess.
2. Make an informed guess about what $\lim_{n\to\infty}\mathrm{T}^nf$ will be. Give a proof using induction.

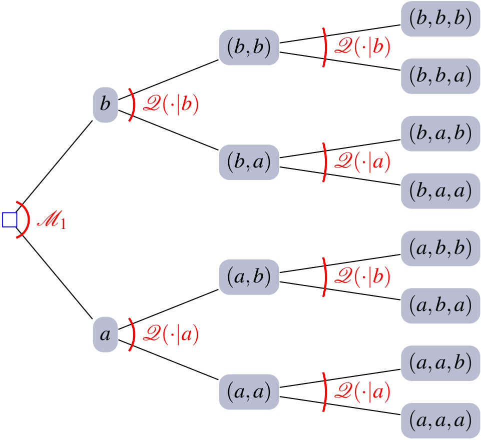
## Definition

The uncertain process is a stationary imprecise Markov chain when the Markov Condition is satisfied:

$$\mathscr{M}_{(x_1,\dots,x_n)} = \mathscr{Q}(\cdot|x_n).$$

# IMPRECISE MARKOV CHAINS

## DEFINITION



# IMPRECISE MARKOV CHAINS

## DEFINITION

### DEFINITION

The uncertain process is a stationary imprecise Markov chain when the Markov Condition is satisfied:

$$\mathscr{M}_{(x_1,\ldots,x_n)} = \mathscr{Q}(\cdot|x_n).$$

An imprecise Markov chain can be seen as an infinity of probability trees.

For each $x \in \mathscr{X}$, the local transition model $\mathscr{Q}(\cdot|x)$ corresponds to lower and upper expectation operators:

$$\underline{E}(f|x) = \min\left\{E_p(f) : p \in \mathscr{Q}(\cdot|x)\right\}$$
$$\overline{E}(f|x) = \max\left\{E_p(f) : p \in \mathscr{Q}(\cdot|x)\right\}.$$

## Definition

Consider the non-linear transformations $\underline{\mathrm{T}}$ and $\overline{\mathrm{T}}$ of $\mathscr{L}(\mathscr{X})$, called lower and upper transition operators:

$$\underline{\mathrm{T}}\colon \mathscr{L}(\mathscr{X}) \to \mathscr{L}(\mathscr{X}) : f \mapsto \underline{\mathrm{T}}f$$
$$\overline{\mathrm{T}}\colon \mathscr{L}(\mathscr{X}) \to \mathscr{L}(\mathscr{X}) : f \mapsto \overline{\mathrm{T}}f$$

where the real maps $\underline{\mathrm{T}}f$ and $\overline{\mathrm{T}}f$ are given by:

$$\underline{\mathrm{T}}f(x) := \underline{E}(f|x) = \min\left\{E_p(f) : p \in \mathscr{Q}(\cdot|x)\right\}$$
$$\overline{\mathrm{T}}f(x) := \overline{E}(f|x) = \max\left\{E_p(f) : p \in \mathscr{Q}(\cdot|x)\right\}$$

# IMPRECISE MARKOV CHAINS

## LOWER AND UPPER TRANSITION OPERATORS

### DEFINITION

Consider the non-linear transformations $\underline{\mathrm{T}}$ and $\overline{\mathrm{T}}$ of $\mathscr{L}(\mathscr{X})$, called lower and upper transition operators:

$$\underline{\mathrm{T}}\colon \mathscr{L}(\mathscr{X}) \to \mathscr{L}(\mathscr{X}) \colon f \mapsto \underline{\mathrm{T}}f$$
$$\overline{\mathrm{T}}\colon \mathscr{L}(\mathscr{X}) \to \mathscr{L}(\mathscr{X}) \colon f \mapsto \overline{\mathrm{T}}f$$

where the real maps $\underline{\mathrm{T}}f$ and $\overline{\mathrm{T}}f$ are given by:

$$\underline{\mathrm{T}}f(x) := \underline{E}(f|x) = \min\left\{E_p(f) \colon p \in \mathscr{Q}(\cdot|x)\right\}$$
$$\overline{\mathrm{T}}f(x) := \overline{E}(f|x) = \max\left\{E_p(f) \colon p \in \mathscr{Q}(\cdot|x)\right\}$$

