A tutorial on

Imprecise Markov chains

by Jasper De Bock & Thomas Krak

SMPS/BELIEF 2018

September 17

now :-)

GHENT UNIVERSITY

Jasper De Bock & Thomas Krak

GHENT
UNIVERSITY

GHENT
UNIVERSITY

A tutorial on

Imprecise Markov chains

by Jasper De Bock & Thomas Krak

2018

September 17

now :-)

(Walley 1991)
(Augustin et al. 2014)

A tutorial on

Imprecise Markov chains

by Jasper De Bock & Thomas Krak

?

2018

September 17

now :-)

$$P \in \mathbb{P} \rightarrow \begin{cases} \overline{E}(f) = \sup_{P \in \mathbb{P}} E(f) = -\underline{E}(-f) \\ \underline{E}(f) = \inf_{P \in \mathbb{P}} E(f) \end{cases}$$

A tutorial on

Imprecise Markov chains

by Jasper De Bock & Thomas Krak

2018

September 17

now :-)

GHENT UNIVERSIT

A tutorial on

Imprecise Markov chains

by Jasper De Bock & Thomas Krak

SMPS/BELIEF 2018

September 17

now :-)

GHENT UNIVERSITY

A tutorial on

Imprecise Markov chains

by Jasper De Bock & Thomas Krak

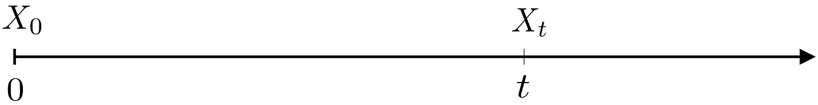
SMPS/BELIEF 2018

September 17

now :-)

GHENT UNIVERSITY

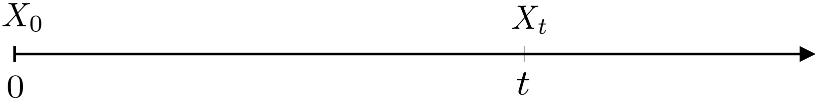
stochastic process



$\mathbb{R}_{\geq 0}$

$\left\{ \begin{matrix} \textbf{continuous} \\ \textbf{discrete} \end{matrix} \right\}$ **-time stochastic process**

$\mathbb{N}$



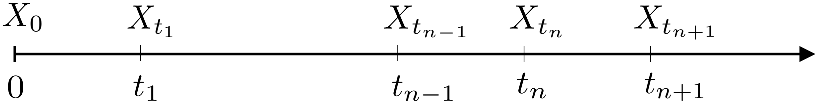
$\mathbb{R}_{\geq 0}$

$$\left\{ \begin{array}{c} \textbf{continuous} \\ \textbf{discrete} \end{array} \right\}\textbf{-time stochastic process}$$

$\mathbb{N}$

$$P(X_0 = x_0)$$

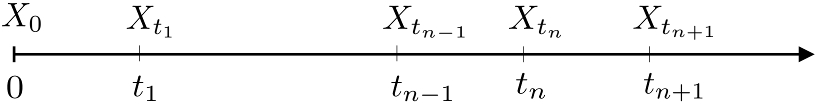
$$P(X_{t_{n+1}} = y \mid X_{t_1} = x_{t_1}, ..., X_{t_{n-1}} = x_{t_{n-1}}, X_{t_n} = x)$$



$$\overset{\mathbb{R}_{\geq 0}}{\underset{\mathbb{N}}{\left\{\begin{matrix}\textbf{continuous}\\ \textbf{discrete}\end{matrix}\right\}}}\textbf{-time Markov chain}$$

$$P(X_0 = x_0)$$

$$P(X_{t_{n+1}} = y \mid X_{t_1} = x_{t_1}, \ldots, X_{t_{n-1}} = x_{t_{n-1}}, X_{t_n} = x)$$
$$= P(X_{t_{n+1}} = y \mid X_{t_1} = X_{t_n} = x)$$



$$\left\{\begin{array}{c} \overset{\mathbb{R}_{\geq 0}}{\textbf{continuous}} \\ \underset{\mathbb{N}}{\textbf{discrete}} \end{array}\right\}\textbf{-time} \overset{\textbf{homogeneous}}{\vee} \textbf{Markov chain}$$

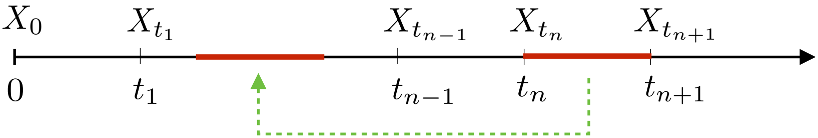
$$P(X_0 = x_0)$$

$$P(X_{t_{n+1}} = y \mid X_{t_1} = x_{t_1}, \ldots, X_{t_{n-1}} = x_{t_{n-1}}, X_{t_n} = x)$$

$$= P(X_{t_{n+1}} = y \mid X_{t_1} = X_{t_n} = x)$$

$$= T_\Delta(x, y)$$

only the time difference $\Delta = t_{n+1} - t_n$ matters!



$$\begin{array}{c}\mathbb{R}_{\geq 0}\\ \left\{\begin{array}{c}\textbf{continuous}\\ \textbf{discrete}\end{array}\right\}\\ \mathbb{N}\end{array}$$

-time homogeneous Markov chain

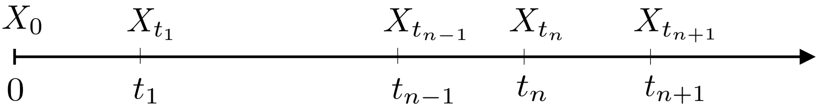
$$\boxed{P(X_0 = x_0)}$$

$$P(X_{t_{n+1}} = y\,|X_{t_1} = x_{t_1}, ..., X_{t_{n-1}} = x_{t_{n-1}}, X_{t_n} = x)$$

$$= P(X_{t_{n+1}} = y\,|X_{t_1} = X_{t_n} = x)$$

$$= \boxed{T_\Delta}(x, y)$$

only the time difference
$\Delta = t_{n+1} - t_n$ **matters!**



$\mathbb{R}_{\geq 0}$

continuous
discrete

$\mathbb{N}$

homogeneous

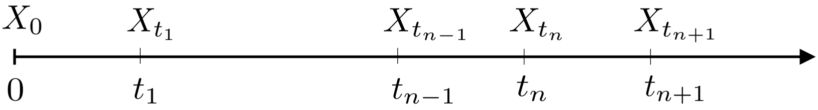
-time Markov chain

$P(X_0 = x_0)$

initial distribution

that's just a probability mass function $\pi_0(x_0)$

$\boxed{T_\Delta}$



$\mathbb{R}_{\geq 0}$

$\left\{ \begin{array}{c} \text{continuous} \\ \text{discrete} \end{array} \right\}$-time Markov chain

$\mathbb{N}$

homogeneous

$P(X_0 = x_0)$

initial distribution

that's just a probability mass function $\pi_0(x_0)$

$\sum_y T(x, y) = 1$

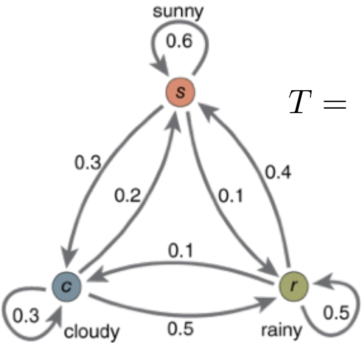
$(\forall y)\ T(x, y) \geq 0$

transition matrix

$T_\Delta = T^\Delta$ with $T := T_1$

$$X_0 \quad X_{t_1} \quad X_{t_{n-1}} \quad X_{t_n} \quad X_{t_{n+1}}$$

$$0 \quad t_1 \quad t_{n-1} \quad t_n \quad t_{n+1}$$

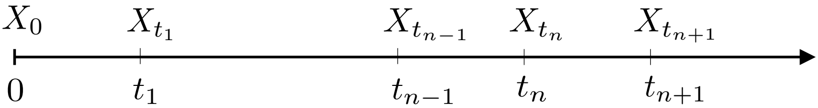


$$T = \begin{array}{c} \\ s \\ c \\ r \end{array} \begin{array}{c} \begin{array}{ccc} s & c & r \end{array} \\ \begin{bmatrix} 0.6 & 0.3 & 0.1 \\ 0.2 & 0.3 & 0.5 \\ 0.4 & 0.1 & 0.5 \end{bmatrix} \end{array}$$

$$\sum_y T(x,y) = 1$$
$$(\forall y)\ T(x,y) \geq 0$$

transition matrix

$$T_\Delta = T^\Delta \quad \text{with} \quad T := T_1$$



$\mathbb{R}_{\geq 0}$

continuous

discrete

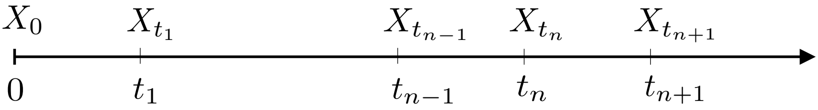
$\mathbb{N}$

-time homogeneous Markov chain

$P(X_0 = x_0)$

initial distribution

that's just a probability mass function $\pi_0(x_0)$



$$\left\{\begin{array}{c} \overset{\mathbb{R}_{\geq 0}}{\textbf{continuous}} \\ \underset{\mathbb{N}}{\textbf{discrete}} \end{array}\right\}\textbf{-time } \textbf{homogeneous } \textbf{Markov chain}$$

$P(X_0 = x_0)$ that's just a probability mass function $\pi_0(x_0)$

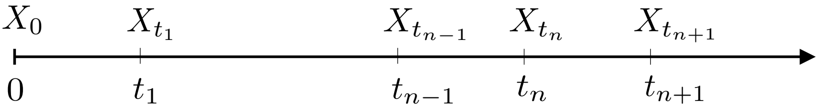
initial distribution

$\sum_y Q(x, y) = 0$

$(\forall y \neq x)\, Q(x, y) \geq 0$

transition rate matrix

$$T_\Delta = e^{Q\Delta} := \lim_{n\to\infty} \left(I + \frac{t}{n}Q\right)^n \quad \text{with} \quad Q := \lim_{\Delta\to 0} \frac{T_\Delta - I}{\Delta}$$

