ISRAEL JOURNAL OF MATHEMATICS 107 (1998), 327–331

PROCESSES WITH NO STANDARD EXTENSION

BY

MEIR SMORODINSKY*

School of Mathematics, Tel Aviv University
Ramat Aviv, Tel Aviv 69978, Israel
e-mail: meir@math.tau.ac.il

ABSTRACT

Using an example of A. Vershik, a class of processes is introduced with the property that they do not admit a *standard extension*. This provides a simple proof that Vershik's example process does not admit a standard extension.

1. Introduction

Recently, the theory of *Systems of decreasing measurable partitions* which was developed by A. Vershik [4–6] in the late sixties (1967, 1970, 1971) was used to solve a problem in the theory of Brownian motion and stochastic integration [1]. The main ingredient of the paper was a construction of a process with no *standard extension*. The first example of such a process was found by Vershik himself [7]. In this note, I introduce a process which is derived from Vershik's example and prove that it does not admit a standard extension. It follows from this proof that Vershik's example does not admit a standard extension as well. Therefore, this note serves as a simple proof of Vershik's claim.

* Supported in part by The Israel Science Foundation.
Received August 21, 1997

2. Parameterization and standard extension

In this section we introduce the notions of standard extension of Markov processes. The restriction to Markov processes is without loss of generality since every process X_n, $n = 1, 2, \ldots$ can be represented by the Markov process X_n^*, $n = 1, 2, \ldots$, where $X_n^* = (X_n, X_{n+1}, \ldots)$.

Let $U_1, U_2, \ldots$ be an i.i.d. process, where U_1 is uniformly distributed in the unit interval I.

Definition: Let $X' = (X_1', X_2', \ldots)$, where the state space of X_n' is Ξ_n, be a Markov process. A joining of the processes X', U is called a **parameterization** of X' if the following holds for all $n \geq 1$:

(1) U_n is jointly independent of $U_{n+1}, U_{n+2}, \ldots$ and X_{n+1}'.

(2) There exist measurable functions $f_1, f_2, \ldots; f_n : (I, \Xi_{n+1}) \to \Xi_n$ such that

$$X_n' = f_n(U_n, X_{n+1}'), \quad n = 1, 2, \ldots.$$

If, in addition, for each n, $(X_n', X_{n+1}', \ldots)$ is $(U_n, U_{n+1}, \ldots)$-measurable, the parameterization is said to be **generating** and the processes are said to admit **a standard extension**.

Notice that every independent process admits a standard extension and that a process that admits a standard extension is **tail trivial**. Also, if X_n, $n = 1, 2, \ldots$ admits a standard extension, so does every subsequent process X_{k_n}, $n = 1, 2, \ldots$. M. Rosenblatt [2, 3] proved that every countable state, stationary mixing Markov chain admits a standard extension.

3. The examples

Example: Vershik [7]. It is possible to construct a Markov chain $Y_1, Y_2, \ldots$ such that each Y_n is distributed uniformly on $[0, 1]$. When $Y_{n+1}, \ldots$ are given, the conditional distribution of Y_n depends on Y_{n+1} in the following manner: if Y_{n+1} takes the value y, which we write in binary as $.y_1y_2y_3\cdots$, then Y_n is equal to either $.y_1y_3y_5\cdots$ or $.y_2y_4y_6\cdots$, each with probability $1/2$. Let $C = \{0, 1\}$. Out of the process Y_n we produce the (discrete state) process X_n as follows. Each X_n is a function of Y_n only. X_1 assumes values in C according to the first binary digit of Y_1. X_2 assumes values in C^2 according to the first 2 binary digits of Y_2. In general, X_n assumes values in C^{2^n} according to the first 2^n digits of Y_n. The resulting process is obviously a Markov process.

Abstractly we have the following class V of Markov chains.

