Exact Computation and Bounds for the Coupling Time

Sebastian Samain and Ana Bušić
INRIA
DI ENS, École Normale Supérieure, CNRS, PSL Research University, Paris
sebastien.samain@inria.fr

March, 12th 2017

Coupling Time

Objective : Compute the coupling time for an homogenous Markov chain with representation

$$X_{n+1} = X_n \triangleright \xi_{n+1}$$

where ξ is i.i.d over a finite state space $\mathcal{A}$.

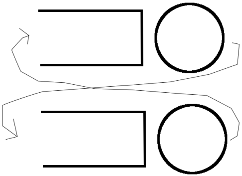
Definition

The coupling time T for a Markov Chain X with state space $\mathcal{X}$ and representation $X_{n+1} = X_n \triangleright \xi_{n+1}$ is the first (random) instant T such that

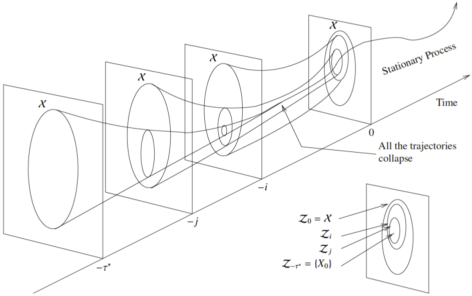
$$\forall (x, y) \in \mathcal{X} \times \mathcal{X}, x \triangleright^* \xi_1 ... \xi_T = y \triangleright^* \xi_1 ... \xi_T$$

Applications

- Mixing time for Markov chains (bounds for the Monte-Carlo algorithm)
- Expected complexity of the Propp and Wilson algorithm (based on the CFTP coupling scheme)
- Queueing networks



Coupling from the past



Methodology

- Language of coupling words : the prefix-free language of words $w \in \mathcal{A}^*$ such that :

$$\forall (x, y) \in \mathcal{X} \times \mathcal{X}, x \triangleright^* w = y \triangleright^* w$$

- Language of coupling words = hitting language in some automaton
- Unambiguous decomposition
- Generating series $G(p_1, \ldots, p_n)$
- $\mathbb{E}[T] = D(G)(\mathbb{P}(p_1), \ldots, \mathbb{P}(p_n))$
  with $D = \sum_{i=1}^{n} p_i \frac{\partial}{\partial p_i}$ and $(\{p_1, \ldots, p_n\}, \mathbb{P})$ a probability space

First step

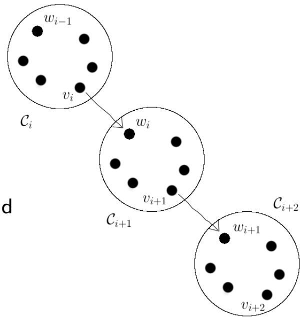
We define the enveloppe automaton $\mathcal{A}_{env} = (\mathcal{A}, \mathcal{E}, \triangleright_{\mathcal{E}})$, with

- $\mathcal{A}$ the (finite) state space of the process ξ
- $\mathcal{E} = \mathcal{P}(\mathcal{X})$
- $\forall (a, P) \in \mathcal{A} \times \mathcal{E}, P \triangleright_{\mathcal{E}} a = \{p \triangleright a | p \in P\}$

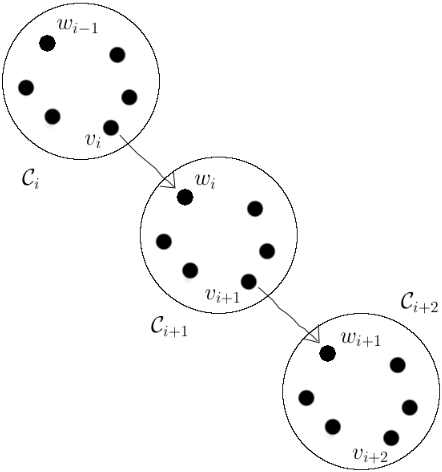
The set of coupling words corresponds then to the hitting language of $\mathcal{S}(\mathcal{X})$ the set of singletons of elements of $\mathcal{X}$ starting from $\mathcal{X}$ in $\mathcal{A}_{env}$.