Then the Law of Iterated Expectation yields:

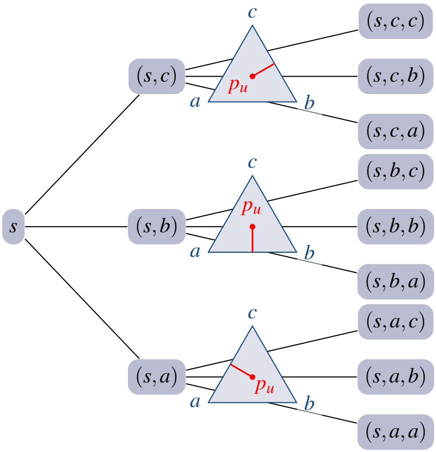
$$\underline{E}_n(f) = \underline{E}_1(\underline{\mathrm{T}}^{n-1}f) \text{ and } \overline{E}_n(f) = \overline{E}_1(\overline{\mathrm{T}}^{n-1}f).$$

Complexity is still linear in the number of time steps $n$.

# Exercise

## Imprecise Markov chains

Given is the following partial imprecise probability tree characterising an imprecise Markov chain:




Here $p_u = (\frac{1}{3}, \frac{1}{3}, \frac{1}{3})$.

# Exercise

## Imprecise Markov chains

1. Given the gamble $f = (0\ 1\ -1)^T$, calculate $[\underline{\mathrm{T}}f, \overline{\mathrm{T}}f]$ and $[\underline{\mathrm{T}}^2 f, \overline{\mathrm{T}}^2 f]$.
   *Hint:* It may be easiest to do the backwards recursion calculations iteratively in the partial tree.
2. Given the initial mass function described by $m_1 = (0\ 1\ 0)^T$, calculate the lower and upper expectations $[\underline{E}_1(f), \overline{E}_1(f)]$, $[\underline{E}_2(f), \overline{E}_2(f)]$ and $[\underline{E}_3(f), \overline{E}_3(f)]$.
3. Based on $[\underline{\mathrm{T}}^2 f, \overline{\mathrm{T}}^2 f]$, what bounds can you put on $\lim_{n\to\infty}[\underline{E}_n(f), \overline{E}_n(f)]$?

# HOMEWORK PROBLEMS

## IMPRECISE MARKOV CHAINS

1. How would you calculate lower and upper mass functions after $n$ steps, i.e., which gambles would you need to calculate the different components of the corresponding vector?
2. For general imprecise Markov chains, do the lower and upper mass functions after $n$ steps fully characterise the uncertainty about the state after $n$ steps? Why (not)?
3. Investigate the complexity of working with precise and imprecise Markov chains; focus on the number and type of computations and memory necessary for calculating the expectation or lower expectation of a gamble after $n$ steps for $m$-state chains.

# Random Walks

## An example with a two-element state space

Consider a two-element state space:

$$\mathscr{X} = \{a, b\},$$

with upper expectation $\overline{E}_1$ for the first state, and for each $(x_1, \ldots, x_n) \in \{a, b\}^n$, with $\varepsilon \in [0, 1]$,

$$\mathscr{M}_{(x_1,\ldots,x_n)} = \mathscr{Q}(\cdot|x_n) = (1-\varepsilon)\{q(\cdot|x_n)\} + \varepsilon \Sigma_{\{a,b\}}$$

or in other words, for the upper transition operator

$$\boxed{\overline{\mathrm{T}} = (1-\varepsilon)\mathrm{T} + \varepsilon \max}$$

where $\mathrm{T}$ is the linear transition operator determined by

$$M := \begin{bmatrix} \mathrm{T}I_{\{a\}}(a) & \mathrm{T}I_{\{b\}}(a) \\ \mathrm{T}I_{\{a\}}(b) & \mathrm{T}I_{\{b\}}(b) \end{bmatrix} = \begin{bmatrix} q(a|a) & q(b|a) \\ q(a|b) & q(b|b) \end{bmatrix}.$$

