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
Assume we are uncertain about:

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

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

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

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

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

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

$$P_p(A) = \sum_{x \in A} p(x) = E_p(I_A).$$
Consider the simplex $\Sigma_X$ of all mass functions on $X$:

$$\Sigma_X := \left\{ p \in \mathbb{R}_+^X : \sum_{x \in X} p(x) = 1 \right\}.$$
PRECISE PROBABILITY MODELS

GEOMETRICAL INTERPRETATION OF EXPECTATION

ASSESSMENTS LEAD TO CONSTRAINTS
Specifying an expectation $E(f)$ for a map $f : \mathcal{X} \rightarrow \mathbb{R}$

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

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

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

$E(2I\{b\} - I\{c\}) = 0$

$E(I\{a\}) = 1/2$
Imprecise probability models
Linear inequality constraints

More flexible assessments
Impose linear inequality constraints on $p$ in $\Sigma X$:

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

Corresponds to intersecting $\Sigma X$ with affine semi-spaces:
Imprecise probability models

Imprecise immediate predictions

Imprecise probabilities

Precise probability trees

Huynghen’s tree

Imprecise probability trees

Imprecise Markov chains

Perron–Frobenius theorem

First passage

Towards credal nets

Literature

More flexible assessments

Impose linear inequality constraints on \( p \) in \( \Sigma \mathcal{X} \):

\[
\bar{E}(f) \leq \sum_{x \in \mathcal{X}} p(x)f(x) \quad \text{or} \quad \sum_{x \in \mathcal{X}} p(x)f(x) \leq \bar{E}(f).
\]

Corresponds to intersecting \( \Sigma \mathcal{X} \) with affine semi-spaces:
Imprecise probability models

Linear inequality constraints

**More flexible assessments**

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

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

Corresponds to intersecting $\Sigma_\mathcal{X}$ with affine semi-spaces:
Imprecise probability models

Linear inequality constraints

More flexible assessments

Impose linear inequality constraints on \( p \) in \( \Sigma \mathcal{X} \):

\[
E(f) \leq \sum_{x \in \mathcal{X}} p(x)f(x) \quad \text{or} \quad \sum_{x \in \mathcal{X}} p(x)f(x) \leq E(f).
\]

Corresponds to intersecting \( \Sigma \mathcal{X} \) with affine semi-spaces:
Imprecise probability models

Linear inequality constraints

More flexible assessments
Impose linear inequality constraints on $p$ in $\Sigma X$:

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

Corresponds to intersecting $\Sigma X$ with affine semi-spaces:
Imprecise probability models

Credal sets

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

**Definition**

A credal set $\mathcal{M}$ is a convex closed subset of $\Sigma^X$. 

![Diagram of credal sets](image)
**Imprecise Probability Models**

**Lower and Upper Expectations**

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

\[
\begin{align*}
E_M(f) &= \min \{ E_p(f) : p \in \mathcal{M} \} \text{ lower expectation} \\
\overline{E}_M(f) &= \max \{ E_p(f) : p \in \mathcal{M} \} \text{ upper expectation.}
\end{align*}
\]

Observe that [conjugacy]

\[
\overline{E}_M(f) = -E_M(-f).
\]
**IMPRECISE PROBABILITY MODELS**

**Basic properties of upper expectations**

**Definition**
We call a real functional \( \overline{E} \) on \( \mathcal{L}(\mathcal{X}) \) an upper expectation if it satisfies the following properties:
For all \( f \) and \( g \) in \( \mathcal{L}(\mathcal{X}) \) 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 \( E \). Then for all real numbers \( \mu \) and all \( f \) and \( g \) in \( \mathcal{L}(\mathcal{X}) \):

1. \( E(f) \leq \overline{E}(f) \);
2. \( E(f) + E(g) \leq \overline{E}(f + g) \leq \overline{E}(f) + \overline{E}(g) \leq \overline{E}(f + g) \leq \overline{E}(f) + \overline{E}(g) \);
3. \( \overline{E}(f + \mu) = \overline{E}(f) + \mu \);
4. \( \overline{E}|f|) \geq |E(f)| \) and \( \overline{E}|f|) \geq |E(f)| \).
Theorem (Lower Envelope Theorem)

