ANY MODIFICATION OF MÜLLER'S MARKOV PROCESS IS TRANSIENT

JACOB HILTON AND JÁNOS KRAMÁR

The purpose of this note is to extend a result of Müller [1].

Müller constructs a digraph on of the set of universal prefix computers, and considers the Markov process in which the transition probabilities are uniformly distributed among the outgoing edges. He shows that the Markov process on the set Ξ of all machines is transient, and that the Markov process is not positive recurrent on any set Φ of machines that is invariant under computable output transformations.

Here we show that even if the transition probabilities are allowed to be chosen non-uniformly, then providing they depend only on the structure of the digraph, the Markov process is in fact transient on any such invariant set of machines.

This note is intended to be read only after reading Müller's paper, and we will not recap the introductory material contained there.

Why modify Müller's Markov process?

Müller's Markov process is constructed to penalise computers that are in some sense hard to simulate. However, it is transient thanks to his *Markoff Chaney Virus*, which in some sense it is hard to use to simulate other machines. One might hope that the transition probabilities of the Markov process could be modified in some way to penalise such machines, thereby making it recurrent.

For example, the probabilities could be proportional to the length of the shortest path back to the starting vertex, or to some sum over all such paths, or over all such prefix-free paths. Since the point is to construct a machine-independent measure though, it seems only fair that the transition probabilities should still depend only on the structure of the digraph.

Müller also identifies an obstruction to positive recurrence due to the possibility of applying a computable permutation to the output string, which induces an automorphism of the digraph. As Müller claims, it

This research was conducted at a workshop held in May 2015 at the Machine Intelligence Research Institute.

seems that any "natural" choice of computer set should be invariant under these automorphisms.

Unfortunately, it turns out that this obstruction in fact ensures that any said modification to Müller's Markov process must result in a Markov process that is not only not positive recurrent, but in fact transient.

Basic definitions and statement of main result

Throughout this note we borrow Müller's notation. In particular, Ξ is the set of universal prefix computers. We view Ξ as a digraph, where for all $C, D \in \Xi$ there is a directed edge from C to D if and only if either $C \xrightarrow{0} D$ or $C \xrightarrow{1} D$.

Let Δ be the group of computable permutations of the set $\{0,1\}^*$. For each $\sigma \in \Delta$, define $\Gamma_\sigma : \Xi \to \Xi$ by $\Gamma_\sigma(C) = \sigma \circ C$ in the sense defined by Müller, i.e. the computer obtained by running C and applying σ to the output string. Write $\mathrm{Aut}(\Xi)$ for the group of digraph-automorphisms of Ξ.

The following straightforward result is essentially due to Müller [1, Theorem 4.6].

Lemma 1. *The map $\sigma \mapsto \Gamma_\sigma$ is an injective homomorphism $\Delta \to \mathrm{Aut}(\Xi)$.* □

Given a subgroup $G \leq \Delta$, write $\Gamma_G \leq \mathrm{Aut}(\Xi)$ for the image of G under this map, and for each $C \in \Xi$, write

$$\Gamma_G(C) = \{\Gamma_\sigma(C) : \sigma \in G\}$$

for the orbit of C under Γ_G.

The main result of this note is the following.

Theorem 1. *Consider any Markov process on Ξ with transition probabilities $(p_{CD})_{C,D\in\Xi}$ such that for all $C, D \in \Xi$:*

1. *$p_{CD} > 0$ if and only if either $C \xrightarrow{0} D$ or $C \xrightarrow{1} D$; and*
2. *for all $\Gamma \in \mathrm{Aut}(\Xi)$, $p_{CD} = p_{\Gamma(C)\Gamma(D)}$.*

Let $\Phi \subseteq \Xi$ be a branching, irreducible and aperiodic set such that for all $\sigma \in \Delta$,

$$\{\Gamma_\sigma(C) : C \in \Phi\} = \Phi.$$

Then the induced Markov process on Φ (as defined by Müller) is transient.

Müller essentially proves that this Markov process is not positive recurrent, so this is a slight extension of that result.

For the rest of this note, fix Φ and the Markov process on it as defined in the above result.

Proof of main result

Suppose for contradiction the Markov process in consideration is recurrent. Fix machine $U \in \Phi$ at which to start the process. Given any subgroup $G \leq \Delta$, the process will return to the orbit $\Gamma_G(U)$ with probability 1. Thus there is an induced a Markov process on this orbit and hence on Γ_G. Moreover, if this process is transient then so is the original Markov process.

It turns out that only a very small subgroup $G \leq \Delta$ is required to achieve the result.

