Sensitivity analysis for finite Markov chains in discrete time

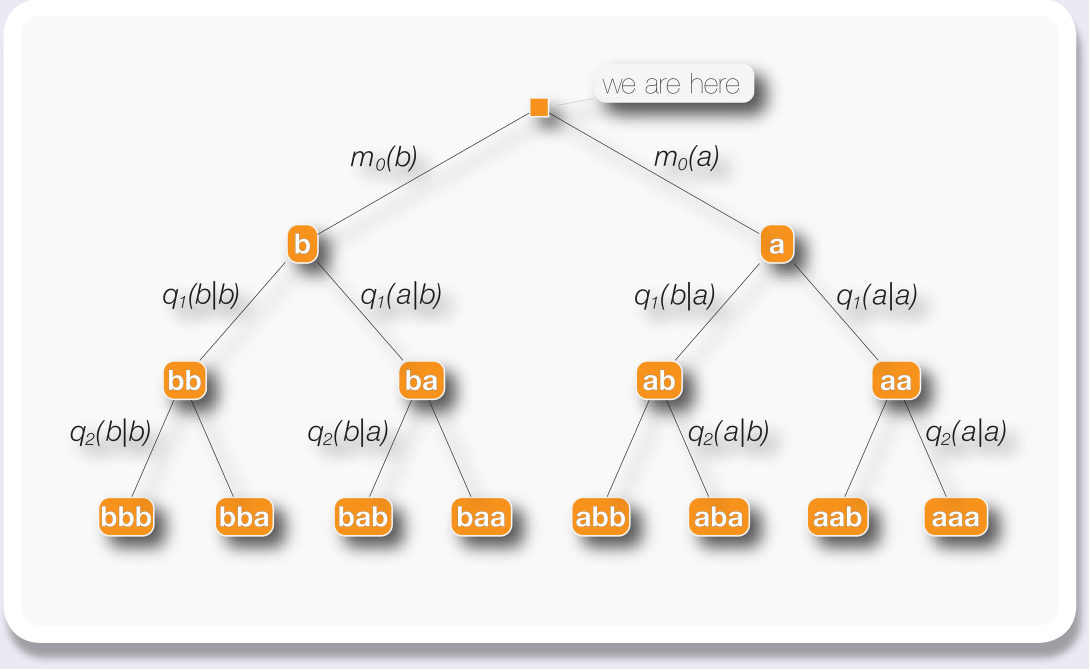
Gert de Cooman, Filip Hermans & Erik Quaeghebeur

SYSTeMS Research Group, EESA Department, Ghent University, Belgium
{Gert.deCooman, Filip.Hermans, Erik.Quaeghebeur}@UGent.be

UNIVERSITEIT GENT

1. Markov tree

Consider a Markov chain with the state space {*a*,*b*} and the probability mass functions m_1(initial) and $q_k(.|a)$, $q_k(.|b)$.



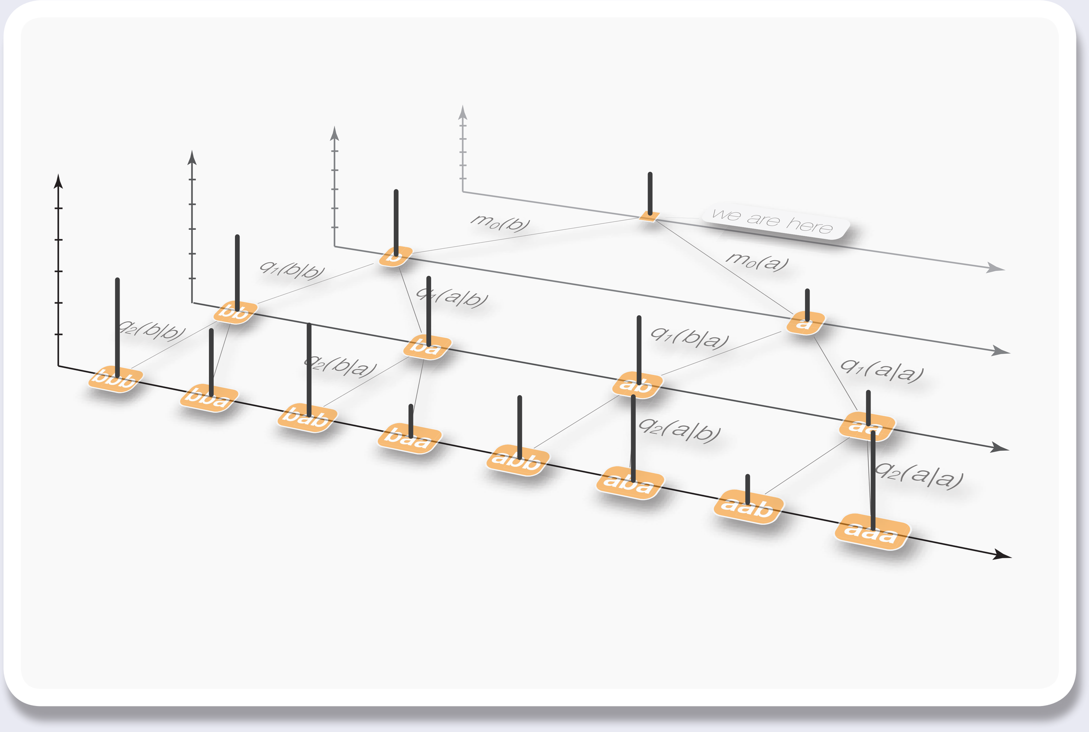
The probability distribution after the k-th time step is given by