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

Proof.
Use $\mathcal{M} = \{ p \in \Sigma_{\mathcal{X}} : (\forall f \in \mathcal{L}(\mathcal{X}))(E_p(f) \leq \bar{E}(f)) \}$.  \[ \square \]
We consider an uncertain process with variables $X_1, X_2, \ldots, X_n, \ldots$ assuming values in a finite set of states $\mathcal{X}$.

This leads to a standard event tree with nodes

$$s = (x_1, x_2, \ldots, x_n), \quad x_k \in \mathcal{X}, \quad n \geq 0.$$
The standard event tree becomes a probability tree by attaching to each node \( s \) a local probability mass function \( p_s \) on \( X \) with associated expectation operator \( E_s \).
We consider an uncertain process with variables $X_1, X_2, \ldots, X_n, \ldots$ assuming values in a finite set of states $\mathcal{X}$.

This leads to a standard event tree with nodes

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

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

**Discrete-time uncertain processes**

**Precise probability trees**
Discrete-time uncertain processes

Precise probability trees

\[
\begin{align*}
&\text{b} \xrightarrow{p_b} (b, b) \xrightarrow{P(b,b)} (b, b, b) \\
&\quad \xrightarrow{P(b,a)} (b, a, b) \\
&\quad \xrightarrow{P(a,b)} (a, b, b) \\
&\quad \xrightarrow{P(a,a)} (a, a, b) \\
&\text{a} \xrightarrow{p_a} (a, b) \xrightarrow{P(a,b)} (a, b, a) \\
&\quad \xrightarrow{P(a,a)} (a, a, a) \\
&\text{b} \xrightarrow{p_b} (b, b) \xrightarrow{P(b,a)} (b, b, a) \\
&\quad \xrightarrow{P(a,b)} (b, a, a) \\
&\quad \xrightarrow{P(a,a)} (b, a, b) \\
&\text{a} \xrightarrow{p_a} (a, a) \xrightarrow{P(a,a)} (a, a, a)
\end{align*}
\]
Precise probability trees
Calculating global expectations from local ones

Consider a function \( g : \mathcal{X}^n \to \mathbb{R} \) of the first \( n \) variables:

\[
g = g(X_1, X_2, \ldots, X_n)
\]

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

**Theorem (Law of Iterated Expectation)**

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

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

\[
E(g|s) = p_s(a)E(g|s,a) + p_s(b)E(g|s,b)
\]

\[
(s,a) \to E(g|s,a)
\]

\[
(s,b) \to E(g|s,b)
\]
PRECISE PROBABILITY TREES
CALCULATING GLOBAL EXPECTATIONS FROM LOCAL ONES

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

1. start in the final cut $\mathcal{X}^n$ and let:
   
   $$E(g|x_1, x_2, \ldots, x_n) = g(x_1, x_2, \ldots, x_n);$$

2. do backwards recursion using the Law of Iterated Expectation:
   
   $$E(g|x_1, \ldots, x_k) = E(x_1, \ldots, x_k) (E(g|x_1, \ldots, x_k, \cdot)_{\text{local}})$$

3. go on until you get to the root node $\Box$, where:
   
   $$E(g|\Box) = E(g).$$
**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:

```
   a
    / \   /  \
   /   \ /   \n b     c     (c, b)
   / \
0.2/  \
0.7/    \
   a    (c, a)
   /  \\
0.1/    \
   (a, c)
   /  \\
0.1/    \
   (a, b)
```

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

\[ p(p + qx) + q(px) \]

\[ \begin{align*}
0,0 & \quad \text{0,0} \\
1,0 & \quad \text{1,0} \quad p + qx \\
0,1 & \quad \text{0,1} \quad p \\
1,1 & \quad \text{1,1} \quad q x \\
0,2 & \quad \text{0,2} \\
2,0 & \quad \text{2,0} \quad 1 \\
1,1 & \quad \text{1,1} \quad x \\
0,2 & \quad \text{0,2} \quad 0
\end{align*} \]
HUYGENS’S SOLUTION
AN ELEGANT SOLUTION

So we get

\[ x = p(p + qx) + q(px) \]

and this leads to:

\[ x = \frac{p^2}{p^2 + q^2}. \]

The general solution when the score difference is \( n \):

\[ x = \frac{p^n}{p^n + q^n}. \]
**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 $\mathcal{X}$ of possible values for the next state.

