# Efficient Computation of Updated Lower Expectations for Imprecise Continuous-Time Hidden Markov Chains

Thomas Krak, Jasper De Bock, Arno Siebes

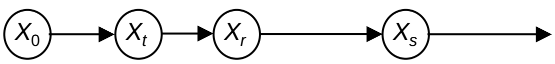
**Abstract** We consider the problem of performing inference with *imprecise continuous-time hidden Markov chains*, that is, *imprecise continuous-time Markov chains* that are augmented with random *output* variables whose distribution depends on the hidden state of the chain. The prefix `imprecise' refers to the fact that we do not consider a classical continuous-time Markov chain, but replace it with a robust extension that allows us to represent various types of model uncertainty, using the theory of *imprecise probabilities*. The inference problem amounts to computing lower expectations of functions on the state-space of the chain, given observations of the output variables. We develop and investigate this problem with very few assumptions on the output variables; in particular, they can be chosen to be either discrete or continuous random variables. Our main result is a polynomial runtime algorithm to compute the lower expectation of functions on the state-space at any given time-point, given a collection of observations of the output variables.

## "Precise" Continuous-Time Markov Chains

State-space $\mathcal{X}$ (e.g., $\mathcal{X} = \{\text{healthy, sick}\}$)

Continuous-time Markov chain $P$ specifies r.v. $X_t$ at each time $t \in \mathbb{R}_{\geq 0}$

For any finite number of time-points, e.g. $0 < t < r < s$, $P$ induces a *Bayesian network:*



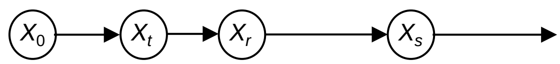

Satisfies Markov property: $P(X_s \mid X_0, X_t, X_r) = P(X_s \mid X_r)$

## Imprecise Continuous-Time Markov Chains

Now a set $\mathcal{P}$ of distributions.

Each $P \in \mathcal{P}$ specifies r.v. $X_t$ at each time $t \in \mathbb{R}_{\geq 0}$

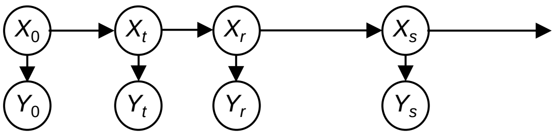
For any finite number of time-points, e.g. $0 < t < r < s$, $\mathcal{P}$ induces a *credal network:*




Satisfies *imprecise Markov property*: $\underline{P}(X_s \mid X_0, X_t, X_r) = \underline{P}(X_s \mid X_r)$

## Imprecise CT Hidden Markov Chains

States $X_t$ cannot be directly observed. Instead we observe $Y_t$, which "correlates" with $X_t$ (e.g., symptoms of a disease).




For simplicity, we use a *precise, homogeneous* output model:

$$\underline{P}(Y_t \mid X_t) = P(Y_t \mid X_t) = P(Y \mid X), t \in \mathbb{R}_{\geq 0}$$

We are interested in inferences about the states given observations. For example, given some $O \subseteq \mathcal{Y}$, we want to know $\underline{\mathbb{E}}[f(X_s) \mid Y_t \in O]$.

## Outputs with Positive (Upper) Probability

If the observation $(Y_t \in O)$ has positive probability, we use Bayes' rule:

$$\mathbb{E}_P[f(X_s) \mid Y_t \in O] := \sum_{x \in \mathcal{X}} f(x) \frac{P(X_s = x, Y_t \in O)}{P(Y_t \in O)}$$

For the imprecise model, we use *regular extension*:

$$\underline{\mathbb{E}}[f(X_s) | Y_t \in O] := \inf\{\mathbb{E}_P[f(X_s) | Y_t \in O] : P \in \mathcal{P}, P(Y_t \in O) > 0\},$$

whenever $\overline{P}(Y_t \in O) > 0$.

This lower expectation satisfies a *generalised Bayes' rule*:

$$\underline{\mathbb{E}}[f(X_s) | Y_t \in O] = \max\{\mu \in \mathbb{R} : \underline{\mathbb{E}}[P(Y_t \in O \mid X_t)(f(X_s) - \mu)] \geq 0\}$$

## Continuous Outputs

If $Y_t$ is continuous, then usually $P(Y_t = y) = 0$ for all $P \in \mathcal{P}$.
Assume a (conditional) probability density function $\phi : \mathcal{Y} \times \mathcal{X} \to \mathbb{R}$:

$$P(Y_t \in O \mid X_t = x) = \int_O \phi(y|x)\, dy$$

Take a sequence $\{O_i\}_{i \in \mathbb{N}}$ such that $\lim_{i \to \infty} O_i = \{y\}$. Then define

$$\mathbb{E}_P[f(X_s) \mid Y_t = y] := \lim_{i \to \infty} \mathbb{E}_P[f(X_s) \mid Y_t \in O_i]$$

This limit exists under suitable assumptions; if $\mathbb{E}_P[\phi(y \mid X_t)] > 0$:

$$\mathbb{E}_P[f(X_s) \mid Y_t = y] = \frac{\mathbb{E}_P[f(X_s)\phi(y \mid X_t)]}{\mathbb{E}_P[\phi(y \mid X_t)]}$$

## Continuous Outputs, Imprecise Case

For the imprecise case, when $\underline{\mathbb{E}}[\phi(y \mid X_t)] > 0$ we define

$$\underline{\mathbb{E}}[f(X_s) \mid Y_t = y] := \inf\{\mathbb{E}_P[f(X_s) \mid Y_t = y] : P \in \mathcal{P}\}$$

This lower expectation satisfies a limit interpretation

$$\underline{\mathbb{E}}[f(X_s) \mid Y_t = y] = \lim_{i \to \infty} \underline{\mathbb{E}}[f(X_s) \mid Y_t \in O_i]$$

and a *generalised Bayes' rule for (finite) mixtures of densities*:

$$\underline{\mathbb{E}}[f(X_s) \mid Y_t = y] = \max\{\mu \in \mathbb{R} : \underline{\mathbb{E}}[\phi(y \mid X_t)(f(X_s) - \mu)] \geq 0\}$$

## Solving the Generalised Bayes' Rule(s)

In both cases, we have a generalised Bayes' rule:

$$\underline{\mathbb{E}}[f(X_s) | Y_t \in O] = \max\{\mu \in \mathbb{R} : \underline{\mathbb{E}}[P(Y_t \in O \mid X_t)(f(X_s) - \mu)] \geq 0\}$$
$$\underline{\mathbb{E}}[f(X_s) \mid Y_t = y] = \max\{\mu \in \mathbb{R} : \underline{\mathbb{E}}[\phi(y \mid X_t)(f(X_s) - \mu)] \geq 0\}$$

See the paper for a polynomial runtime algorithm to solve these.

Universiteit Utrecht

GHENT UNIVERSITY