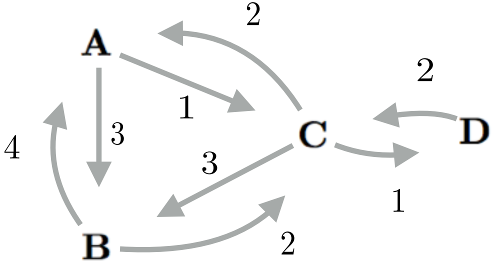


Amorous Bickering Confusion Depression

$$
Q = \begin{array}{c} \\ \mathbf{A} \\ \mathbf{B} \\ \mathbf{C} \\ \mathbf{D} \end{array}
\begin{array}{c}
\begin{array}{cccc} \mathbf{A} & \mathbf{B} & \mathbf{C} & \mathbf{D} \end{array} \\
\begin{bmatrix} -4 & 3 & 1 & 0 \\ 4 & -6 & 2 & 0 \\ 2 & 3 & -6 & 1 \\ 0 & 0 & 2 & -2 \end{bmatrix}
\end{array}
$$

$$\sum_y Q(x,y) = 0$$

$$(\forall y \neq x)\, Q(x,y) \geq 0$$



transition rate matrix

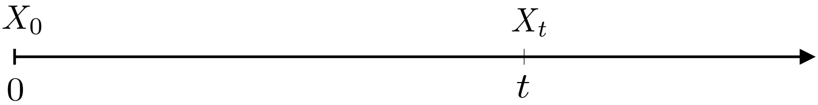
$$Q := \lim_{\Delta \to 0} \frac{T_\Delta - I}{\Delta}$$

$\left\{\begin{matrix}\overset{\mathbb{R}_{\geq 0}}{\textbf{continuous}}\\ \underset{\mathbb{N}}{\textbf{discrete}}\end{matrix}\right\}$**-time** $\overset{\textbf{homogeneous}}{\curlyvee}$ **Markov chain**

initial distribution π_0

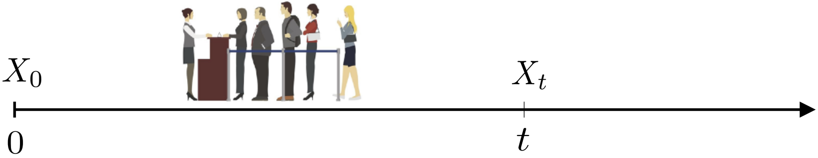
+

$\left\{\begin{matrix}\textbf{transition rate matrix } Q\\ \textbf{or}\\ \textbf{transition matrix } T\end{matrix}\right.$



$$
\begin{aligned}
E(f(X_t)|X_0 = x) &= [T_t f](x) \\
&= \sum_y T_t(x, y) f(y) = \begin{cases} \sum_y e^{Qt}(x, y) f(y) \\ \sum_y T^t(x, y) f(y) \end{cases}
\end{aligned}
$$

$$
\begin{aligned}
x &= 0 \\
f(X_t) &= X_t
\end{aligned}
$$



$$
\begin{aligned}
E(f(X_t)|X_0 = x) &= [T_t f](x) \\
&= \sum_y T_t(x,y) f(y) = \begin{cases} \sum_y e^{Qt}(x,y) f(y) \\ \sum_y T^t(x,y) f(y) \end{cases}
\end{aligned}
$$

$$
P(X_t = y | X_0 = x) = E(\mathbb{I}_y(X_t)|X_0 = x) = [T_t \mathbb{I}_y](x)
$$

$$
y = x = 0
$$

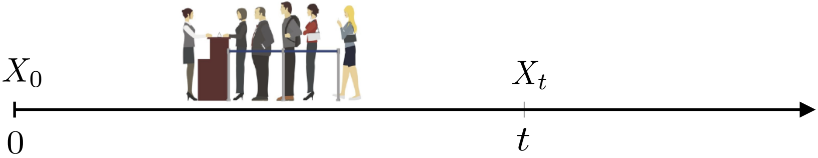
$$
f(X_t) = \mathbb{I}_y(X_t) = \begin{cases} 1 & \text{if } X_t = y \\ 0 & \text{otherwise} \end{cases}
$$

$$E(f(X_t)|X_0 = x) = [T_t f](x)$$

$$\boxed{E_\infty(f) := \lim_{t\to+\infty} E(f(X_t)|X_0 = x)}$$

$$P(X_t = y|X_0 = x) = E(\mathbb{I}_y(X_t)|X_0 = x) = [T_t \mathbb{I}_y](x)$$

$$\boxed{\pi_\infty(y) := \lim_{t\to+\infty} P(X_t = y|X_0 = x)}$$



Reliability engineering (failure probabilities, …)

Queuing theory (waiting in line …)

- optimising supermarket waiting times
- dimensioning of call centers
- airport security lines
- router queues on the internet

Chemical reactions (time-evolution …)

Pagerank

…

So how about imprecision?

So how about imprecision?

What if we don’t know T or Q exactly?

Sets of transition (rate) matrices

Don't know T (or Q) exactly

But confident that $T \in \mathscr{T}$ for some *set* $\mathscr{T}$ of transition matrices

- (or that $Q \in \mathscr{Q}$ for some *set* $\mathscr{Q}$ of rate matrices)

Induces *imprecise Markov chain*; *set* of processes *compatible* with $\mathscr{T}$.

Sets of transition (rate) matrices

Don't know T (or Q) exactly

But confident that $T \in \mathscr{T}$ for some *set* $\mathscr{T}$ of transition matrices

- (or that $Q \in \mathscr{Q}$ for some *set* $\mathscr{Q}$ of rate matrices)

Induces *imprecise Markov chain*; *set* of processes *compatible* with $\mathscr{T}$.

Different versions:

- $\mathbb{P}^{\mathrm{HM}}_{\mathscr{T}}$: all homogeneous Markov chains with $T \in \mathscr{T}$

Sets of transition (rate) matrices

Don't know T (or Q) exactly

But confident that $T \in \mathscr{T}$ for some *set* $\mathscr{T}$ of transition matrices

- (or that $Q \in \mathscr{Q}$ for some *set* $\mathscr{Q}$ of rate matrices)

Induces *imprecise Markov chain*; *set* of processes *compatible* with $\mathscr{T}$.

Different versions:

- $\mathbb{P}^{\mathrm{HM}}_{\mathscr{T}}$: all homogeneous Markov chains with $T \in \mathscr{T}$
- $\mathbb{P}^{\mathrm{M}}_{\mathscr{T}}$: all (**non**-homogeneous) Markov chains with $T^{(t)} \in \mathscr{T}$

Sets of transition (rate) matrices

Don't know T (or Q) exactly

But confident that $T \in \mathscr{T}$ for some *set* $\mathscr{T}$ of transition matrices

- (or that $Q \in \mathscr{Q}$ for some *set* $\mathscr{Q}$ of rate matrices)

Induces *imprecise Markov chain*; *set* of processes *compatible* with $\mathscr{T}$.

Different versions:

- $\mathbb{P}^{\mathrm{HM}}_{\mathscr{T}}$: all homogeneous Markov chains with $T \in \mathscr{T}$
- $\mathbb{P}^{\mathrm{M}}_{\mathscr{T}}$: all (**non**-homogeneous) Markov chains with $T^{(t)} \in \mathscr{T}$
- $\mathbb{P}_{\mathscr{T}}$: all (**non**-Markov) processes with $T^{(t,x_{\mathbf{u}})} \in \mathscr{T}$

Sets of transition (rate) matrices

Don't know T (or Q) exactly

But confident that $T \in \mathscr{T}$ for some *set* $\mathscr{T}$ of transition matrices

- (or that $Q \in \mathscr{Q}$ for some *set* $\mathscr{Q}$ of rate matrices)

Induces *imprecise Markov chain*; *set* of processes *compatible* with $\mathscr{T}$.

Different versions:

- $\mathbb{P}^{\mathrm{HM}}_{\mathscr{T}}$: all homogeneous Markov chains with $T \in \mathscr{T}$
- $\mathbb{P}^{\mathrm{M}}_{\mathscr{T}}$: all (**non**-homogeneous) Markov chains with $T^{(t)} \in \mathscr{T}$
- $\mathbb{P}_{\mathscr{T}}$: all (**non**-Markov) processes with $T^{(t,x_{\mathbf{u}})} \in \mathscr{T}$

Clearly

$$\mathbb{P}^{\mathrm{HM}}_{\mathscr{T}} \subseteq \mathbb{P}^{\mathrm{M}}_{\mathscr{T}} \subseteq \mathbb{P}_{\mathscr{T}}$$

Lower expectations and lower probabilities

Given an imprecise Markov chain $\mathbb{P}^*_{\mathscr{T}}$, we are interested in

$$\underline{\mathbb{E}}^*_{\mathscr{T}}\bigl[f(X_t)\,|\,X_0 = x\bigr] = \inf_{P \in \mathbb{P}^*_{\mathscr{T}}} \mathbb{E}_P\bigl[f(X_t)\,|\,X_0 = x\bigr]$$

(And $\overline{\mathbb{E}}^*_{\mathscr{T}}\bigl[f(X_t)\,|\,X_0 = x\bigr]$ by conjugacy)

Lower- (and upper) probabilities a special case:

$$\underline{P}^*_{\mathscr{T}}\bigl(X_t = y\,|\,X_0 = x\bigr) = \inf_{P \in \mathbb{P}^*_{\mathscr{T}}} P\bigl(X_t = y\,|\,X_0 = x\bigr) = \underline{\mathbb{E}}^*_{\mathscr{T}}\bigl[\mathbb{I}_y(X_t)\,|\,X_0 = x\bigr]$$

Because different types are nested,

$$\underline{\mathbb{E}}_{\mathscr{T}}\bigl[f(X_t)\,|\,X_0 = x\bigr] \leq \underline{\mathbb{E}}^{\mathrm{M}}_{\mathscr{T}}\bigl[f(X_t)\,|\,X_0 = x\bigr] \leq \underline{\mathbb{E}}^{\mathrm{HM}}_{\mathscr{T}}\bigl[f(X_t)\,|\,X_0 = x\bigr]$$

Computing lower expectations, first try

Recall that for a homogeneous Markov chain P with transition matrix T,

$$\mathbb{E}_P\big[f(X_1)\,|\,X_0=x\big] = \big[Tf\big](x).$$

Now consider $\mathbb{P}^{\mathrm{HM}}_{\mathscr{T}}$. Then,

$$\begin{aligned}\underline{\mathbb{E}}^{\mathrm{HM}}_{\mathscr{T}}\big[f(X_1)\,|\,X_0=x\big] &:= \inf_{P\in\mathbb{P}^{\mathrm{HM}}_{\mathscr{T}}} \mathbb{E}_P\big[f(X_1)\,|\,X_0=x\big] \\ &= \inf_{T\in\mathscr{T}} \big[Tf\big](x)\end{aligned}$$