A CLASS OF EXAMPLES. Let $X = (X_1, X_2, \ldots)$ be a process with a Markov distribution defined as follows: C is a finite set (colors), $|C| = K$. X_1 assumes values from C with uniform distribution. X_2 assumes as values the pair (ξ_1, η_1), where each ξ_1 and η_1 are values from the same set of values of X_1, i.e. from C. The probability distribution of X_2 is uniform on C^2. Inductively, X_n will assume as values pairs (ξ_{n-1}, η_{n-1}), where ξ_{n-1} and η_{n-1} are values from the same set of values of X_{n-1} and with the product of the same measure. So, actually X_n assumes values from the set Ξ_n of 2^n-tuples of colors with uniform distribution. The transition from X_n to X_{n-1} is with probability 1/2 to ξ_{n-1} and 1/2 to η_{n-1}, provided $\xi_{n-1} \neq \eta_{n-1}$. Otherwise, X_{n-1} assumes the common value with probability 1.

THEOREM 1: *Each process $X \in V$ is tail-trivial.*

Proof: Since X is Markov, it suffices to show that for each $n \geq 1$, $P[X_n = x^{(n)} | X_{n+m}]$ converges in probability to $P[X_n = x^{(n)}]$ as m goes to infinity, for each $x^{(n)} \in \Xi_n$. Any $x^{(n+m)} \in \Xi_{n+m}$ can be viewed as a sequence of length 2^m of 2^n-tuples. It is evident that $P[X_n | X_{n+m} = x^{(n+m)}]$ is uniformly distributed over the 2^n-tuples which comprise $x^{(n+m)}$. Hence the above required convergence follows from the weak law of large numbers.

4. The main result

THEOREM 2: *No process $X \in V$ admits a standard extension.*

Before proceeding with the proof of Theorem 2 we shall introduce some more notation and definitions.

Since a point $x \in \Xi_n$ is a 2^{n-1}-tuple of colors from C, $x = (x_1, \ldots, x_{2^{n-1}})$, $x_i \in C$. Each index i can be represented by a sequence of length $n-1$, $\omega(i) = (\omega_1(i), \ldots, \omega_{n-1}(i))$ of zeros and ones, the binary expansion of i. The set Ω_n of the binary sequences of length $n-1$ has a natural tree structure. Consider A_n, the set of the tree automorphisms of Ω_n. A surjective map a is a tree automorphism, if it maps surjectively the set of all sequences beginning with zero either onto itself or onto the set of all sequences beginning with one, each of these maps being a tree map on Ω_{n-1}. The cardinality of A_n is 2^{2^n-1}. We define an n-distance $n \geq 0$, between any 2 points $x, x' \in \Xi_n$, as follows.

Definition: $d_n(x, x') = \min_{a \in A_n} 1/2^{n-1} \#\{i | x_i \neq x'_{a(i)}\}$.

Let $x = (\xi, \eta)$, $x' = (\xi', \eta')$, observe the following recursive relation:

$$(1) \qquad d_n(x, x') = \tfrac{1}{2} \min\{d_{n-1}(\xi, \xi'\} + d_{n-1}(\eta, \eta'); d_{n-1}(\xi, \eta') + d_{n-1}(\eta, \xi')\}.$$

Next, consider a given parameterization p and the functions f_n associated with it. Let us define the functions $g_n : (I^{n-1}, \Xi_n) \to \Xi_1$, $n > 0$, by putting $g_1 = f_1$ and inductively

$$g_n(u_1, \ldots, u_n, x_{n+1}) = g_{n-1}(u_0, \ldots, u_{n-2}, f_n(u_{n-1}, x_n)).$$

Verify by induction that for all n

$$X_1 = g_n(U_1, \ldots, U_n, X_{n+1}).$$

Now we define a new n-distance, associated with p, on the pairs (x, x'):

Definition: $d_1^p = d_1$, and

$$d_n^p(x, x') = P[g_n(U_1, \ldots, U_{n-1}, x) \neq g_n(U_1, \ldots, U_{n-1}, x')], \quad n > 1.$$