$$m_{k+1}^T = m_k^T \cdot T_k = m_k^T \cdot \begin{bmatrix} q_k(a|a) & q_k(b|a) \\ q_k(a|b) & q_k(b|b) \end{bmatrix} \text{ and } m_1 = \begin{bmatrix} m_1(a) \\ m_1(b) \end{bmatrix}.$$

Expectation in the Markov tree

Consider a real-valued function *f* on $\{a,b\}^n$. Its expectation *P(f)* in different parts of the tree is given by the equation

$$P_k(f) = p_k^T \cdot f = p_0^T \cdot T_k \cdot f = P_0(T_k f) \text{ where}$$
$$P_0(f) = f(a)m_0(a) + f(b)m_0(b).$$



4. Convergence theorem

We can prove a Perron-Frobenius theorem for the nonlinear (transition) operator $\overline{T}$:

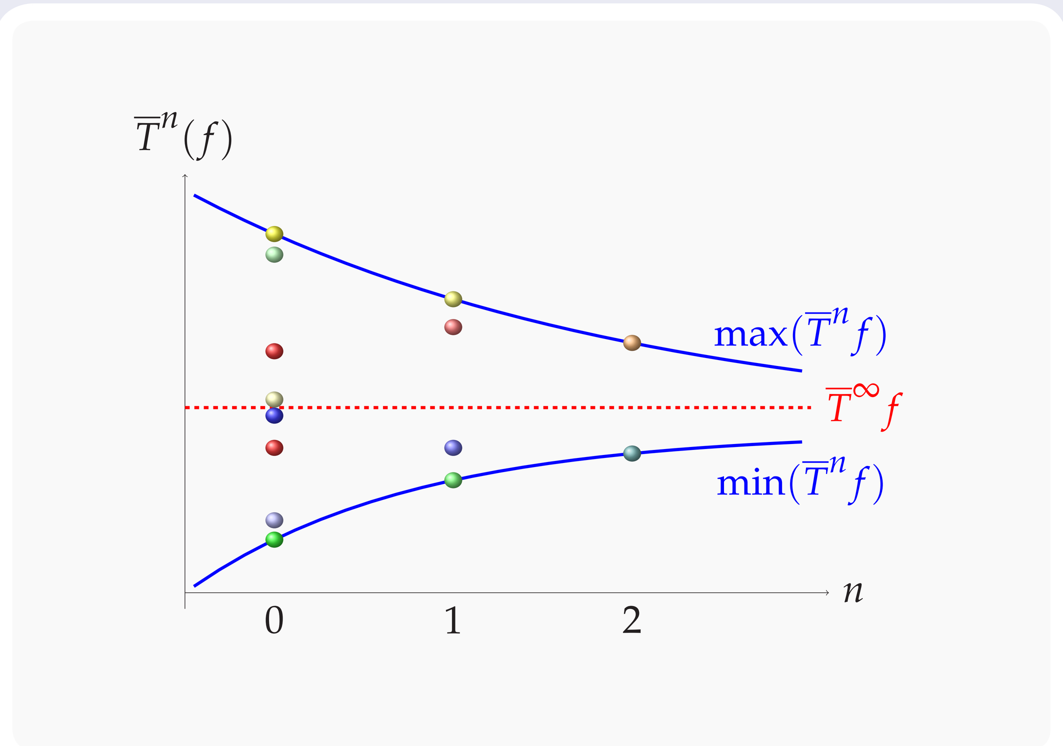
Convergence Theorem

If the transition operator $\overline{T}$ is regular, then the sequence $\overline{T}^n, n \in \mathbb{N}$ converges pointwise to some operator $\overline{T}^\infty$ and for any real-valued function f, $\overline{T}^\infty f$ is some constant.