MORE GENERAL UNCERTAINTY MODELS
We consider credal sets as more general uncertainty models: closed convex subsets of $\Sigma \mathcal{X}$.
An imprecise probability tree is a probability tree where in each node \( s \) the local uncertainty model is an imprecise probability model \( \mathcal{M}_s \), or equivalently, its associated upper expectation \( \overline{E}_s \):\[\overline{E}_s(f) = \max \{ E_p(f) : p \in \mathcal{M}_s \} \text{ for all real maps } f \text{ on } X.\]
**Imprecise Probability Trees**

**Definition and Interpretation**

An imprecise probability tree can be seen as an infinity of compatible precise probability trees: choose in each node a probability mass function $p_s$ from the set $M_s$. 
**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 $\mathcal{M}_s$, or equivalently, its associated upper expectation $\overline{E}_s$:

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

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 $\mathcal{M}_s$. 

---

**Imprecise immediate predictions**

**GdC,EQ,FH**

**Imprecise probabilities**

**Precise probability trees**

**Huygens’s tree**

**Imprecise probability trees**

**Imprecise Markov chains**

**Perron–Frobenius Theorem**

**First passage**

**Towards credal nets**

**Literature**
IMPRECISE PROBABILITY TREES

DEFINITION AND INTERPRETATION

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

- $\mathcal{M}_b = \{ (b, b), (b, b, b) \}$
- $\mathcal{M}_a = \{ (a, b), (a, a), (a, a, a) \}$

The tree structure and the compatible sets of probability mass functions illustrate the concept of imprecise probability trees.
IMPRECISE PROBABILITY TREES
ASSOCIATED LOWER AND UPPER EXPECTATIONS

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

\[
E(g|x_1, \ldots, x_k)
\]

using the backwards recursion method described before.

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

\[
[E(g|x_1, \ldots, x_k), \overline{E}(g|x_1, \ldots, x_k)]
\]

We want a better, more efficient method to calculate these lower and upper expectations \( E(g|x_1, \ldots x_k) \) and \( \overline{E}(g|x_1, \ldots, x_k) \).
Theorem (Law of Iterated Expectation)

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

$$\bar{E}(g|s) = \max_{p_s \in \mathcal{M}_s} \sum_{x \in \mathcal{X}} p_s(x) \bar{E}(g|s,x).$$

The complexity of calculating the $\bar{E}(g|s)$, 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 $\frac{1}{4}$ 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.
**Definition**

The uncertain process is a stationary precise Markov chain when all $M_s$ are singletons (precise), and

1. $M_{\square} = \{m_1\}$,
2. the Markov Condition is satisfied:

$$M(x_1, \ldots, x_n) = \{q(\cdot|x_n)\}.$$
Precise Markov Chains

Definition

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

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

Diagram:

- Node $m_1$ with transitions $(a, a), (a, b), (b, a), (b, b)$.
- Node $b$ with transitions $(b, a), (b, b), q(\cdot | b)$.
- Node $a$ with transitions $(a, a), (a, b), q(\cdot | a)$. 
- Leaves: $(a, a, a), (a, b, a), (b, a, a), (b, b, b)$. 

Transition mass functions: $q(\cdot | a), q(\cdot | b)$. 