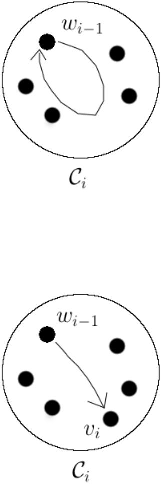
Second step

- Longest path into the maximal stronlgy connected components (MSCC) for $\triangleright_{\mathcal{E}}$.
- Queueing systems : linked to distance between top and bottom element.



Third step


w_{i-1}
v_i
$\mathcal{C}_i$
w_i
v_{i+1}
$\mathcal{C}_{i+1}$
$\mathcal{C}_{i+2}$
w_{i+1}
v_{i+2}


w_{i-1}
$\mathcal{C}_i$
w_{i-1}
v_i
$\mathcal{C}_i$

First result

Theorem

Let X a Markov Chain with representation

$$X_{n+1} = X_n \triangleright \xi_{n+1}$$

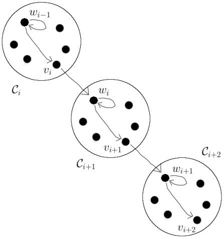
Then:

$$A_{\mathcal{X},\mathcal{S}(\mathcal{X})}(p)=\sum_{y\in\mathcal{S}(\mathcal{X})}\sum_{\mathcal{C}\in\mathcal{C}h\mathcal{C}on(C_x,C_{y,A})} L_{\mathcal{C}}(p)$$

$$L_{\mathcal{C}}(p)=\sum_{(v,w)\in\mathfrak{P}ass(\mathcal{C})} \frac{1}{1-R_{C_x,x}(p)} A_{C_x,x,\{v_0\}}(p)$$

$$\gamma\left(\sum_{a\in(v_0\triangleright .)^{-1}(w_0)} a\right) S_{\mathcal{C},v,w}(p) \frac{1}{1-R_{C_{y,A},w_{|\mathcal{C}|-1}}(p)} A_{C_{y,A},w_{|\mathcal{C}|-1},\{y\}}(p)$$

$$S_{\mathcal{C},v,w}(p)=\left(\prod_{i=2}^{|\mathcal{C}|-2} \frac{1}{1-R_{\mathcal{C}_i,w_{i-1}}(p)} A_{\mathcal{C}_i,w_{i-1},\{v_i\}}(p)\gamma\left(\sum_{a\in(v_i\triangleright .)^{-1}(w_i)} a\right)\right)$$



Compare formal series

- Generating series of events A for the reprensentation $X_{n+1} = X_n \triangleright \xi_{n+1}$ define conditional expectations for the coupling time T :

$$\frac{D(G)(\mathbb{P}(p_1), ..., \mathbb{P}(p_n))}{G(\mathbb{P}(p_1), ..., \mathbb{P}(p_n))} = \mathbb{E}_{\mathbb{P}}[T|A]$$

with $D = \sum_{i=1}^{n} p_i \frac{\partial}{\partial p_i}$

- This leads to the following criterion of comparison $\preceq_{ratio}$.

$\preceq_{ratio}$

Definition

We define the binary relation $\preceq_{ratio}$ over the generating series in $p_1, \ldots, p_n$ by $G \preceq_{ratio} H$ iff :

$$\forall \lambda \in \mathcal{FM}(\{p_i | i \in [1, n]\}), \frac{D(G)(\lambda)}{G(\lambda)} \leq \frac{D(H)(\lambda)}{H(\lambda)}$$

where $\mathcal{FM}(\{p_i | i \in [1, n]\})$ is the set of finite measures over the formal variables $p_1, \ldots, p_n$ and $\forall I$ formal series in $p_1, \ldots, p_n$, $I(\lambda) = I(\lambda(p_1), \ldots, \lambda(p_n))$

Compare formal series

Definition

We define the binary relation $\preceq$ over the generating series in $p_1, \dots, p_n$ by: $G \preceq H$ iff $\exists M$ a set of monomials in the variables $p_1, \dots, p_n$ such that

- $H = \sum_{m \in M} m H_{[m]}$
- $\forall m \in M, \forall (j, k)$ monomials in $p_1, \dots, p_n$ such that

$$deg_{tot}(j) > deg_{tot}(k) \implies ([j]H_{[m]})([k]G) \geq ([k]H_{[m]})([j]G)$$

Compare formal series

Theorem

$$G \preceq H \implies G \preceq_{ratio} H$$

Proof.

