Iterating Random Functions on a Finite Set

William M.Y. Goh
MCS Department, Drexel University
Philadelphia, Pa. 19104
wgoh@mcs.drexel.edu

Pawel Hitczenko
MCS Department, Drexel University
Philadelphia, Pa. 19104
phitczen@mcs.drexel.edu

Eric Schmutz
MCS Department, Drexel University
Philadelphia, Pa., 19104
Eric.Jonathan.Schmutz@drexel.edu

November 4, 2002

Abstract

Choose random functions $f_1, f_2, f_3, \ldots$ independently and uniformly from among the n^n functions from $[n]$ into $[n]$. For $t > 1$, let $g_t = f_t \circ f_{t-1} \circ \cdots \circ f_1$ be the composition of the first t functions, and let T be the smallest t for which g_t is constant(i.e. $g_t(i) = g_t(j)$ for all i, j). We prove that, for any positive real number x,

$$\lim_{n\to\infty} \Pr(\frac{T}{n} \leq x) = \int_0^x f(y)dy,$$

where

$$f(y) = \sum_{k\geq 2} (-1)^k e^{-y\binom{k}{2}} (2k-1)\binom{k}{2}.$$

We make our proof available here, but acknowledge that the result is already "well known."

1 Introduction

Let $f_1, f_2, f_3, \ldots$ be a sequence of functions chosen independently and uniformly randomly from the n^n functions on $[n]$. Let $g_1 = f_1$, and for $t > 1$ let $g_t = f_t \circ g_{t-1}$ be the composition of the first t random functions. Define $T(\langle f_i \rangle_{i=1}^{\infty})$ to be the smallest t for which g_t is a constant function. (i.e. $g_t(i) = g_t(j)$ for all $i \neq j$.) This manuscript contains a simple derivation of the asymptotic distribution of T. We had originally intended to publish it in a journal, but we recently learned that the the asymptotic distribution of T is "well known". It was apparently known to Kingman twenty years ago [7],[8],[9], and stronger results are fully proved in Donnelly[3]. There is a lot of related work by Mőhle and others,e.g. [11],[12], and [5].

For $m > 1$, let $\tau_m = |\{t : |Range(g_t)| = m\}|$ be the the number of iterates for which the range has exactly m elements. Thus $T = \sum_{m=2}^{n} \tau_m$. The random variables $\{\tau_m\}_{m=2}^{n}$ are not independent. They are however *conditionally* independent once we specify the set of visited states. Fortunately this set is well behaved, has some convenient properties that enable us to do computations.

Let $\xi = \lfloor \log \log n \rfloor$, and decompose T as $T = T_1 + T_2$, where $T_1 = \sum\limits_{m=2}^{\xi} \tau_m$ and $T_2 = \sum\limits_{m=\xi+1}^{n} \tau_m$. Let $\mathcal{A} = \bigcap\limits_{m=1}^{\xi} [\tau_m > 0]$. The following facts from [2] will be needed (See also Theorem 5 of [9]):

Theorem 1 $\Pr(\mathcal{A}) = 1 - o(1)$, *and* $E(T_2) = o(n)$.

2 Characteristic Function

Let $\lambda_k = \prod\limits_{j=1}^{k-1} (1 - \frac{j}{n})$. Then we have

Theorem 2 $E(e^{itT_1} | \mathcal{A}) = e^{it(\xi-1)} \prod\limits_{k=2}^{\xi} \frac{(1-\lambda_k)}{1-\lambda_k e^{it}}$.

Proof:

Suppose g_{t-1} has an m element range $R = \{r_1, r_2, \ldots, r_m\}$. What is the chance that the next function g_t still has an m element range? On R we have n choices for $f_t(r_1)$, then $n-1$ choices for $f_t(r_2)$ etc. For $x \notin R$, $f_t(x)$ can be chosen arbitrarily. Hence the number of functions f_t for which $g_t = f_t \circ g_{t-1}$ has an an m element range is

$n^{n-m}\prod_{j=0}^{m-1}(n-j)$. Hence

$$\Pr(\tau_m = k|\tau_m > 0) = \lambda_m^{k-1}(1-\lambda_m), \tag{1}$$

and consequently

$$E(e^{it\tau_m}|\tau_m > 0) = \sum_{k=1}^{\infty} \lambda_m^{k-1}(1-\lambda_m)e^{ikt} = \frac{(1-\lambda_m)e^{it}}{1-\lambda_m e^{it}}$$

■

Now let $\phi_n(t) = E(e^{itT_1/n}|\mathcal{A})$ be the characteristic function of the normalized random variable T_1/n on $\mathcal{A}$. Then the following corollary follows immediately from Theorem 2.