Let $c = P[f(U_{n-1}, x) = \xi; f(U_{n-1}, x') = \xi']$; observe the following recursive relation:

$$(2) \quad d_n^p(x, x') = c\{d_{n-1}^p(\xi, \xi') + d_{n-1}^p(\eta, \eta')\} + (\tfrac{1}{2} - c)\{d_{n-1}^p(\xi, \eta') + d_{n-1}^p(\eta, \xi')\}.$$

LEMMA 1: *For any n, (x, x') and parameterization p,*

$$d_n^p(x, x') \geq d_n(x, x').$$

Proof: Use the recursive relations (1) and (2) to induct on n.

For an integer m and $\epsilon > 0$, consider the function

$$\varphi(m, \epsilon) = \sum_{i=0}^{[m\epsilon]} \binom{m}{i}.$$

Put $m = 2^{n-1}$. Using this notation we have the following estimate for any $x' \in \Xi_n$:

$$\#\{x \in \Xi : d_n(x, x') < \epsilon\} < 2^m \varphi(m, \epsilon) K^{m\epsilon}.$$

Known estimates on the binomial coefficients imply the existence of a function $h(\epsilon)$ such that

(a) $\varphi(m, \epsilon) < 2^{mh(\epsilon)}$,
(b) $h(\epsilon) \to 0$ as $\epsilon \to 0$.

LEMMA 2: *Given any fixed point $x \in \Xi_n$,*

$$P[d_n^p(x, X_n) < \epsilon] < 2^{m(1+h(\epsilon))} K^{-m(1-\epsilon)}.$$

Proof: Use Lemma 1 and the estimates of the cardinality of the sets Ξ_n, A_n.

Proof of Theorem 2: Assume first that $K > 2$. Suppose p is a generating parameterization. In particular, X_1 has to be $(U_1, U_2, \ldots)$-measurable. This implies that for given $\epsilon > 0$ there is an n and sets E_n and G_n such that E_n is $(U_1, \ldots, U_n)$-measurable, $P[E_n] > 3/4$ and G_n is X_n-measurable; $P(G_n) > 3/4$ and for any $(y_1, \ldots, y_n) \in E_n$, $x, x' \in G_n$,

$$g_n(y_1, \ldots, y_n, x) = g_n(y_1, \ldots, y_n, x').$$

That would imply the existence of a point $x \in \Xi_n$ such that

$$P[d_n^p(x, X_n) < \epsilon] > 1/2.$$

This contradicts Lemma 2, since for $\epsilon > 0$ sufficiently small the above probability will be less than 1/2.

Finally, in the case $K = 2$, consider X_2 instead of X_1. Under the assumption that p generates, X_2 must be generated by $U_2, U_2, \ldots$. Since $|\Xi_1| = 4$, the same reasoning as above will imply a contradiction.

Reference

[1] L, Dubins, J. Feldman, M. Smorodinsky and B. Tsirelson, *Decreasing sequence of σ-fields and a measure change for Brownian motion*, The Annals of Probability **24** (1996), 882–904.

[2] M. Rosenblatt, *Stationary processes as shifts of functions of independent random variables*, Journal of Mathematical Mechanics **8** (1959), 665–681.

[3] M. Rosenblatt, *Stationary Markov chains and independent random variables*, Journal of Mathematical Mechanics **9** (1960), 945–949.

[4] A. M. Vershik, *A theorem on lacunary isomorphisms*, Functional Analysis and its Applications **2** (1968), 200–203.

[5] A. M. Vershik, *Decreasing sequences of measurable partitions, and their applications*, Soviet Mathematics Doklady **11** (1970), 1007–1011.

[6] A. M. Vershik, *A continuum of pairwise nonisomorphic dyadic sequences*, Functional Analysis and its Applications **5** (1971), 16–18.

[7] A. M. Vershik, *Approximation in measure theory*, Dissertation, Leningrad University, 1973 (in Russian).