# Random Walks

## Stationary distribution

It is a matter of simple verification that for $n \geq 1$ and $f \in \mathscr{L}(\mathscr{X})$:

$$\overline{\mathrm{T}}^n f = (1-\varepsilon)^n \mathrm{T}^n f + \varepsilon \sum_{k=0}^{n-1} (1-\varepsilon)^k \max \mathrm{T}^k f,$$

and therefore, using the Law of Iterated Expectation,

$$\overline{E}_{n+1}(f) = \overline{E}_1(\overline{\mathrm{T}}^n f) = (1-\varepsilon)^n \overline{E}_1(\mathrm{T}^n f) + \varepsilon \sum_{k=0}^{n-1} (1-\varepsilon)^k \max \mathrm{T}^k f.$$

If we now let $n \to \infty$, we see that the limit exists and is independent of the initial upper expectation $\overline{E}_1$:

$$\boxed{\overline{E}_\infty(f) = \varepsilon \sum_{k=0}^{\infty} (1-\varepsilon)^k \max \mathrm{T}^k f.}$$

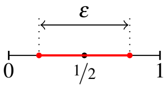
### Contaminated Random Walk

When

$$\mathrm{T}f(a) = \mathrm{T}f(b) = 1/2[f(a) + f(b)], \text{ i.e., } M = \begin{bmatrix} 1/2 & 1/2 \\ 1/2 & 1/2 \end{bmatrix}$$

then we find that

$$\boxed{\overline{E}_\infty(f) = (1 - \varepsilon)1/2[f(a) + f(b)] + \varepsilon \max f.}$$



# Random Walks

## Special cases

### Contaminated Cycle

When

$$\mathrm{T}f(a) = f(b) \text{ and } \mathrm{T}f(b) = f(a), \text{ i.e., } M = \begin{bmatrix} 0 & 1 \\ 1 & 0 \end{bmatrix}$$

then we find that

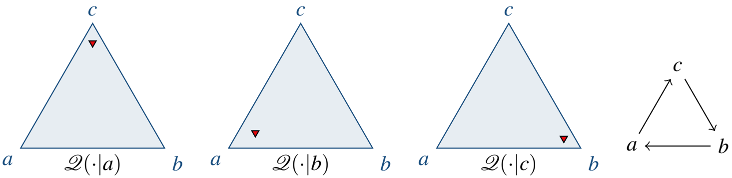
$$\boxed{\overline{E}_{\infty}(f) = \max f.}$$

# Lower and upper mass functions

## Another example with $\mathscr{X} = \{a,b,c\}$

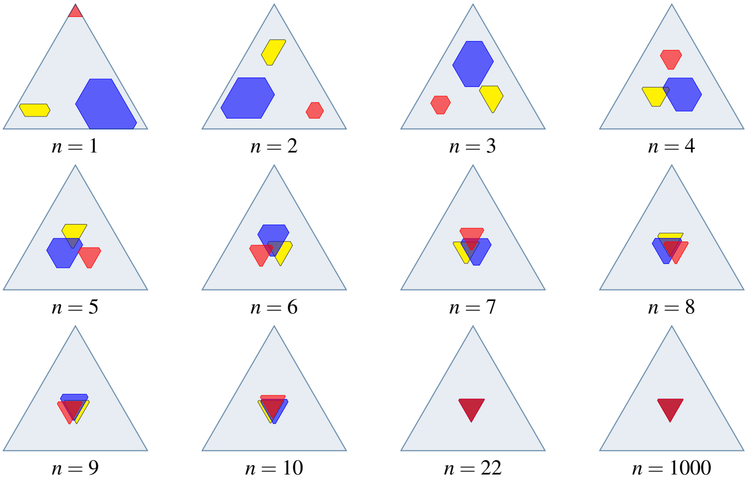
$$\begin{bmatrix}\underline{\mathrm{T}}I_{\{a\}} & \underline{\mathrm{T}}I_{\{b\}} & \underline{\mathrm{T}}I_{\{c\}}\end{bmatrix} = \begin{bmatrix}\underline{q}(a|a) & \underline{q}(b|a) & \underline{q}(c|a)\\ \underline{q}(a|b) & \underline{q}(b|b) & \underline{q}(c|b)\\ \underline{q}(a|c) & \underline{q}(b|c) & \underline{q}(c|c)\end{bmatrix} = {}^1\!/_{200}\begin{bmatrix}9 & 9 & 162\\ 144 & 18 & 18\\ 9 & 162 & 9\end{bmatrix}$$