Corollary 3 $\phi_n(t) = \prod_{m=2}^{\xi} \frac{(1-\lambda_m)e^{it/n}}{1-\lambda_m e^{it/n}} = e^{it(\xi-1)/n} \prod_{m=2}^{\xi} \frac{(1-\lambda_m)}{1-\lambda_m e^{it/n}}$

Lemma 4 $\phi_n(t) = \prod_{m=2}^{\infty} \frac{\binom{m}{2}}{\binom{m}{2}-it} + o(1)$.

Proof: Note that, for $m \le \xi$,

$$1-\lambda_m = \frac{1}{n}\binom{m}{2} + O(\frac{\xi^4}{n^2}) \tag{2}$$

and

$$1-\lambda_m e^{it/n} = \frac{1}{n}\Big(\binom{m}{2} - it\Big) + O(\frac{\xi^4}{n^2}) \tag{3}$$

Therefore

$$\frac{(1-\lambda_m)e^{it/n}}{1-\lambda_m e^{it/n}} = \frac{\binom{m}{2} + O(\frac{\xi^4}{n})}{\binom{m}{2} - it + O(\frac{\xi^4}{n})} = \frac{\binom{m}{2}}{\binom{m}{2} - it}(1 + O(\frac{\xi^4}{n})). \tag{4}$$

Therefore

$$\phi_n(t) = (1 + O(\frac{\xi^4}{n}))^{\xi} \prod_{m=2}^{\xi} \frac{\binom{m}{2}}{\binom{m}{2} - it}$$

$$= (1 + o(1)) \prod_{m=2}^{\xi} \frac{\binom{m}{2}}{\binom{m}{2} - it}.$$

Finally, note that the infinite product $\prod_{m=2}^{\infty} \frac{\binom{m}{2}}{\binom{m}{2}-it} = \prod_{m=2}^{\infty} \frac{1}{1-\frac{it}{\binom{m}{2}}}$ converges since $\sum \binom{m}{2}^{-1}$ is convergent.

■

3 Simplification

To facilitate inversion, we reexpress the characteristic function ϕ in a more convenient form. Working with the reciprocal, we have

$$\frac{1}{\phi_n(t)+o(1)} = \prod_{k\geq 1}(1-\frac{2it}{k(k+1)}) = \prod_{k\geq 1}\frac{(k-\alpha)(k-\beta)}{k(k+1)}, \tag{5}$$

where $\alpha = \frac{-1-\sqrt{1+8it}}{2}, \beta = \frac{-1+\sqrt{1+8it}}{2}$. It is well known [4] that

$$\frac{1}{\Gamma(z)} = ze^{\gamma z}\prod_{n\geq 1}\Big((1+\frac{z}{n})e^{-z/n}\Big).$$

Since $\alpha+\beta=-1$, the right side of equation (5) becomes

$$\prod_{k\geq 1}\frac{(1-\alpha/k)e^{\alpha/k}(1-\beta/k)e^{\beta/k}}{(1+\frac{1}{k})e^{-1/k}} = \frac{1}{\alpha\beta\Gamma(-\alpha)\Gamma(-\beta)} = \frac{\cos(\frac{\pi}{2}\sqrt{1+8it})}{-2\pi it}.$$

Hence

$$\phi_n(t) = \frac{-2\pi it}{\cos(\frac{\pi}{2}\sqrt{1+8it})} + o(1). \tag{6}$$

4 Fourier Inversion

Inverting, we get the conditional density function f_n for T_1/n:

$$f_n(x) = \frac{1}{2\pi}\int_{-\infty}^{\infty} e^{-itx}(\frac{-2\pi it}{\cos(\frac{\pi}{2}\sqrt{1+8it})} + o(1))dt. \tag{7}$$