- Linear optimisation problem with constraints given by $\mathscr{T}$
- Relatively straightforward if $\mathscr{T}$ is "nice"
- Essentially solving a linear programming problem

Computing lower expectations, first try

Recall that for a homogeneous Markov chain P with transition matrix T,

$$\mathbb{E}_P\big[f(X_t)\,|\,X_0 = x\big] = \big[T^t f\big](x)\,.$$

Now consider $\mathbb{P}^{\mathrm{HM}}_{\mathscr{T}}$. Then,

$$\begin{aligned}\underline{\mathbb{E}}^{\mathrm{HM}}_{\mathscr{T}}\big[f(X_t n)\,|\,X_0 = x\big] &:= \inf_{P \in \mathbb{P}^{\mathrm{HM}}_{\mathscr{T}}} \mathbb{E}_P\big[f(X_t)\,|\,X_0 = x\big] \\ &= \inf_{T \in \mathscr{T}} \big[T^t f\big](x)\end{aligned}$$

- **Non**-linear optimisation problem with constraints given by $\mathscr{T}$
- **Not** straightforward even if $\mathscr{T}$ is “nice”

Computing lower expectations, first try

Recall that for a homogeneous Markov chain P with transition matrix T,

$$\mathbb{E}_P\big[f(X_t)\,|\,X_0 = x\big] = \big[T^t f\big](x)\,.$$

Now consider $\mathbb{P}^{\mathrm{HM}}_{\mathscr{T}}$. Then,

$$\begin{aligned}\underline{\mathbb{E}}^{\mathrm{HM}}_{\mathscr{T}}\big[f(X_t n)\,|\,X_0 = x\big] &:= \inf_{P\in\mathbb{P}^{\mathrm{HM}}_{\mathscr{T}}} \mathbb{E}_P\big[f(X_t)\,|\,X_0 = x\big] \\ &= \inf_{T\in\mathscr{T}} \big[T^t f\big](x)\end{aligned}$$

- **Non**-linear optimisation problem with constraints given by $\mathscr{T}$
- **Not** straightforward even if $\mathscr{T}$ is “nice”

See e.g. (Kozine and Utkin, 2002) and (Campos *et al.*, 2003) for analyses of this approach.

What about the non-Markov case?

$\mathbb{P}_{\mathscr{T}}$: all (**non**-Markov) processes with $T^{(t,x_{\mathbf{u}})} \in \mathscr{T}$

How to interpret this?
$\Rightarrow$ Helps to draw a picture

What about the non-Markov case?

$\mathbb{P}_{\mathscr{T}}$: all (**non**-Markov) processes with $T^{(t,x_{\mathbf{u}})} \in \mathscr{T}$

How to interpret this?
$\Rightarrow$ Helps to draw a picture

Example with binary state space $\mathscr{X} = \{a, b\}$

Use *event tree* / *probability tree*
Illustration of behaviour *over time*

Need notation

$$\mathscr{T}_x := \left\{ T(x,\cdot) \,\middle|\, T \in \mathscr{T} \right\} \quad \forall x \in \mathscr{X}$$

What about the non-Markov case?

$\mathbb{P}_{\mathscr{T}}$: all (**non**-Markov) processes with $T^{(t,x_{\mathbf{u}})} \in \mathscr{T}$

How to interpret this?
$\Rightarrow$ Helps to draw a picture

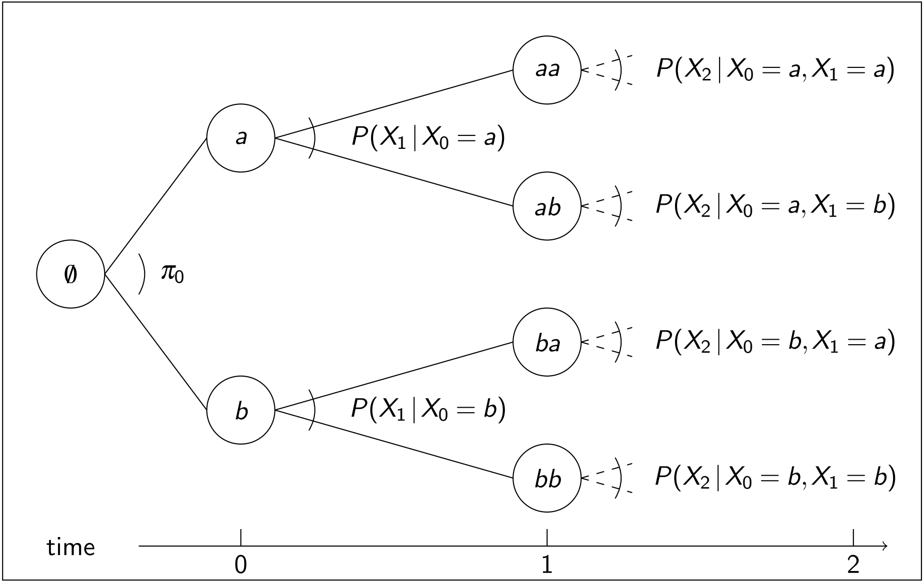
Example with binary state space $\mathscr{X} = \{a, b\}$

Use *event tree* / *probability tree*
Illustration of behaviour *over time*

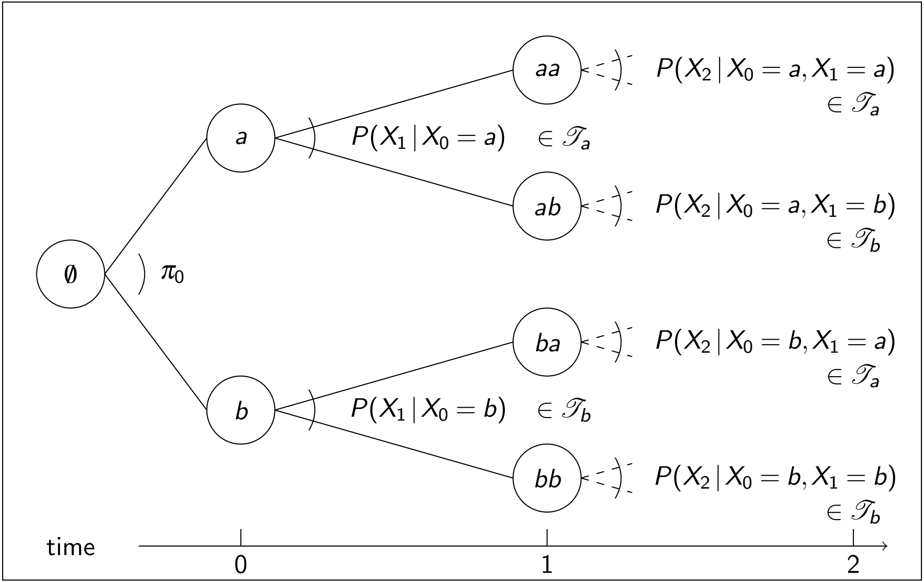
This setting explored by (De Cooman *et al.*, 2009).

Tree representation related to (Shafer and Vovk, 2001) game-theoretic probabilities. Connection to (Walley's) imprecise probabilities in (De Cooman and Hermans, 2008).

Visualising a stochastic process



Visualising a stochastic process in $\mathbb{P}_{\mathscr{T}}$



Computations by iterated lower expectation

For the set $\mathbb{P}_{\mathscr{T}}$, it can be shown that

$$\underline{\mathbb{E}}_{\mathscr{T}}\left[f(X_t)\,\middle|\,X_0=x\right]=\underline{\mathbb{E}}_{\mathscr{T}}\left[\underline{\mathbb{E}}_{\mathscr{T}}\left[f(X_t)\,|\,X_0=x,X_s\right]\,\middle|\,X_0=x\right]\qquad\forall s\leq t$$

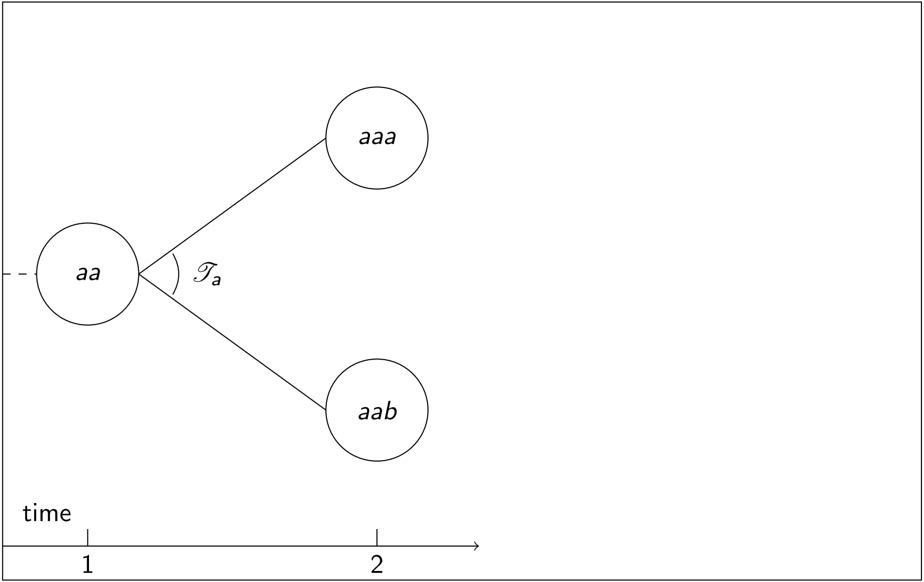
This is the *law of iterated lower expectation*.