For the limit distribution we find that

$$\lim_{n\to\infty} \overline{P}_n(f) = \overline{T}^\infty f = c_f.$$

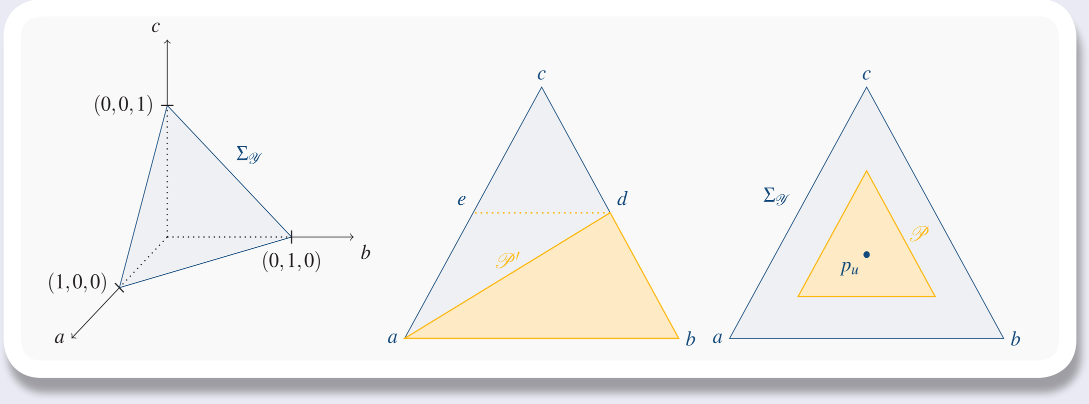
So its value is independent of the initial distribution $\overline{P}_0$, as in the classical case.



2. Imprecise Markov trees

Every probability mass function can be seen as a point of some unit simplex

$$\Sigma_{\mathcal{Y}} = \{p \in \mathbb{R}^{\mathcal{Y}} \colon \sum_{y \in \mathcal{Y}} p(y) = 1 \text{ and } (\forall y \in \mathcal{Y})(p(y) \geq 0)\}$$



For imprecise Markov chains, the assumption is that we only know that the mass functions belong to a credal set: a closed convex subset of the unit simplex.

Expectations in imprecise Markov chains

The linear expectation operator is replaced by a sublinear lower expectation and a superlinear upper expectation operator. They are the minimum and maximum of the expectation over all the probability mass functions of the credal set.

$$\underline{P}(f) = \min\{p^T f : p \in \mathcal{M}\} \text{ and } \overline{P}(f) = \max\{p^T f : p \in \mathcal{M}\}.$$

Using the credal sets assigned to each node of the tree, these functionals can be used to define the lower and upper transition operators.

$$\underline{T}_k f(y) = \min\{q^T f : q \in \mathcal{M}_k(\cdot|y)\} \text{ and}$$
$$\overline{T}_k f(y) = \max\{q^T f : q \in \mathcal{M}_k(\cdot|y)\}.$$

Similarly, the lower and upper previsions in the initial situation are calculated

$$\underline{P}_0 f(y) = \min\{m^T f : m \in \mathcal{M}_0\} \text{ and}$$
$$\overline{P}_0 f(y) = \max\{m^T f : m \in \mathcal{M}_0\}.$$

Calculating the joint distribution

Because of the nonlinear character of the upper and lower expecation functionals, the joint distribution can not be calculated directly from the imprecise joint mass functions. Luckily, the approach using transition operators can be generalized to the imprecise case:

$$\underline{P}_k(f) = \underline{P}_0(\underline{T}^k f) \text{ and } \overline{P}_k(f) = \overline{P}_0(\overline{T}^k f)$$

The backpropagation of f (i.e. $f \to \underline{T}_n f \to \underline{T}_{n-1}\underline{T}_n f \to \ldots$) is linear in the number of transition steps. The complexity of each transition step depends on the nature of the credal sets.