Note that $\frac{-2\pi it}{\cos(\frac{\pi}{2}\sqrt{1+8it})}$ has simple poles at $t=-i\binom{k}{2}$, for $k=2,3,4,\ldots$ Since the residue at $t=-i\binom{k}{2}$ is $i(-1)^k(2k-1)\binom{k}{2}e^{-\binom{k}{2}x}$, contour integration yields, for $x>0$, $f_n(x)=f(x)+o(1)$ where

$$f(x) = \sum_{k\geq 2}(-1)^k e^{-\binom{k}{2}x}\binom{k}{2}(2k-1).$$

5 Main Result

For $x>0$, let $F(x)=\int_0^x f(t)dt$. Our main result is

Theorem 5 *For any $x > 0$, $\lim_{n\to\infty} \Pr(T/n \le x) = F(x)$.*

Proof: For any x,

$$\Pr(T/n \le x) \le \Pr(T_1/n \le x)$$
$$= \Pr(T_1/n \le x|\mathcal{A}) \Pr(\mathcal{A}) + \Pr(T_1/n \le x|\mathcal{A}^c) \Pr(\mathcal{A}^c)$$
$$= \Pr(T_1/n \le x|\mathcal{A})(1 + o(1)) + o(1)$$
$$= F(x)(1 + o(1)) + o(1).$$

In the other direction, let ϵ be a fixed but arbitrarily small positive number. Then

$$\Pr(T/n \le x) \ge \Pr(T_1/n \le x - \epsilon \text{ and } T_2/n \le \epsilon)$$
$$\ge \Pr(T_1/n \le x - \epsilon) - \Pr(T_2/n > \epsilon)$$
$$\ge \Pr(T_1/n \le x - \epsilon) - \frac{E(T_2/n)}{\epsilon}$$
$$= F(x - \epsilon) + o(1).$$

The theorem follows from this and the fact that F is continuous. ∎

6 Discussion

Although the ultimate behaviour of our chain is like Kingman's coalescent [7], there are differences. In that process every state is visited, whereas in our process few of the high numbered states are visited. Let $N = \sum_{m=2}^{n} I_{[\tau_m > 0]}$, the number of states visited. In an earlier version of this manuscript, we conjectured that $E(N) \sim \sqrt{2\pi n}$. Robin Pemantle recently proved our conjecture and the corresponding central limit theorem. He may also be able to prove stronger and more general versions of this result, e.g. a functional limit theorem. [14].

Acknowledgement There is a large literature in applied probability that can be traced back to Kingman's work, and it was not immediately obvious to us what is relevant. We are grateful to Simon Tavare for pointing our way to the the work of Donnelly and Mőhle.

References

[1] D.Aldous, Deterministic and Stochastic Models for Coalescence (Aggregation,Coagulation): a Review of Mean Field Theory for Probabilists, *Bernoulli* **5** (1999) 3-48.

[2] A.Dalal and E.Schmutz, Compositions of Random Functions on a Finite Set, *Electronic Journal of Combinatorics*, **9** R26 (2002).

[3] P.Donnelly, Weak convergence to a Markov chain with an entrance boundary: ancestral processes in population genetics, *The Annals of Probability* **19** No.3, (1991) 1102–1117.

[4] Handbook of Mathematical Functions, Abramowitz and Stegun, ed.,Dover Publications,page 255.

[5] J.Fill,On Compositions of Random Functions on a Finite Set, *Preprint.*

[6] R.C.Griffiths, Lines of descent in the diffusion approximation of neutral Wright-Fisher models, *Theoretical Population Biology* **17** (1980) 35–50.

[7] J.F.Kingman, The Coalescent, *Stochastic Proc. Appl.* **13**(1982) 235–248.

[8] J. F. C. Kingman, On the genealogy of large populations. Essays in statistical science. *J. Appl. Probab.* **19A** (1982) 27–43.

[9] J. F. C.Kingman, Exchangeability and the evolution of large populations, in Exchangeability in probability and statistics, pp. 97–112, North-Holland, Amsterdam-New York, 1982.

[10] J. F. C. Kingman,. Mathematics of genetic diversity. CBMS-NSF Regional Conference Series in Applied Mathematics, 34. Society for Industrial and Applied Mathematics (SIAM), Philadelphia, Pa., 1980. ISBN: 0-89871-166-5.

[11] M.Möhle, The time back to the most recent common ancestor in exchangeable population models,preprint submitted for publication.

[12] M.Möhle,Total variation distances and rates of convergence for ancestral coalescenct processes in exchangeable population models, *Adv. Appl. Prob.* **32** (2000)983–993.

[13] M.Möhle, Weak convergence to the coalescent in neutral population models, *J.Appl.Probab.***36**(1999) 446–460.

[14] R.Pemantle, personal communication.

[15] S.Tavare, Line-of-descent and geneological processes and their applications in population genetics models, *Theoretical Population Biology* **26** (1984) 119–164.