Study $D(G)H$ and $GD(H)$ monomial by monomial. □

Compare formal series

Theorem

*If I is a formal series over the set of monomials M,
$G = \sum_{m \in M} \frac{[m]I}{I} m G_{[m]}$, $H = \sum_{m \in M} \frac{[m]I}{I} m G_{[m]}$ and $\forall m \in M$, $G_{[m]} \preceq_{ratio} H_{[m]}$, then $G \preceq_{ratio} H$.*

Proof.

If M_I is the set of monomials of I, we use the fact that :

$$\frac{D(G)(\lambda)}{G(\lambda)} = \sum_{m \in M_I} \frac{[m]I}{I(\lambda)} m(\lambda) \left(|m| + \frac{D(G_{[m]})(\lambda)}{G_{[m]}(\lambda)} \right)$$

$$\leq \sum_{m \in M_I} \frac{[m]I}{I(\lambda)} m(\lambda) \left(|m| + \frac{D(H_{[m]})(\lambda)}{H_{[m]}(\lambda)} \right) = \frac{D(H)(\lambda)}{H(\lambda)}$$

□

Theorem

$$G \preceq_{ratio} H \implies GI \preceq_{ratio} HI$$

Preliminaries

Definition

Let G and H two formal series with monomial sets M_G and M_H respectively. We define $G \sqcup\!\sqcup H$ the mixing series of G and H as :

$$G \sqcup\!\sqcup H = \sum_{m_G \in M_G} \sum_{m_H \in M_H} \binom{|m_A| + |m_B|}{|m_A|} [m_G]G[m_H]Hm_G m_H$$

and $G \sqcup\!\sqcup_{last} H$ the mixing series except the last letter of G and H as :

$$G \sqcup\!\sqcup_{last} H = \sum_{m_G \in M_G} \sum_{m_H \in M_H} \binom{|m_G| + |m_H| - 1}{|m_G|} [m_G]G[m_H]Hm_G m_H$$

Preliminaries

Theorem

Let G and H two formal series over disjoint sets of respective variables. Then :

$$GH \preceq G \sqcup\!\sqcup H$$

and

$$GH \preceq G \sqcup\!\sqcup_{last} H$$

Proof.

$\forall n \in \mathbb{N}$, the function $f_n : \left(\begin{array}{c} \mathbb{N} \longrightarrow \mathbb{N} \\ m \mapsto \binom{n+m}{m} \end{array} \right)$ is increasing.

□

Second result

Theorem

The coupling time for the "classical" representation of the embedded Markov chain of a queueing network of k independent queues is lower bounded by :

$$\min_{j\in[1,k]} \sum_{i=1}^{k} \frac{D(G_i)(\mathbb{P})}{G_i(\mathbb{P})} + \sum_{\substack{i=1\\ i\neq j}}^{k} \frac{D(L_i)(\mathbb{P})}{L_i(\mathbb{P})} \leq E_{\mathbb{P}}[T]$$

with $G_i = G(p_{2i}, p_{2i-1}), L_i = \sum_{n=0}^{+\infty}(p_{2i-1} + p_{2i})^n = \frac{1}{1-(p_{2i-1}+p_{2i})}$, *G the generating series of one queue and* $\mathbb{P}$ *a probability distribution on* $p_1, \ldots, p_{2k}$.

Proof

If A_i is the event that the last queue to couple is the i th queue, then :

- $[T < +\infty] = \uplus_{i \in [1,k]} A_i$
- $\mathbb{E}_{\mathbb{P}}[T|A_i] = \frac{D(K_i)(\mathbb{P})}{K_i(\mathbb{P})}$ where $K_i = \left(\sqcup_{j \neq i} H_j\right) \sqcup_{last} G_i$ and $H_j = G_j L_j$
- $\left(\prod_{j \neq i} H_j\right) G_i \preceq_{ratio} K_i$

Further research

- Upper bound
- General open queueing networks of finite capacity