5. Contamination model

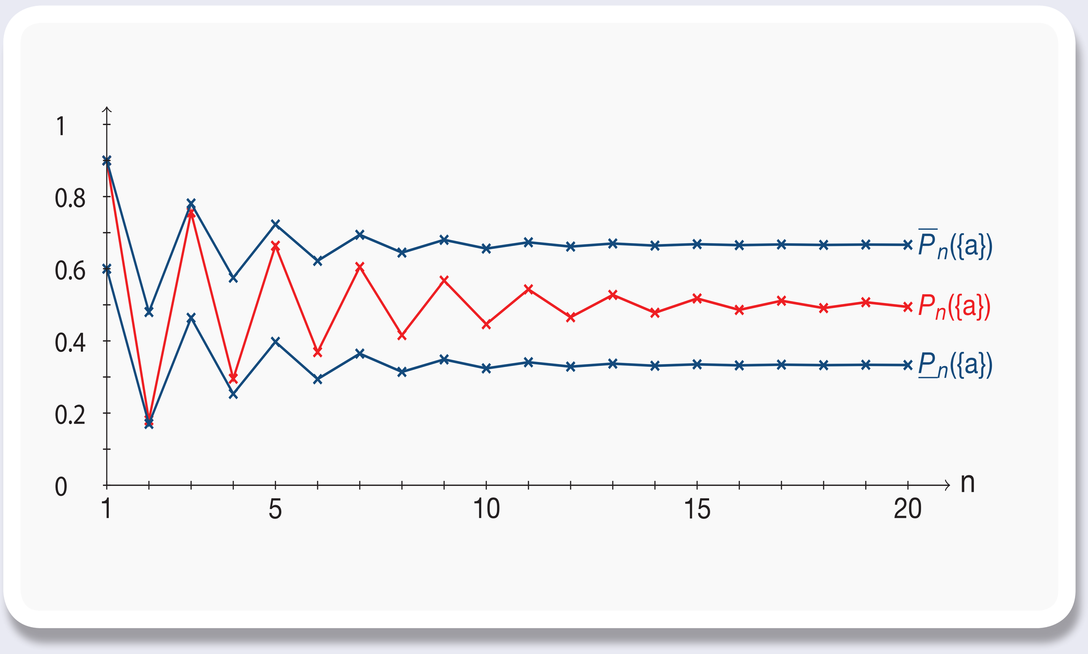
Suppose we consider a precise stationary Markov chain, with precise transition operator which we contaminate with a vacuous model, i.e., we take a convex mixture with the upper transition operator $\overline{T}$ = *max*. This leads to the upper transition operator

$$\overline{T}f = (1-\epsilon)Th + \epsilon \max f \text{ with } \epsilon \in [0,1].$$

Consider for example the stationary imprecise Markov chain with $\mathcal{X} = \{a,b\}$ and the initial credal set defined by

$$\mathcal{M}_0 = \{m \in \Sigma_{\{a,b\}} : 0.6 \leq m(a) \leq 0.9\}.$$

Let moreover $\epsilon = 0.1$ and the precise transition matrix $T = \begin{bmatrix} 0.15 & 0.85 \\ 0.85 & 0.15 \end{bmatrix}$.