Perron–Frobenius Theorem

First Passage

Towards Credal Nets

Literature
**Precise Markov chains**

**Definition**

The uncertain process is a stationary precise Markov chain when all $M_s$ are singletons (precise), and

1. $M_\square = \{ m_1 \}$,
2. the Markov Condition is satisfied:

$$M(x_1, \ldots, x_n) = \{ q(\cdot | x_n) \}.$$

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

$$E(f | x) = \sum_{z \in \mathcal{X}} q(z | x)f(z).$$
Precise Markov chains

Transition operators

Definition

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

$$ T: \mathcal{L}(\mathcal{X}) \rightarrow \mathcal{L}(\mathcal{X}): f \mapsto Tf $$

where $Tf$ is the real map given by, for any $x \in \mathcal{X}$:

$$ Tf(x) := E(f|x) = \sum_{z \in \mathcal{X}} q(z|x)f(z) $$

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

**TRANSITION OPERATORS**

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

$$T: \mathcal{L}(\mathcal{X}) \rightarrow \mathcal{L}(\mathcal{X}): f \mapsto Tf$$

where $Tf$ is the real map given by, for any $x \in \mathcal{X}$:

$$Tf(x) := E(f|x) = \sum_{z \in \mathcal{X}} q(z|x)f(z)$$

$T$ 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:

$$E_n(f) = E_1(T^{n-1}f), \text{ and dually, } m^T_n = m^T_1 M^{n-1}.$$  

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 $Tf$, $T^2f$, $T^3f$ and $T^4f$.
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 $E_n(f) = m_1^T (T^{n-1}f)$. 

**Literature**

- Imprecise immediate predictions
- GdC.EQ.FH
- Imprecise probabilities
- Precise probability trees
- Huygens’ tree
- Imprecise probability trees
- Imprecise Markov chains
- Perron–Frobenius theorem
- First passage
- Towards credal nets
Consider your results $m_2, m_3, m_4$ and $m_5$, and $Tf, T^2f, T^3f$ and $T^4f$ 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 $\lim_{n \to \infty} T^n f$ will be. Give a proof using induction.
**Imprecise Markov Chains**

**Definition**

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

\[ \mathcal{M}(x_1, \ldots, x_n) = \mathcal{D}(\cdot | x_n). \]
An imprecise Markov chain can be seen as an infinity of probability trees. For each $x \in \mathcal{X}$, the local transition model $\mathcal{Q}(\cdot|x)$ corresponds to lower and upper expectation operators:

$$
\mathcal{E}(f|x) = \min_{p \in \mathcal{Q}(\cdot|x)} \mathcal{E}_p(f)
$$

$$
\mathcal{E}(f|x) = \max_{p \in \mathcal{Q}(\cdot|x)} \mathcal{E}_p(f)
$$
**Imprecise Markov chains**

**Definition**

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

\[
\mathcal{M}(x_1, \ldots, x_n) = \mathcal{D}(\cdot | x_n).
\]

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

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

\[
\underline{E}(f|x) = \min \{ E_p(f) : p \in \mathcal{D}(\cdot | x) \}
\]

\[
\overline{E}(f|x) = \max \{ E_p(f) : p \in \mathcal{D}(\cdot | x) \}.
\]
**Imprecise Markov chains**

**Lower and upper transition operators**

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

$$\underline{T} : \mathcal{L}(\mathcal{X}) \rightarrow \mathcal{L}(\mathcal{X}) : f \mapsto \underline{T}f$$
$$\overline{T} : \mathcal{L}(\mathcal{X}) \rightarrow \mathcal{L}(\mathcal{X}) : f \mapsto \overline{T}f$$

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

$$\underline{T}f(x) := \underline{E}(f|x) = \min \left\{ E_p(f) : p \in \mathcal{Q}(\cdot|x) \right\}$$
$$\overline{T}f(x) := \overline{E}(f|x) = \max \left\{ E_p(f) : p \in \mathcal{Q}(\cdot|x) \right\}$$
**Imprecise Markov Chains**