Provides *backwards recursive* scheme for computations.
$\Rightarrow$ Intuitive in the tree representation

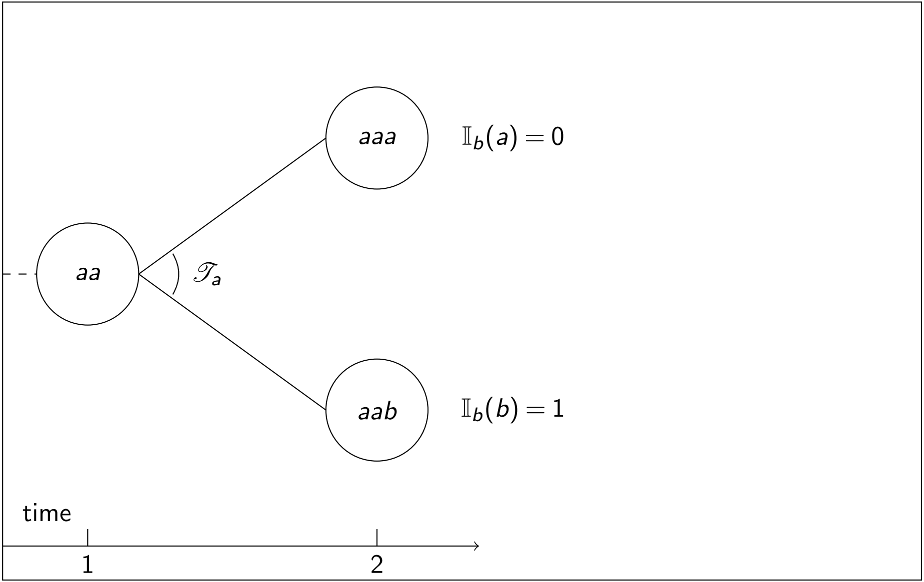
Example: compute $\underline{\mathbb{E}}_{\mathscr{T}}\left[\mathbb{I}_b(X_2)\,\middle|\,X_0=a\right]=\underline{P}\left(X_2=b\,|\,X_0=a\right)$

$$\mathscr{T}_a:=\left\{T(a,\cdot)\,\middle|\,T(a,a)\in[0.4,0.6]\right\}$$
$$\mathscr{T}_b:=\left\{T(b,\cdot)\,\middle|\,T(b,a)\in[0.1,0.3]\right\}$$

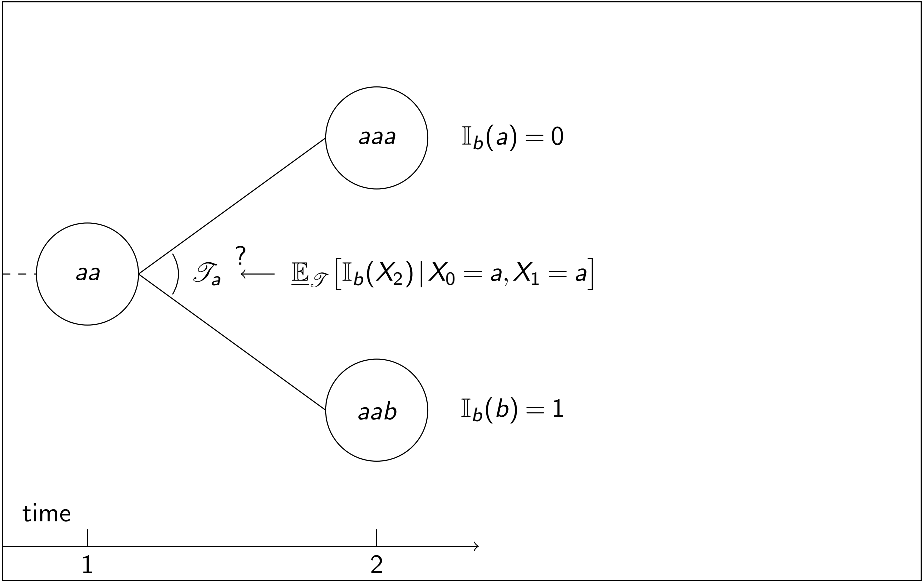
Example $\mathbb{E}_{\mathscr{T}}\left[\mathbb{I}_b(X_2)\,\middle|\,X_0 = a\right]$ base case



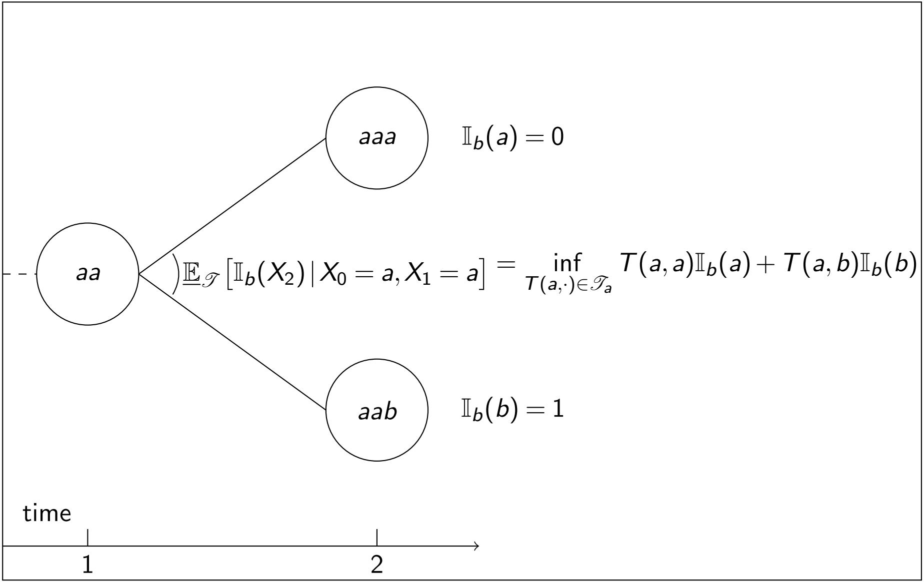
Example $\mathbb{E}_{\mathscr{T}}\left[\mathbb{I}_b(X_2) \,\middle|\, X_0 = a\right]$ base case



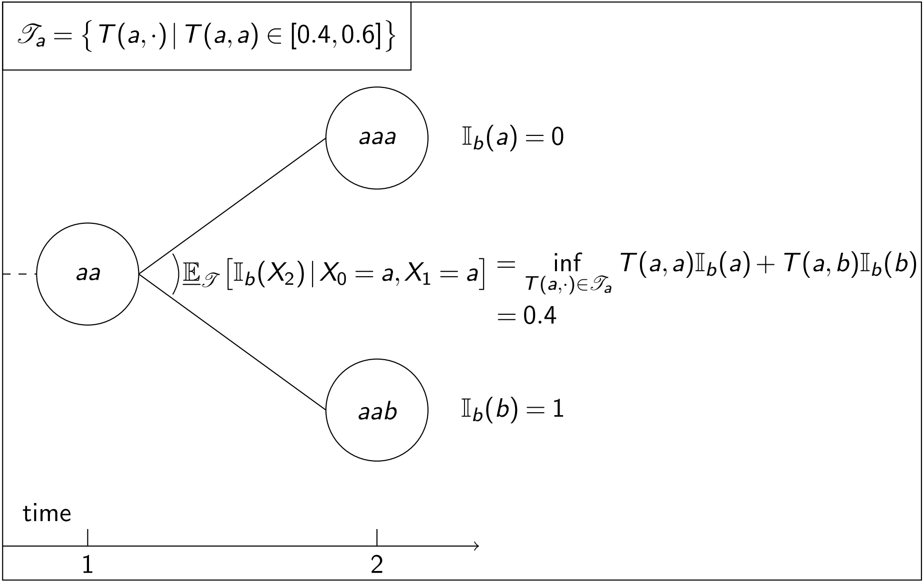
Example $\underline{\mathbb{E}}_{\mathscr{T}}\left[\mathbb{I}_b(X_2)\,\middle|\,X_0 = a\right]$ base case



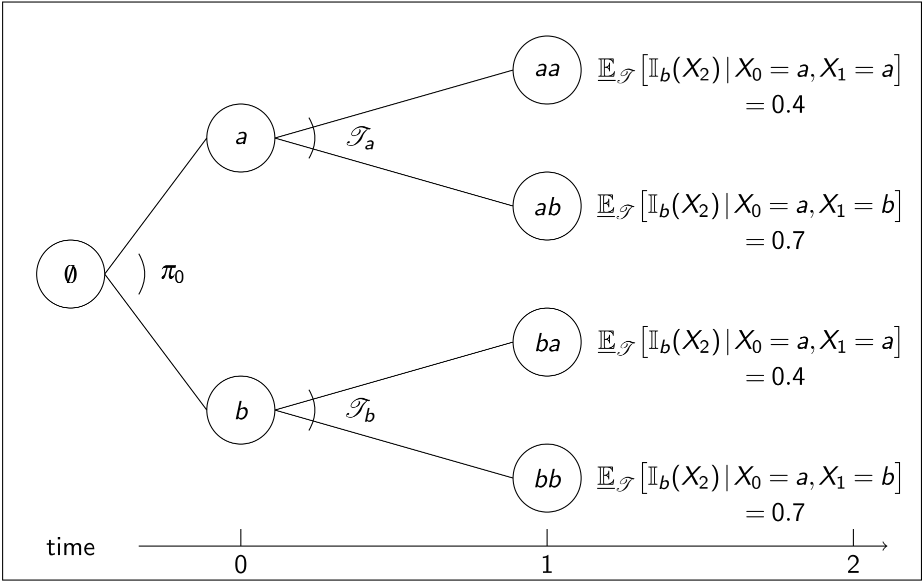
Example $\mathbb{E}_{\mathscr{T}}\left[\mathbb{I}_b(X_2)\,\middle|\,X_0 = a\right]$ base case



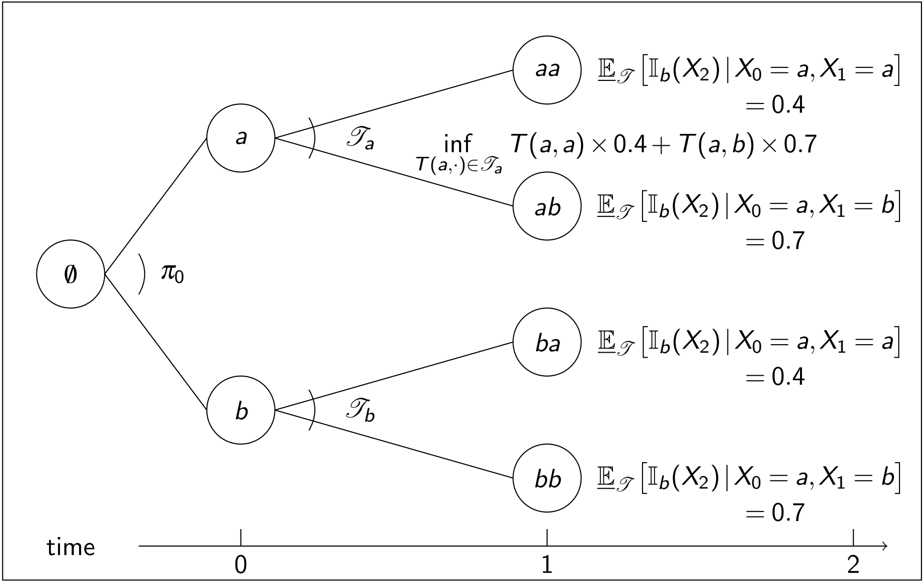
Example $\mathbb{E}_{\mathscr{T}}\left[\mathbb{I}_b(X_2)\,\middle|\,X_0 = a\right]$ base case