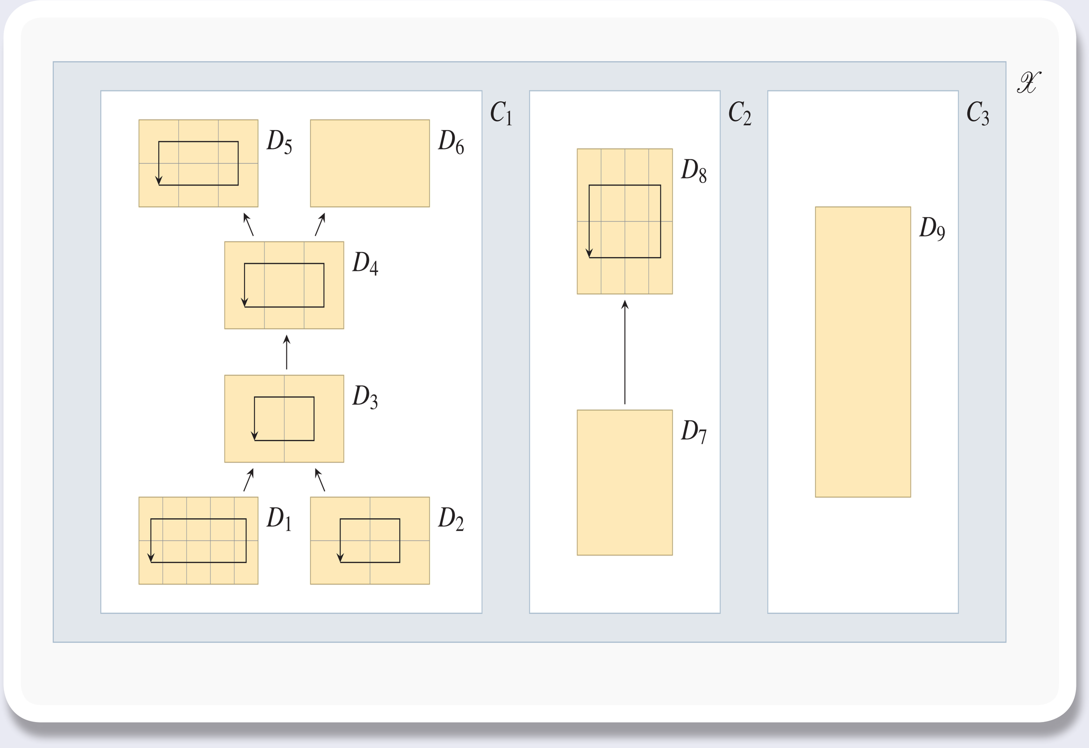
3. State classification

If we define $\overline{P}_n(y|x) := \overline{P}_0(\overline{T}_n y(x))$, i.e. the upper expectation that the system is evolving from state x into state y in n steps then the following inequality can be proved.

For any state *x* and *y* of an imprecise Markov tree, and for any natural number *n* and *m*,

$$\overline{P}_{xy}^{n+m} \geq \overline{P}_{xz}^{n}\overline{P}_{zy}^{m}.$$

Using this property, we can define the accessibility relation *x* → *y* (*y* is accessible from *x*). This relation defines a communication relation which is an equivalence relation and divides the possibility space in communication classes *x*↔*y*. Moreover, within each communication class, states can be subdivided in cyclic classes.



Acyclic maximal classes are called ergodic. If a transition operator has only one class which is ergodic, then we call this transition operator regular.

6. Quasi cyclic example

Consider a three-state stationary imprecise Markov model with $\mathcal{X} = \{a,b,c\}$ and with marginal and transition probabilities given by probability intervals.

The upper transition operator $\overline{T}$ is fully determined by the upper and lower Markov matrices:

$$\underline{T} := [\underline{T}\{a\} \quad \underline{T}\{b\} \quad \underline{T}\{c\}] = \begin{bmatrix} \underline{q}(a|a) & \underline{q}(b|a) & \underline{q}(c|a) \\ \underline{q}(a|b) & \underline{q}(b|b) & \underline{q}(c|b) \\ \underline{q}(a|c) & \underline{q}(b|c) & \underline{q}(c|c) \end{bmatrix} = \frac{1}{200}\begin{bmatrix} 9 & 9 & 162 \\ 144 & 18 & 18 \\ 9 & 162 & 9 \end{bmatrix},$$

$$\overline{T} := [\overline{T}\{a\} \quad \overline{T}\{b\} \quad \overline{T}\{c\}] = \begin{bmatrix} \overline{q}(a|a) & \overline{q}(b|a) & \overline{q}(c|a) \\ \overline{q}(a|b) & \overline{q}(b|b) & \overline{q}(c|b) \\ \overline{q}(a|c) & \overline{q}(b|c) & \overline{q}(c|c) \end{bmatrix} = \frac{1}{200}\begin{bmatrix} 19 & 19 & 172 \\ 154 & 28 & 28 \\ 19 & 172 & 19 \end{bmatrix}.$$

Similarly, the initial upper expectation $\overline{P}_0$ is completely determined by the matrices:

$$\underline{P}_0 \, [\underline{m}_0(a) \quad \underline{m}_0(b) \quad \underline{m}_0(c)] \quad \text{and} \quad \overline{P}_0 \, [\overline{m}_0(a) \quad \overline{m}_0(b) \quad \overline{m}_0(c)].$$