**Lower and Upper Transition Operators**

**Definition**

Consider the non-linear transformations \( T \) and \( \bar{T} \) of \( \mathcal{L}(\mathcal{X}) \), called lower and upper transition operators:

\[
T: \mathcal{L}(\mathcal{X}) \rightarrow \mathcal{L}(\mathcal{X}): f \mapsto Tf \\
\bar{T}: \mathcal{L}(\mathcal{X}) \rightarrow \mathcal{L}(\mathcal{X}): f \mapsto \bar{T}f
\]

where the real maps \( Tf \) and \( \bar{T}f \) are given by:

\[
Tf(x) := E(f|x) = \min \{ E_p(f) : p \in \mathcal{Q}(\cdot|x) \} \\
\bar{T}f(x) := \bar{E}(f|x) = \max \{ E_p(f) : p \in \mathcal{Q}(\cdot|x) \}
\]

Then the Law of Iterated Expectation yields:

\[
E_n(f) = E_1(T^{n-1}f) \quad \text{and} \quad \bar{E}_n(f) = \bar{E}_1(\bar{T}^{n-1}f).
\]

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:

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

Here \( p_u = \left( \frac{1}{3}, \frac{1}{3}, \frac{1}{3} \right) \).
**Exercise**

**Imprecise Markov chains**

1. Given the gamble $f = (0 \ 1 \ -1)^T$, calculate $[Tf, \overline{Tf}]$ and $[T^2f, \overline{T^2f}]$.
   
   **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 $[E_1(f), \overline{E_1(f)}]$, $[E_2(f), \overline{E_2(f)}]$ and $[E_3(f), \overline{E_3(f)}]$.

3. Based on $[T^2f, \overline{T^2f}]$, what bounds can you put on $\lim_{n \to \infty} [E_n(f), \overline{E_n(f)}]$?
**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:

\[ X = \{a, b\}, \]

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

\[ M(x_1, \ldots, x_n) = \mathcal{D}(\cdot|x_n) = (1 - \varepsilon) \{q(\cdot|x_n)\} + \varepsilon \Sigma_{a,b} \]

or in other words, for the **upper transition operator**

\[ \overline{T} = (1 - \varepsilon)T + \varepsilon \max \]

where \( T \) is the **linear transition operator** determined by

\[ M := \begin{bmatrix} TI_{\{a\}}(a) & TI_{\{b\}}(a) \\ TI_{\{a\}}(b) & TI_{\{b\}}(b) \end{bmatrix} = \begin{bmatrix} q(a|a) & q(b|a) \\ q(a|b) & q(b|b) \end{bmatrix}. \]
It is a matter of simple verification that for $n \geq 1$ and $f \in \mathcal{L}(X)$:

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

and therefore, using the Law of Iterated Expectation,

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

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

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

When

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

then we find that

\[ E_\infty(f) = (1 - \varepsilon)^{1/2} [f(a) + f(b)] + \varepsilon \max f. \]
RANDOM WALKS

SPECIAL CASES

CONTAMINATED CYCLE

When

\[ Tf(a) = f(b) \text{ and } Tf(b) = f(a), \text{ i.e., } M = \begin{bmatrix} 0 & 1 \\ 1 & 0 \end{bmatrix} \]

then we find that

\[ \bar{E}_\infty(f) = \max f. \]
LOWER AND UPPER MASS FUNCTIONS

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

$$\begin{bmatrix}
T\{a\} & T\{b\} & T\{c\} \\
\end{bmatrix} =
\begin{bmatrix}
\bar{q}(a|a) & \bar{q}(b|a) & \bar{q}(c|a) \\
\bar{q}(a|b) & \bar{q}(b|b) & \bar{q}(c|b) \\
\bar{q}(a|c) & \bar{q}(b|c) & \bar{q}(c|c) \\
\end{bmatrix} = \frac{1}{200}
\begin{bmatrix}
9 & 9 & 162 \\
144 & 18 & 18 \\
9 & 162 & 9 \\
\end{bmatrix}$$

$$\begin{bmatrix}
\overline{T}\{a\} & \overline{T}\{b\} & \overline{T}\{c\} \\
\end{bmatrix} =
\begin{bmatrix}
\underline{q}(a|a) & \underline{q}(b|a) & \underline{q}(c|a) \\
\underline{q}(a|b) & \underline{q}(b|b) & \underline{q}(c|b) \\
\underline{q}(a|c) & \underline{q}(b|c) & \underline{q}(c|c) \\
\end{bmatrix} = \frac{1}{200}
\begin{bmatrix}
19 & 19 & 172 \\
154 & 28 & 28 \\
19 & 172 & 19 \\
\end{bmatrix}$$
LOWER AND UPPER MASS FUNCTIONS