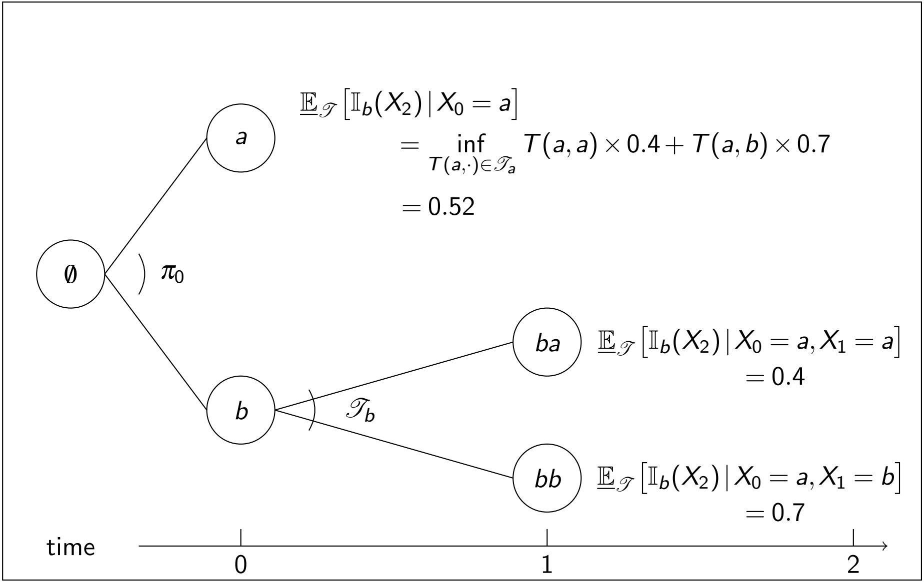
Recursive, local computations



Recursive, local computations



Recursive, local computations



Local computations in operator form

Consider $\mathscr{T}_x$, and define for any $f : \mathscr{X} \to \mathbb{R}$,

$$\left[\underline{T}f\right](x) := \inf_{T(x,\cdot)\in\mathscr{T}_x} \sum_y T(x,y)f(y)$$

Linear optimisation problem, and

$$\left[\underline{T}f\right](x) = \inf_{T\in\mathscr{T}} \left[Tf\right](x)$$

We call $\underline{T}$ the *lower transition operator* for $\mathscr{T}$.
We can write

$$\underline{\mathbb{E}}_{\mathscr{T}}\left[f(X_{t+1})\,|\,X_{0:t} = x_{0:t}\right] = \left[\underline{T}f\right](x_t)$$

Local computations in operator form

Consider $\mathscr{T}_x$, and define for any $f : \mathscr{X} \to \mathbb{R}$,

$$\big[\underline{T}f\big](x) := \inf_{T(x,\cdot)\in\mathscr{T}_x} \sum_y T(x,y)f(y)$$

Linear optimisation problem, and

$$\big[\underline{T}f\big](x) = \inf_{T\in\mathscr{T}} \big[Tf\big](x)$$

We call $\underline{T}$ the *lower transition operator* for $\mathscr{T}$.
We can write

$$\underline{\mathbb{E}}_{\mathscr{T}}\big[f(X_{t+1})\,|\,X_{0:t}=x_{0:t}\big] = \big[\underline{T}f\big](x_t)$$

We find

$$\underline{\mathbb{E}}_{\mathscr{T}}\big[f(X_{t+1})\,|\,X_{0:t}=x_{0:t}\big] = \big[\underline{T}f\big](x_t) = \underline{\mathbb{E}}_{\mathscr{T}}\big[f(X_{t+1})\,|\,X_t=x_t\big]$$

Lower envelope for imprecise Markov chain $\mathbb{P}_{\mathscr{T}}$ has “Markov” property

- But contains **non**-Markov models!

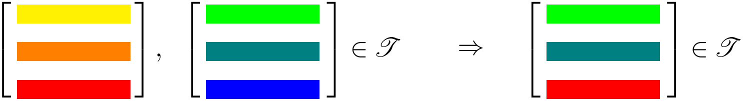
Similarly the lower envelope is also homogeneous!

Multiple time steps

By repeating the local computations,

$$\underline{\mathbb{E}}_{\mathscr{T}}\left[f(X_2)\,|\,X_0 = x\right] = \left[\underline{T}\,\underline{T}f\right](x),$$

if the set $\mathscr{T}$ has *separately specified rows*:



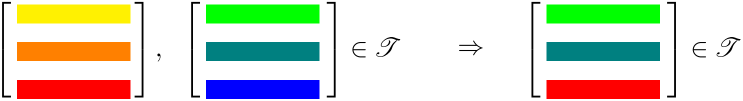
($\mathscr{T}$ is closed under recombination of rows)

Multiple time steps

By repeating the local computations,

$$\underline{\mathbb{E}}_{\mathscr{T}}\left[f(X_2)\,|\,X_0 = x\right] = \left[\underline{T}\,\underline{T}f\right](x)\,,$$

if the set $\mathscr{T}$ has *separately specified rows*:



($\mathscr{T}$ is closed under recombination of rows)

By induction we get

$$\underline{\mathbb{E}}_{\mathscr{T}}\left[f(X_t)\,|\,X_0 = x\right] = \left[\underline{T}^t f\right](x)$$

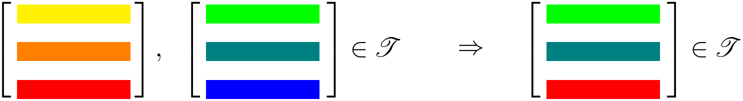
- *Local*, *linear* optimisations only
- Can be efficiently computed

Multiple time steps

By repeating the local computations,

$$\underline{\mathbb{E}}_{\mathscr{T}}\left[f(X_2)\,|\,X_0 = x\right] = \left[\underline{T}\,\underline{T}f\right](x)\,,$$

if the set $\mathscr{T}$ has *separately specified rows*:



($\mathscr{T}$ is closed under recombination of rows)

By induction we get

$$\underline{\mathbb{E}}_{\mathscr{T}}\left[f(X_t)\,|\,X_0 = x\right] = \left[\underline{T}^t f\right](x)$$

Imprecise Markov chain $\mathbb{P}_{\mathscr{T}}$ can be seen as *credal network* under *epistemic irrelevance*. Gives a graphical model representation.

"Separately specified rows" is a well-known condition in that context.

That's two extremes. What about the intermediate one?

So far ignored $\mathbb{P}^{\mathrm{M}}_{\mathscr{T}}$

That's two extremes. What about the intermediate one?

So far ignored $\mathbb{P}^{\mathrm{M}}_{\mathscr{T}}$

Turns out that if $\mathscr{T}$ has separately specified rows, then

$$\underline{\mathbb{E}}^{\mathrm{M}}_{\mathscr{T}}\big[f(X_t)\,|\,X_0=x\big] = \big[\underline{T}^t f\big](x)$$

If follows that

$$\underline{\mathbb{E}}^{\mathrm{M}}_{\mathscr{T}}\big[f(X_t)\,|\,X_0=x\big] = \underline{\mathbb{E}}_{\mathscr{T}}\big[f(X_t)\,|\,X_0=x\big]$$

- Does **not** hold for functions on multiple time points
- Then only $\mathbb{P}_{\mathscr{T}}$ remains tractable

That's two extremes. What about the intermediate one?

So far ignored $\mathbb{P}^{\mathrm{M}}_{\mathscr{T}}$

Turns out that if $\mathscr{T}$ has separately specified rows, then

$$\underline{\mathbb{E}}^{\mathrm{M}}_{\mathscr{T}}\left[f(X_t)\,|\,X_0 = x\right] = \left[\underline{T}^t f\right](x)$$

If follows that

$$\underline{\mathbb{E}}^{\mathrm{M}}_{\mathscr{T}}\left[f(X_t)\,|\,X_0 = x\right] = \underline{\mathbb{E}}_{\mathscr{T}}\left[f(X_t)\,|\,X_0 = x\right]$$

- Does **not** hold for functions on multiple time points
- Then only $\mathbb{P}_{\mathscr{T}}$ remains tractable

First pioneered by Hartfiel, *Markov Set-Chains* (Hartfiel, 1998)
$\Rightarrow$ No explicit connection to imprecise probabilities

Exploration with imprecise probabilities by (Škulj, 2009)

Limit behaviour?

Limit inference often of interest:

$$\mathbb{E}\big[f(X_\infty)\,|\,X_0 = x\big] = \lim_{t\to+\infty} \mathbb{E}\big[f(X_t)\,|\,X_0 = x\big]$$

In imprecise setting, often exists:

$$\underline{\mathbb{E}}_{\mathscr{T}}\big[f(X_\infty)\,|\,X_0 = x\big] := \lim_{t\to+\infty} \big[\underline{T}^t f\big](x),$$

and often independent of x.

See e.g. (De Cooman *et al.*, 2009) and (Škulj, 2009)

Summary for imprecise Markov chains in discrete time

Parameterisation through *set* $\mathscr{T}$ of transition matrices.

Can induce **three** different *imprecise Markov chains*:

- $\mathbb{P}^{\mathrm{HM}}_{\mathscr{T}}$: all homogeneous Markov chains compatible with $\mathscr{T}$
- $\mathbb{P}^{\mathrm{M}}_{\mathscr{T}}$: all (**non**-homogeneous) Markov chains compatible with $\mathscr{T}$
- $\mathbb{P}_{\mathscr{T}}$: all (**non**-Markov) processes compatible with $\mathscr{T}$

For $\mathbb{P}^{\mathrm{HM}}_{\mathscr{T}}$, computations are difficult.

For $\mathbb{P}^{\mathrm{M}}_{\mathscr{T}}$ and $\mathbb{P}_{\mathscr{T}}$, computations using *lower transition operator*

$$\underline{\mathbb{E}}^{\mathrm{M}}_{\mathscr{T}}\left[f(X_t)\,|\,X_0=x\right]=\underline{\mathbb{E}}_{\mathscr{T}}\left[f(X_t)\,|\,X_0=x\right]=\left[\underline{T}^t f\right](x)$$

The imprecise Markov chain $\mathbb{P}_{\mathscr{T}}$ satisfies an *imprecise Markov property*

The limit $\lim_{t\to+\infty}\left[\underline{T}^t f\right](x)$ often exists, and often independent of x.