$$\begin{bmatrix}\overline{\mathrm{T}}I_{\{a\}} & \overline{\mathrm{T}}I_{\{b\}} & \overline{\mathrm{T}}I_{\{c\}}\end{bmatrix} = \begin{bmatrix}\overline{q}(a|a) & \overline{q}(b|a) & \overline{q}(c|a)\\ \overline{q}(a|b) & \overline{q}(b|b) & \overline{q}(c|b)\\ \overline{q}(a|c) & \overline{q}(b|c) & \overline{q}(c|c)\end{bmatrix} = {}^1\!/_{200}\begin{bmatrix}19 & 19 & 172\\ 154 & 28 & 28\\ 19 & 172 & 19\end{bmatrix}$$



# Lower and upper mass functions

## Another example with $\mathscr{X} = \{a, b, c\}$



# Non-linear Perron–Frobenius Theorem

## Generalising the linear case

### Theorem (De Cooman, Hermans and Quaeghebeur, 2008)

*Consider a stationary imprecise Markov chain with finite state set* $\mathscr{X}$ *and an upper transition operator* $\overline{\mathrm{T}}$*. Suppose that* $\overline{\mathrm{T}}$ *is* *regular, meaning that there is some* $n > 0$ *such that* $\min \overline{\mathrm{T}}^n I_{\{x\}} > 0$ *for all* $x \in \mathscr{X}$*. Then for every initial upper expectation* $\overline{E}_1$*, the upper expectation* $\overline{E}_n = \overline{E}_1 \circ \overline{\mathrm{T}}^{n-1}$ *for the state at time* $n$ *converges point-wise to the same upper expectation* $\overline{E}_\infty$*:*

$$\lim_{n\to\infty} \overline{E}_n(h) = \lim_{n\to\infty} \overline{E}_1(\overline{\mathrm{T}}^{n-1} h) := \overline{E}_\infty(h)$$

*for all* $h$ *in* $\mathscr{L}(\mathscr{X})$*. Moreover, the corresponding limit upper expectation* $\overline{E}_\infty$ *is the only* $\overline{\mathrm{T}}$*-invariant upper expectation on* $\mathscr{L}(\mathscr{X})$*, meaning that* $\overline{E}_\infty = \overline{E}_\infty \circ \overline{\mathrm{T}}$*.*

# MEAN FIRST PASSAGE TIMES

## DEFINITION





Let the random process $\tau_{xy}$ be the first time $n > 0$ such that $X_{n+1} = y$, if the process starts out in $X_1 = x$.

We are interested in the lower and upper mean first passage times:

$$\boxed{\underline{M}_{xy} = \underline{E}(\tau_{xy}|x) \text{ and } \overline{M}_{xy} = \overline{E}(\tau_{xy}|x).}$$

If $x = y$, we call

$$\boxed{\underline{R}_x := \underline{M}_{xx} = \underline{E}(\tau_{xx}|x) \text{ and } \overline{R}_x := \overline{M}_{xx} = \overline{E}(\tau_{xx}|x)}$$

lower and upper mean recurrence times.