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

$n = 1$

$n = 2$

$n = 3$

$n = 4$

$n = 5$

$n = 6$

$n = 7$

$n = 8$

$n = 9$

$n = 10$

$n = 22$

$n = 1000$
**Theorem (De Cooman, Hermans and Quaeghebeur, 2008)**

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

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

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

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

If $x = y$, we call

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

lower and upper mean recurrence times.
Now for any trajectory \((x, x_2, x_3, \ldots)\) starting in \(x\):

\[
\tau_{xy}(x, x_2, x_3, \ldots) = \begin{cases} 
1 & ; \ x_2 = y \\
1 + \tau_{x_2y}(x_2, x_3, \ldots) & ; \ x_2 \neq y 
\end{cases}
\]

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:

\[
M. y = 1 + T[(1 - \delta_y)M. y] \quad \text{and} \quad \overline{M}. y = 1 + \overline{T}[(1 - \delta_y)\overline{M}. y].
\]
MEAN FIRST PASSAGE TIMES

EXAMPLES

We find after solving the non-linear equations that:

CONTAMINATED RANDOM WALK

\[
\begin{align*}
R_a &= R_b = M_{ab} = M_{ba} = \frac{2}{1 + \varepsilon} \\
\bar{R}_a &= \bar{R}_b = \bar{M}_{ab} = \bar{M}_{ba} = \frac{2}{1 - \varepsilon}.
\end{align*}
\]

CONTAMINATED CYCLE

\[
\begin{align*}
R_a &= R_b = 2 - \varepsilon \quad \text{and} \quad M_{ab} = M_{ba} = 1 \\
\bar{R}_a &= \bar{R}_b = \frac{2 - \varepsilon}{1 - \varepsilon} \quad \text{and} \quad \bar{M}_{ab} = \bar{M}_{ba} = \frac{1}{1 - \varepsilon}.
\end{align*}
\]
A SPECIAL CREDAL NETWORK
UNDER EPISTEMIC IRRELEVANCE

An imprecise Markov chain can also be depicted as follows:

\[ X_1 \rightarrow X_2 \rightarrow X_3 \rightarrow \cdots \rightarrow X_{n-1} \rightarrow X_n \]

INTERPRETATION OF THE GRAPH
Conditional on \( X_k \) we have that \( X_1, \ldots, X_{k-1} \) are epistemically irrelevant to \( X_{k+1}, \ldots, X_n \):

\[
\overline{E}(f(X_{k+1}, \ldots, X_n)|X_1, \ldots, X_{k-1}, X_k) = \overline{E}(f(X_{k+1}, \ldots, X_n)|X_k)
\]

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

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

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

**Figure:** $I_2$ doesn’t separate $T$ from $I_1$. 

---

### Imprecise immediate predictions
GdC,EQ,FH

### Imprecise probabilities

### Precise probability trees

### Huygens’s tree

### Imprecise probability trees

### Imprecise Markov chains

### Perron–Frobenius Theorem

### First passage

### Towards credal nets

### Literature
A special case

Hidden Markov chains

\[ X_1 \xrightarrow{} X_2 \xrightarrow{} X_3 \xrightarrow{} \ldots \xrightarrow{} X_{n-1} \xrightarrow{} X_n \]

\[ O_1 \xrightarrow{} O_2 \xrightarrow{} O_3 \xrightarrow{} \ldots \xrightarrow{} O_{n-1} \xrightarrow{} O_n \]