Imprecise Continuous-Time Markov Chains

Going to go a bit faster with more intuition

We use the same basic approach:

- Uncertain about Q, but consider a *set* $\mathscr{Q}$
- Three imprecise (continuous-time) Markov chains, *compatible* with $\mathscr{Q}$:
  - $\mathbb{P}_{\mathscr{Q}}^{\mathrm{HM}}$: all homogeneous Markov chains with $Q \in \mathscr{Q}$
  - $\mathbb{P}_{\mathscr{Q}}^{\mathrm{M}}$: all (**non**-homogeneous) Markov chains with $Q_t \in \mathscr{Q}$
  - $\mathbb{P}_{\mathscr{Q}}$: all (**non**-Markov) processes with $Q_{t,x_{\mathbf{u}}} \in \mathscr{Q}$

Similar to discrete-time case,

$$\underline{\mathbb{E}}_{\mathscr{Q}}^{\mathrm{HM}}\left[f(X_t)\,\middle|\,X_0 = x\right] = \inf_{Q\in\mathscr{Q}}\left[e^{Qt}f\right](x)$$

which is difficult due to nonlinearities in the optimisation.

Imprecise Continuous-Time Markov Chains

Going to go a bit faster with more intuition

We use the same basic approach:

- Uncertain about Q, but consider a *set* $\mathscr{Q}$
- Three imprecise (continuous-time) Markov chains, *compatible* with $\mathscr{Q}$:
  - $\mathbb{P}^{\mathrm{HM}}_{\mathscr{Q}}$: all homogeneous Markov chains with $Q \in \mathscr{Q}$
  - $\mathbb{P}^{\mathrm{M}}_{\mathscr{Q}}$: all (**non**-homogeneous) Markov chains with $Q_t \in \mathscr{Q}$
  - $\mathbb{P}_{\mathscr{Q}}$: all (**non**-Markov) processes with $Q_{t,x_{\mathbf{u}}} \in \mathscr{Q}$

Similar to discrete-time case,

$$\underline{\mathbb{E}}^{\mathrm{HM}}_{\mathscr{Q}}\big[f(X_t)\,\big|\,X_0 = x\big] = \inf_{Q\in\mathscr{Q}}\big[e^{Qt}f\big](x)$$

which is difficult due to nonlinearities in the optimisation.

See e.g. (Goldsztejn and Neumaier, 2014) and (Oppenheimer and Michel, 1988) for details on this homogeneous setting

Non-homogeneous case in continuous-time

$\mathbb{P}^{\mathrm{M}}_{\mathscr{Q}}$: all (**non**-homogeneous) Markov chains with $Q_t \in \mathscr{Q}$

How to interpret this?

Homogeneous case, rate matrix is just a derivative,

$$Q := \lim_{\Delta \to 0} \frac{T_\Delta - I}{\Delta} \qquad \text{where} \qquad T_\Delta(x, y) := P(X_\Delta = y \,|\, X_0 = x)$$

Non-homogeneous case in continuous-time

$\mathbb{P}^{\mathrm{M}}_{\mathscr{Q}}$: all (**non**-homogeneous) Markov chains with $Q_t \in \mathscr{Q}$

How to interpret this?

Homogeneous case, rate matrix is just a derivative,

$$Q := \lim_{\Delta \to 0} \frac{T_\Delta - I}{\Delta} \qquad \text{where} \qquad T_\Delta(x,y) := P(X_\Delta = y \,|\, X_0 = x)$$

For non-homogeneous case we write

$$T_t^{t+\Delta}(x,y) := P(X_{t+\Delta} = y \,|\, X_t = x),$$

which has a *time-dependent derivative*,

$$Q_t := \lim_{\Delta \to 0} \frac{T_t^{t+\Delta} - T_t^t}{\Delta} = \lim_{\Delta \to 0} \frac{T_t^{t+\Delta} - I}{\Delta}$$

Non-homogeneous case in continuous-time

$\mathbb{P}^{\mathrm{M}}_{\mathscr{Q}}$: all (**non**-homogeneous) Markov chains with $Q_t \in \mathscr{Q}$

How to interpret this?

Homogeneous case, rate matrix is just a derivative,

$$Q := \lim_{\Delta \to 0} \frac{T_\Delta - I}{\Delta} \qquad \text{where} \qquad T_\Delta(x,y) := P(X_\Delta = y \,|\, X_0 = x)$$

For non-homogeneous case we write

$$T_t^{t+\Delta}(x,y) := P(X_{t+\Delta} = y \,|\, X_t = x),$$

which has a *time-dependent derivative*,

$$Q_t := \lim_{\Delta \to 0} \frac{T_t^{t+\Delta} - T_t^t}{\Delta} = \lim_{\Delta \to 0} \frac{T_t^{t+\Delta} - I}{\Delta}$$

Setting explored by (Hartfiel, 1985) and (Škulj, 2015)

Continuous-time local models

We have

$$Q_t = \lim_{\Delta \to 0} \frac{T_t^{t+\Delta} - I}{\Delta}$$

and so for small Δ,

$$T_t^{t+\Delta} \approx I + \Delta Q_t$$

Continuous-time local models

We have

$$Q_t = \lim_{\Delta \to 0} \frac{T_t^{t+\Delta} - I}{\Delta}$$

and so for small Δ,

$$T_t^{t+\Delta} \approx I + \Delta Q_t$$

Then we can write

$$\underline{\mathbb{E}}_{\mathscr{Q}}^{\mathrm{M}}\left[f(X_{t+\Delta}) \,|\, X_t = x\right] = \inf_{T_t^{t+\Delta}} \left[T_t^{t+\Delta} f\right](x) \approx \inf_{Q \in \mathscr{Q}} \left[(I + \Delta Q) f\right](x)$$

Continuous-time local models

We have

$$Q_t = \lim_{\Delta \to 0} \frac{T_t^{t+\Delta} - I}{\Delta}$$

and so for small Δ,

$$T_t^{t+\Delta} \approx I + \Delta Q_t$$

Then we can write

$$\underline{\mathbb{E}}^{\mathrm{M}}_{\mathscr{Q}}\big[f(X_{t+\Delta})\,|\,X_t = x\big] = \inf_{T_t^{t+\Delta}} \big[T_t^{t+\Delta} f\big](x) \approx \inf_{Q \in \mathscr{Q}} \big[(I + \Delta Q) f\big](x)$$

We get

$$\underline{\mathbb{E}}^{\mathrm{M}}_{\mathscr{Q}}\big[f(X_{t+\Delta})\,|\,X_t = x\big] \approx \big[(I + \Delta \underline{Q}) f\big](x) \approx \underline{\mathbb{E}}^{\mathrm{M}}_{\mathscr{Q}}\big[f(X_{\Delta})\,|\,X_0 = x\big]$$

where we have defined

$$\big[\underline{Q} f\big](x) := \inf_{Q \in \mathscr{Q}} \big[Q f\big](x),$$

Again homogeneous lower expectation!

Arbitrary time points

If $\mathscr{Q}$ has separately specified rows,

$$\underline{\mathbb{E}}^{\mathrm{M}}_{\mathscr{Q}}\left[f(X_t)\,|\,X_0=x\right] \approx \left[(I + {}^{t}\!/\!{}_{n}\underline{Q})^n f\right](x)$$

and in fact

$$\underline{\mathbb{E}}^{\mathrm{M}}_{\mathscr{Q}}\left[f(X_t)\,|\,X_0=x\right] = \lim_{n\to+\infty} \left[(I + {}^{t}\!/\!{}_{n}\underline{Q})^n f\right](x)$$

Allows practical computation

- Solve $\inf_{Q\in\mathscr{Q}}[Q\cdot]$ multiple times
- Each is a linear optimisation problem

Arbitrary time points

If $\mathscr{Q}$ has separately specified rows,

$$\underline{\mathbb{E}}^{\mathrm{M}}_{\mathscr{Q}}\left[f(X_t)\,|\,X_0 = x\right] \approx \left[(I + {}^{t}\!/\!_{n}\underline{Q})^n f\right](x)$$

and in fact

$$\underline{\mathbb{E}}^{\mathrm{M}}_{\mathscr{Q}}\left[f(X_t)\,|\,X_0 = x\right] = \lim_{n\to+\infty}\left[(I + {}^{t}\!/\!_{n}\underline{Q})^n f\right](x)$$

Allows practical computation

- Solve $\inf_{Q\in\mathscr{Q}}[Q\cdot]$ multiple times
- Each is a linear optimisation problem

Better computational method in (Erreygers and De Bock, 2017)

The non-Markov case

For the set $\mathbb{P}_{\mathscr{Q}}$, derivative becomes *history* dependent.
Let $x_{\mathbf{u}} = x_{u_1}, \ldots, x_{u_n}$, $0 \leq u_1 < \cdots < u_n < t$. For all $x, y \in \mathscr{X}$,

$$Q_{t,x_{\mathbf{u}}}(x,y) := \lim_{\Delta \to 0} \frac{P(X_{t+\Delta} = y \,|\, X_{\mathbf{u}} = x_{\mathbf{u}}, X_t = x) - I(x,y)}{\Delta}$$

This is becoming a bit unwieldy...

The non-Markov case

For the set $\mathbb{P}_{\mathscr{Q}}$, derivative becomes *history* dependent.
Let $x_{\mathbf{u}} = x_{u_1}, \ldots, x_{u_n}$, $0 \leq u_1 < \cdots < u_n < t$. For all $x, y \in \mathscr{X}$,

$$Q_{t,x_{\mathbf{u}}}(x,y) := \lim_{\Delta \to 0} \frac{P(X_{t+\Delta} = y \,|\, X_{\mathbf{u}} = x_{\mathbf{u}}, X_t = x) - I(x,y)}{\Delta}$$

This is becoming a bit unwieldy...

Turns out that

$$\underline{\mathbb{E}}_{\mathscr{Q}}\big[f(X_{s+t}) \,|\, X_{\mathbf{u}} = x_{\mathbf{u}}, X_s = x\big] = \lim_{n \to +\infty} \big[(I + {}^{t}/{}_{n}\underline{Q})^n f\big](x)$$

The non-Markov case

For the set $\mathbb{P}_{\mathscr{Q}}$, derivative becomes *history* dependent.
Let $x_{\mathbf{u}} = x_{u_1}, \ldots, x_{u_n}$, $0 \leq u_1 < \cdots < u_n < t$. For all $x, y \in \mathscr{X}$,

$$Q_{t,x_{\mathbf{u}}}(x,y) := \lim_{\Delta \to 0} \frac{P(X_{t+\Delta} = y \,|\, X_{\mathbf{u}} = x_{\mathbf{u}}, X_t = x) - I(x,y)}{\Delta}$$

This is becoming a bit unwieldy…

Turns out that

$$\underline{\mathbb{E}}_{\mathscr{Q}}\left[f(X_{s+t}) \,|\, X_{\mathbf{u}} = x_{\mathbf{u}}, X_s = x\right] = \lim_{n \to +\infty} \left[(I + {}^{t}/{}_{n}\underline{Q})^n f\right](x) = \underline{\mathbb{E}}^{\mathrm{M}}_{\mathscr{Q}}\left[f(X_t) \,|\, X_0 = x\right]$$

Lower expectation for $\mathbb{P}_{\mathscr{Q}}$ has an *imprecise Markov property*!