# Mean first passage times

## Non-linear equations for mean first passage times



Now for any trajectory $(x, x_2, x_3, \dots)$ starting in $x$:

$$\tau_{xy}(x, x_2, x_3, \dots) = \begin{cases} 1 & ; \quad x_2 = y \\ 1 + \tau_{x_2 y}(x_2, x_3, \dots) & ; \quad x_2 \neq y \end{cases}$$

which is a recursive relation, so if we use the Law of Iterated Expectation, stationarity and the Markov Property, we are led to the non-linear equations:

$$\underline{M}_{\cdot y} = 1 + \underline{\mathrm{T}}[(1 - \delta_{\cdot y})\underline{M}_{\cdot y}] \text{ and } \overline{M}_{\cdot y} = 1 + \overline{\mathrm{T}}[(1 - \delta_{\cdot y})\overline{M}_{\cdot y}].$$

# Mean first passage times

## Examples

We find after solving the non-linear equations that:

### Contaminated random walk

$$\underline{R}_a = \underline{R}_b = \underline{M}_{ab} = \underline{M}_{ba} = \frac{2}{1+\varepsilon}$$
$$\overline{R}_a = \overline{R}_b = \overline{M}_{ab} = \overline{M}_{ba} = \frac{2}{1-\varepsilon}.$$

### Contaminated cycle

$$\underline{R}_a = \underline{R}_b = 2-\varepsilon \text{ and } \underline{M}_{ab} = \underline{M}_{ba} = 1$$
$$\overline{R}_a = \overline{R}_b = \frac{2-\varepsilon}{1-\varepsilon} \text{ and } \overline{M}_{ab} = \overline{M}_{ba} = \frac{1}{1-\varepsilon}.$$

# A SPECIAL CREDAL NETWORK

## UNDER EPISTEMIC IRRELEVANCE

An imprecise Markov chain can also be depicted as follows:

$$X_1 \longrightarrow X_2 \longrightarrow X_3 \longrightarrow \cdots \longrightarrow X_{n-1} \longrightarrow X_n$$

## INTERPRETATION OF THE GRAPH

Conditional on $X_k$ we have that $X_1, \dots, X_{k-1}$ are epistemically irrelevant to $X_{k+1}, \dots, X_n$:

$$\overline{E}(f(X_{k+1},\dots,X_n)|X_1,\dots,X_{k-1},X_k) = \overline{E}(f(X_{k+1},\dots,X_n)|X_k)$$

## MORE GENERALLY, FOR A CREDAL NET

Conditionally on the parents, the non-parent non-descendants of each node are epistemically irrelevant to it.

# Separation in credal nets

## under epistemic irrelevance


Figure: $I_2$ separates $T$ from $I_1$.

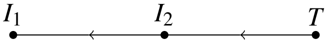
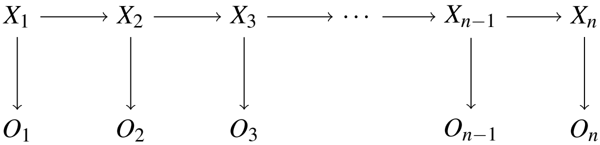



Figure: $I_2$ doesn't separate $T$ from $I_1$.

### Conclusion

For a variable $T$ to be separated from $I_2$ by a variable $I_1$, arrows should point from $I_2$ to $T$.

# A SPECIAL CASE
## HIDDEN MARKOV CHAINS



# LITERATURE


Gert de Cooman and Filip Hermans.
Imprecise probability trees: Bridging two theories of imprecise probability.
*Artificial Intelligence*, 2008, vol. 172, pp. 1400–1427. (arXiv:0801.1196v1).

Gert de Cooman, Filip Hermans, and Erik Quaeghebeur.
Imprecise Markov chains and their limit behaviour.
Submitted for publication (arXiv:0801.0980).

Glenn Shafer and Vladimir Vovk.
*Probability and Finance: It's Only a Game!*
Wiley, New York, 2001.

Peter Walley.
*Statistical Reasoning with Imprecise Probabilities.*
Chapman and Hall, London, 1991.