Lemma 2. *There is a subgroup $G \leq \Delta$ such that $G \cong \mathbb{Z}^3$.*

Proof. Let α be a computable permutation of $\mathbb{N} \cup \{0\}$ of infinite order, such as the infinite cycle $(\ldots 5\,3\,1\,0\,2\,4\ldots)$. For each $i \in \{0,1,2\}$, define $\sigma_i \in \Delta$ by

$$\sigma_i(s) = \begin{cases} 1^{3\alpha(j)+i}, & \text{if } s = 1^{3j+i} \text{ for some } j \in \mathbb{N} \cup \{0\} \\ s & \text{otherwise,} \end{cases}$$

where 1^k denotes the string of length k consisting of 1s. Then the subgroup of Δ generated by $\{\sigma_0, \sigma_1, \sigma_2\}$ is as required. $\square$

Let $G \leq \Delta$ be as in the above result, and consider the induced Markov process on G and hence on $\mathbb{Z}^3$. Write $(p_{\vec{x}\vec{y}})_{\vec{x},\vec{y}\in\mathbb{Z}^3}$ for the transition probabilities of the induced Markov process on $\mathbb{Z}^3$, and for each $n \in \mathbb{N}$ write $\left(p^{(n)}_{\vec{x}\vec{y}}\right)_{\vec{x},\vec{y}\in\mathbb{Z}^3}$ for the nth step transition probabilities. Note that by property 2, for all $\vec{x},\vec{y} \in \mathbb{Z}^3$,

$$p_{\vec{x}\vec{y}} = p_{\vec{0}\,\vec{y}-\vec{x}}$$

and therefore $\left(p^{(n)}_{\vec{0}\vec{x}}\right)_{\vec{x}\in\mathbb{Z}^3}$ is simply the convolution of $(p_{\vec{0}\vec{x}})_{\vec{x}\in\mathbb{Z}^3}$ with itself n times. Note also that $p_{\vec{x}\vec{y}} > 0$ for all $\vec{x},\vec{y} \in \mathbb{Z}^3$ by property 1.

The following result now completes the proof of Theorem 1. Here we write $p = (p_{\vec{x}})_{\vec{x}\in\mathbb{Z}^3} = (p_{\vec{0}\vec{x}})_{\vec{x}\in\mathbb{Z}^3}$.

Theorem 2. *The induced Markov process on $\mathbb{Z}^3$ is transient.*

Proof. Consider the Fourier transform $\hat{p} : \mathbb{R}^3 \to \mathbb{C}$ defined by $\hat{p}(\vec{\theta}) = \sum_{\vec{x}\in\mathbb{Z}^3} p_{\vec{x}} e^{i\vec{x}\cdot\vec{\theta}}$, which has the property that $\widehat{p^{(n)}}(\vec{\theta}) = \hat{p}(\vec{\theta})^n$. Note that

$|\hat{p}(\vec{\theta})| \leq 1$ with equality if and only if $\vec{\theta} = \vec{0}$. Then, the expected number of visits to the starting state is

$$\begin{aligned} N &= \sum_{n=0}^{\infty} p_{\vec{0}}^{(n)} \\ &= \sum_{n=0}^{\infty} (2\pi)^{-3} \iiint_{-\pi}^{\pi} \widehat{p^{(n)}}(\vec{\theta})\, d\vec{\theta} \\ &= \sum_{n=0}^{\infty} (2\pi)^{-3} \iiint_{-\pi}^{\pi} \hat{p}(\vec{\theta})^n\, d\vec{\theta} \\ &\leq (2\pi)^{-3} \iiint_{-\pi}^{\pi} \frac{d\vec{\theta}}{1 - |\hat{p}(\vec{\theta})|}. \end{aligned}$$

Now for all $\vec{\theta}$, it's the case that $\frac{d^2}{dt^2}\hat{p}(t\vec{\theta})|_{t=0} = -\sum_{\vec{x}\in\mathbb{Z}^3} p_{\vec{x}} \cdot (\vec{x}\cdot\vec{\theta})^2$; this is a continuous negative real-valued function on the unit sphere $|\vec{\theta}| = 1$, so it has a negative maximum value. This implies that for some small $c, \epsilon > 0$, $1 - |\hat{p}(\vec{\theta})| > c|\vec{\theta}|^2$ for all $\vec{\theta}$ in the ball $B := \{\vec{\theta} : |\vec{\theta}| < \epsilon\}$. Hence:

$$\begin{aligned} N &\leq (2\pi)^{-3} \int_{[-\pi,\pi]^3\setminus B} \frac{d\vec{\theta}}{1 - |\hat{p}(\vec{\theta})|} + (2\pi)^{-3} \int_B \frac{d\vec{\theta}}{1 - |\hat{p}(\vec{\theta})|} \\ &\leq \frac{1}{1 - \max_{\vec{\theta}\in[-\pi,\pi]^3\setminus B} |\hat{p}(\vec{\theta})|} + (2\pi)^{-3} \int_B \frac{d\vec{\theta}}{c|\vec{\theta}|^2} \\ &\leq \frac{1}{1 - \max_{\vec{\theta}\in[-\pi,\pi]^3\setminus B} |\hat{p}(\vec{\theta})|} + (2\pi)^{-3} \int_0^{\epsilon} \frac{4\pi r^2\, dr}{cr^2} \\ &< \infty. \end{aligned}$$

Hence, the process is transient. □

References

[1] Markus Müller. Stationary algorithmic probability. *Theoretical Computer Science*, 411(1):113–130, 2010.