- And is time-homogeneous!
- **Not** the same as $\mathbb{P}^{\mathrm{M}}_{\mathscr{Q}}$ when f depends on multiple time points!
    - Then only $\mathbb{P}_{\mathscr{Q}}$ remains tractable.

The non-Markov case

For the set $\mathbb{P}_{\mathscr{Q}}$, derivative becomes *history* dependent.
Let $x_{\mathbf{u}} = x_{u_1}, \ldots, x_{u_n}$, $0 \leq u_1 < \cdots < u_n < t$. For all $x, y \in \mathscr{X}$,

$$Q_{t,x_{\mathbf{u}}}(x,y) := \lim_{\Delta \to 0} \frac{P(X_{t+\Delta} = y \,|\, X_{\mathbf{u}} = x_{\mathbf{u}}, X_t = x) - I(x,y)}{\Delta}$$

This is becoming a bit unwieldy...

Turns out that

$$\underline{\mathbb{E}}_{\mathscr{Q}}\left[f(X_{s+t}) \,|\, X_{\mathbf{u}} = x_{\mathbf{u}}, X_s = x\right] = \lim_{n \to +\infty} \left[(I + {}^{t}\!/\!{}_{n}\underline{Q})^n f\right](x) = \underline{\mathbb{E}}^{\mathrm{M}}_{\mathscr{Q}}\left[f(X_t) \,|\, X_0 = x\right]$$

Lower expectation for $\mathbb{P}_{\mathscr{Q}}$ has an *imprecise Markov property*!

- And is time-homogeneous!
- **Not** the same as $\mathbb{P}^{\mathrm{M}}_{\mathscr{Q}}$ when f depends on multiple time points!
    - Then only $\mathbb{P}_{\mathscr{Q}}$ remains tractable.

Explored by (Krak *et al.*, 2017)

Continuous-time limit behaviour?

Limit inference often of interest:

$$\mathbb{E}\big[f(X_\infty)\,\big|\,X_0 = x\big] = \lim_{t\to+\infty} \mathbb{E}\big[f(X_t)\,\big|\,X_0 = x\big]$$

In imprecise setting, limit *always* exists:

$$\underline{\mathbb{E}}_{\mathscr{Q}}\big[f(X_\infty)\,\big|\,X_0 = x\big] = \lim_{t\to+\infty} \underline{\mathbb{E}}_{\mathscr{Q}}\big[f(X_t)\,\big|\,X_0 = x\big]$$

and often independent of x.

See (De Bock, 2017)

Main take away points

If we do not know T or Q, we can consider *sets* $\mathscr{T}$ or $\mathscr{Q}$

Gives rise to three different *imprecise* models:

- Set of homogeneous Markov chains
- Set of **non**-homogeneous Markov chains
- Set of **non**-Markov processes

For homogeneous Markov chains:

- *Difficult* to work with

For non-homogeneous and non-Markov processes:

- Efficient computations using *local models* $\underline{T}$ or $\underline{Q}$
- Have *homogeneous* lower expectations
- Have "*Markov*" lower expectations

That’s all fine and
well, but what can you
use it for?

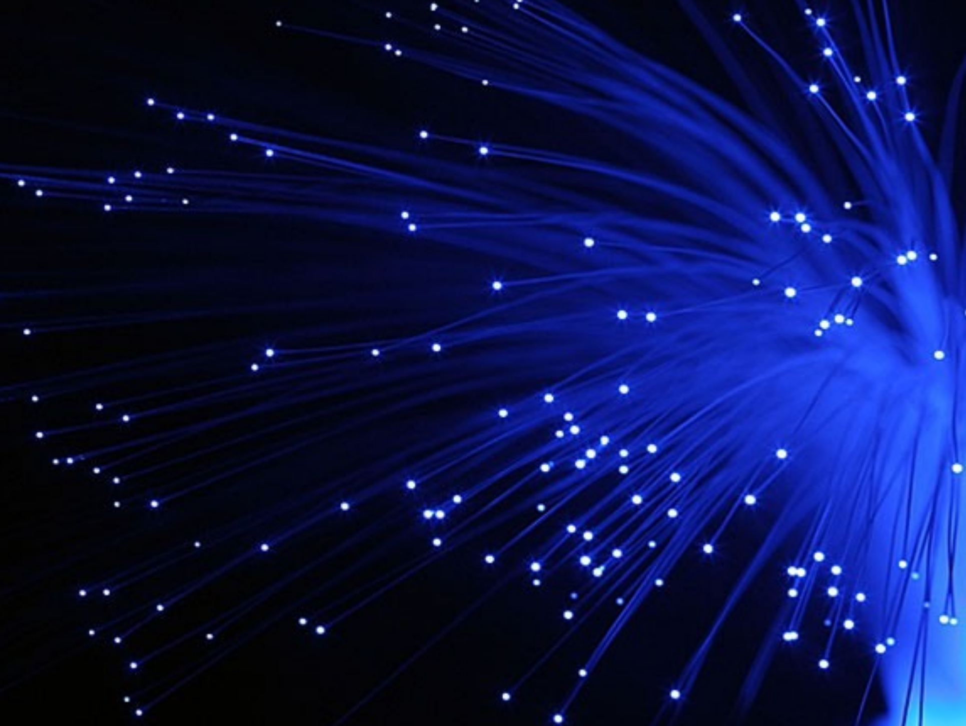
- **Reliability engineering** (failure probabilities, …)
- **Queuing theory** (waiting in line …)
  - optimising supermarket waiting times
  - dimensioning of call centers
  - airport security lines
  - router queues on the internet

- **Chemical reactions** (time-evolution …)

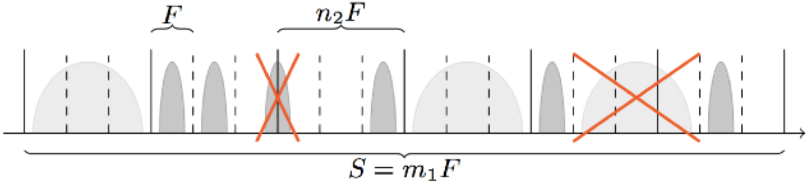
- **Pagerank**

Google

- **…**



Message passing in optical links



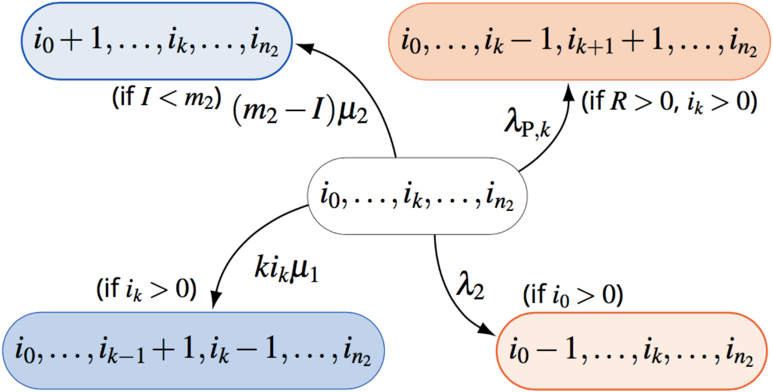
m_1 **channels** $m_2 = \dfrac{m_1}{n_2}$ **superchannels**

type I messages require 1 channel

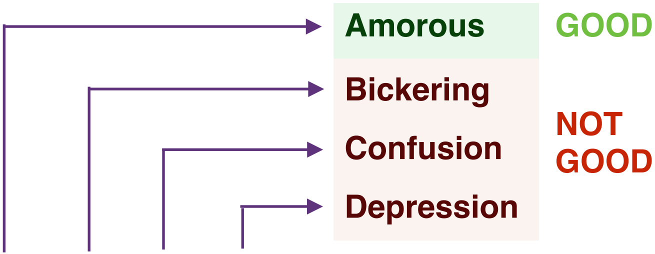
type II messages require n_2 channels

We want to know the blocking probability of messages for a given policy, and optimise it

$$\mathscr{X}_{\text{det}} := \left\{ (i_0, \ldots, i_{n_2}) \in \mathbb{N}^{(n_2+1)} : \sum_{k=0}^{n_2} i_k \leq m_2 \right\}$$



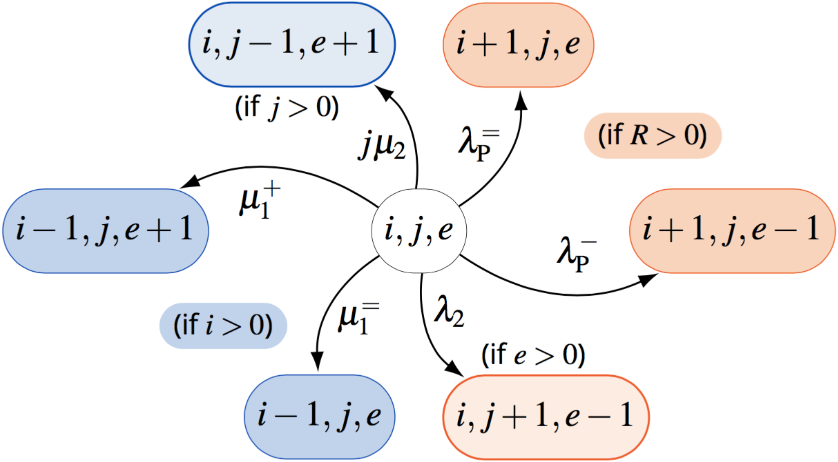
$I := \sum_{k=0}^{n_2} i_k \qquad R := \sum_{k=0}^{n_2-1} i_k (n_2 - k)$



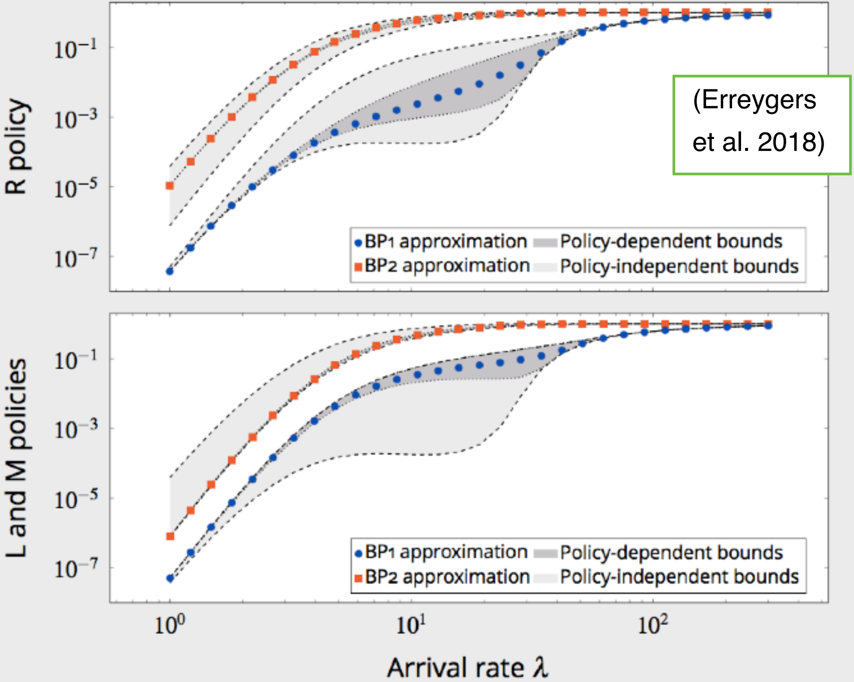
$$
\begin{array}{c c} & \begin{array}{cccc} \mathbf{A} & \mathbf{B} & \mathbf{C} & \mathbf{D} \end{array} \\
Q = \begin{array}{c} \mathbf{A} \\ \mathbf{B} \\ \mathbf{C} \\ \mathbf{D} \end{array} & \begin{bmatrix} -4 & 3 & 1 & 0 \\ 4 & -6 & 2 & 0 \\ 2 & 3 & -6 & 1 \\ 0 & 0 & 2 & -2 \end{bmatrix} \end{array}
$$

(Erreygers & De Bock 2018)

$$\mathscr{X}_{\mathrm{red}} := \{(i,j,e) \in \mathbb{N}^3 : m_2 \le i+j+e, i+(j+e)n_2 \le m_1\}$$



$$R := m_1 - i - jn_2$$



(Erreygers et al. 2018)
R policy
L and M policies
BP1 approximation
BP2 approximation
Policy-dependent bounds
Policy-independent bounds
10^{-1}
10^{-3}
10^{-5}
10^{-7}
10^{0}
10^{1}
10^{2}
Arrival rate λ

Advantages of imprecise Markov chains over their precise counterpart

- ✓ Partially specified π_0 and Q/T are allowed
- ✓ Time homogeneity can be relaxed
- ✓ The Markov assumption can be relaxed
- ✓ Efficient computations remain possible
- ✓ State space explosion can be dealt with

All of this sounds too good to be true! What have you been hiding?

Can we learn these from data?

IDM (Walley 1996)
(Quaegheberur 2009)
(Krak et al. 2018)

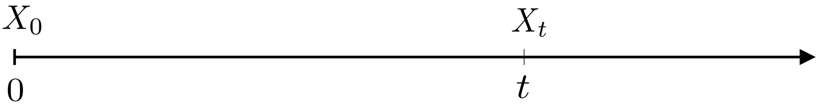
initial distribution π_0

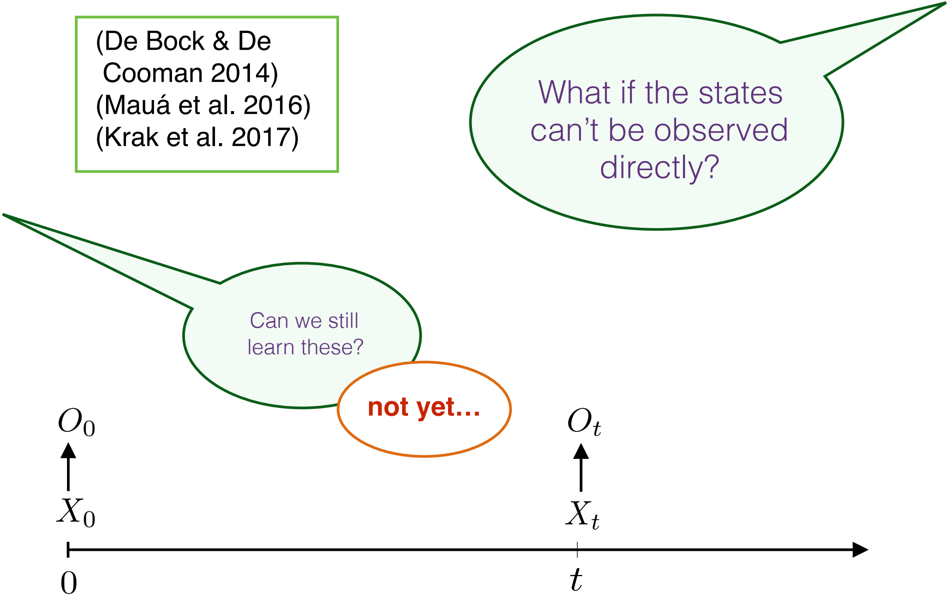
+

transition rate matrix Q

or

transition matrix T




(De Bock & De Cooman 2014)
(Mauá et al. 2016)
(Krak et al. 2017)
What if the states can’t be observed directly?
Can we still learn these?
not yet…
O_0
X_0
O_t
X_t
0
t

How about
more complicated
inferences?

in some cases…

(Troffaes et al. 2015)
(Lopatatzidis 2017)

X_0 X_t

0 t

(Peng 2005)

Can we
do infinite state
spaces?

only in theory…

X_0 X_t

0 t

References (1)

T. Augustin, F. P. A. Coolen, G. De Cooman, M.C.M. Troffaes. *Introduction to Imprecise Probabilities*. Wiley, 2014.

J. De Bock: The limit behaviour of imprecise continuous-time Markov chains. *J. Nonlinear Science*. 27(1), 159–196 (2017)

J. De Bock, G. de Cooman: An efficient algorithm for estimating state sequences in imprecise hidden Markov models. *Journal of Artificial Intelligence Research*, 50: 189–233. 2014.

M. Campos, G.P. Dimuro, A. da Rocha Costa, and V. Kreinovich: Computing 2-step predictions for interval-valued finite stationary Markov chains. *Technical report utep-cs-03-20a*, University of Texas at El Paso, 2003.

G. de Cooman and F. Hermans: Imprecise probability trees: Bridging two theories of imprecise probability. *Artificial Intelligence*, 172:1400–1427, 2008.

G. de Cooman, F. Hermans, and E.Quaeghebeur: Imprecise Markov chains and their limit behavior. *Probability in the Engineering and Informational Sciences*, 23:597–635, 2009

A. Erreygers and J. De Bock: Imprecise continuous-time Markov chains: Efficient computational methods with guaranteed error bounds. In: *Proceedings of ISIPTA 2017*, pp. 145–156, 2017

References (2)

A. Erreygers and J. De Bock: Inferences for Large- Scale Continuous-Time Markov Chains by Combining Lumping with Imprecision. *Accepted for publication in Proceedings of SMPS*, 2018.

A. Erreygers, C. Rottondi, G. Verticale, J. De Bock: Imprecise Markov Models for Scalable Robust Performance Evaluation of Flexi-Grid Spectrum Allocation Policies. *Accepted for publication in IEEE Transactions on Communications*. 2018.

A. Goldsztejn and A. Neumaier: On the exponentiation of interval matrices, *Reliab. Comput*. 20:52–72, 2014

D. Hartfiel: *Markov Set-Chains*. Springer, Berlin, 1998.

D. Hartfiel: On the solutions to $x'(t) = a(t)x(t)$ over all $a(t)$, where $p \leq a(t) \leq q$. *Journal of Mathematical Analysis and Applications*, 108:230–240, 1985.

I. Kozine and L. Utkin: Interval-valued finite Markov chains. *Reliable Computing*, 8:97–113, 2002.

T. Krak, J. De Bock, A. Siebes: Imprecise continuous-time Markov chains. *Int. J. Approx. Reason*. 88, 452–528, 2017

T. Krak, J. De Bock, A. Siebes: Efficient computation of updated lower expectations for imprecise continuous-time hidden Markov chains. *PMLR: proceedings of machine learning research, 62 (proceedings of ISIPTA '17)*, 193–204. 2017.

References (3)

T. Krak, A. Erreygers, J. De Bock: An Imprecise Probabilistic Estimator for the Transition Rate Matrix of a Continuous-Time Markov Chain. *Accepted for publication in the Proceedings of SMPS* 2018.

S. Lopatatzidis: *Robust modelling and optimisation in stochastic processes using imprecise probabilities, with an application to queueing theory*. PhD Thesis, Ghent University. 2017.

D. D. Mauá, A. Antonucci, C. P. de Campos: Hidden Markov models with set-valued parameters. *Neurocomputing*, 180: 94–107. 2016.

E.P. Oppenheimer, A.N. Michel: Application of interval analysis techniques to linear systems. II. The interval matrix exponential function. *IEEE Trans. Circuits Syst.* 35(10):1230–1242,1988

S. Peng. Nonlinear expectations and nonlinear Markov chains. *Chinese Ann. Math. Ser. B*, 26(2):159–184. 2005.

E. Quaeghebeur: *Learning from samples using coherent lower previsions*. PhD Thesis, Ghent University. 2009.

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

D. Škulj: Discrete time Markov chains with interval probabilities. *International Journal Approximate Reasoning*, 50:1314–1329, 2009.

References (4)

D. Škulj: Efficient computations of the bounds of continuous time imprecise Markov chains. *Applied Mathematics and Computation*, 250:165–180, 2015.

M.C.M. Troffaes, J. Gledhill, D. Skulj, S. Blake: Using imprecise continuous time Markov chains for assessing the reliability of power networks with common cause failure and non-immediate repair. *Proceedings of ISIPTA '15*: 287–294, 2015.

P. Walley: Inferences from multinomial data: learning about a bag of marbles. *Journal of the Royal Statistical Society, Series B*, 58:3–57. 1996.

P. Walley: *Statistical Reasoning with Imprecise Probabilities* Chapman and Hall, 1991.

This work was partially supported by H2020-MSCA-ITN-2016 UTOPIAE, grant agreement 722734.

http://twitter.com/utopiae_network http://utopiae.eu