Finite Factors of Bernoulli Schemes and Distinguishing Labelings of Directed Graphs

Andrew Lazowski
Department of Mathematics
Sacred Heart University
Fairfield, CT 06825
lazowskia@sacredheart.edu

Stephen M. Shea
Department of Mathematics
St. Anselm College
Manchester, NH 03102
sshea@anselm.edu

Submitted: May 21, 2011; Accepted: Dec 21, 2011; Published: Jan 6, 2012
Mathematics Subject Classifications: 05C20, 60G10
Keywords: Bell numbers; Bernoulli scheme; directed graph; distinguishing number; finitarily Markovian; Markov; variable-length Markov

Abstract

A labeling of a graph is a function from the vertices of the graph to some finite set. In 1996, Albertson and Collins defined distinguishing labelings of undirected graphs. Their definition easily extends to directed graphs. Let G be a directed graph associated to the k-block presentation of a Bernoulli scheme X. We determine the automorphism group of G, and thus the distinguishing labelings of G. A labeling of G defines a finite factor of X. We define demarcating labelings and prove that demarcating labelings define finitarily Markovian finite factors of X. We use the Bell numbers to find a lower bound for the number of finitarily Markovian finite factors of a Bernoulli scheme. We show that demarcating labelings of G are distinguishing.

1 Introduction

A process is a quadruple $(X, \mathcal{U}, \mu, T)$ where X is the set of doubly infinite sequences on some alphabet A, $\mathcal{U}$ is the σ-algebra generated by the coordinates, μ is a shift invariant probability measure on $(X, \mathcal{U})$, and T is the left shift by one. When we refer to a process X, we are referring to this quadruple. A process is a Bernoulli scheme if $\mu = p^{\mathbb{Z}}$ for some probability vector p. We will need the following definitions.

Definition 1.1. *Let k be a positive integer. The process $X^{(k)}$ called the k-stringing (or k-block presentation) of X is defined as follows. The state space of $X^{(k)}$ is all allowable sequences of length k in X, and $X_n^{(k)} = (X_n, X_{n+1}, \ldots, X_{n+k-1})$ $(n \in \mathbb{Z})$.*

Definition 1.2. *Let $a \in A$ where $P(X_n = a) > 0$. We say a is a renewal state of X if the σ-algebras $\mathcal{U}(X_{n+1}, X_{n+2}, \dots)$ and $\mathcal{U}(\dots, X_{n-2}, X_{n-1})$ are independent given the event $[X_n = a]$. If there exists such an a, we say X is a renewal process.*

Definition 1.3. *A process X is Markov if for all $a \in A$ where $P(X_n = a) > 0$, a is a renewal state in X. A process Y is k-step Markov if $Y^{(k)}$ is Markov. A process Z is finitarily Markovian if for some k, $Z^{(k)}$ is a renewal process.*

In 1983, Rissanen [19] introduced the class of models we are calling finitarily Markovian processes. These processes have gone by several names including variable length Markov chains and Markov chains of variable order. The terminology finitarily Markovian comes from [15].

Let $(X, \mathcal{U}, \mu, T)$ and $(Y, \mathcal{V}, \nu, S)$ be two processes. A factor map f from $(X, \mathcal{U}, \mu, T)$ to $(Y, \mathcal{V}, \nu, S)$ is a measurable equivariant map from a subset of X of full measure to a subset of Y of full measure which takes μ to ν. A factor map $f : X \to Y$ is a k-block factor if for $k \in \mathbb{Z}^+$, $x \in X$, $(f(x))_1 = (f(x_1, x_2, \dots, x_k))_1$. In other words, f is a (not necessarily injective) labeling of the k-blocks in X. A k-block factor of X is a 1-block factor of $X^{(k)}$. If the precise k is not of interest, we will say f is a finite factor. Let $X = \{0, 1\}^{\mathbb{Z}}$ with measure $\mu = (p_0, p_1)^{\mathbb{Z}}$. We are in search of necessary and sufficient conditions for a finite factor of a 2-state Bernoulli scheme to be finitarily Markovian. In the next section we discuss our motivation for studying this problem. Also in Section 2, we present two finite factors of 2-state Bernoulli schemes, one that is finitarily Markovian and one that is not. These examples demonstrate that the conditions which may be necessary when $p_0 \neq p_1$ may not be necessary in the special case that $p_0 = p_1$. When $p_0 \neq p_1$, we conjecture that a combinatorial condition on f is necessary and sufficient. If f is viewed as a labeling of a directed graph G_k associated with $X^{(k)}$, then the condition is that f be r-distinguishing [2]. In Section 3 we define r-distinguishing and determine the r-distinguishing labelings of G_k. In Section 4 we define r-demarcating labelings. We show that r-demarcating is sufficient for the factor to be finitarily Markovian. In Section 5, we show that r-demarcating implies r-distinguishing. In Section 6 we use Bell numbers to count the number of finite factors of a Bernoulli scheme in order to show that most finite factors are finitarily Markovian. We end with some remarks and a discussion of future directions.

2 Motivation and Examples

2.1 Classifying ergodic processes

After Kolmogorov introduced entropy to dynamical systems [10], it was hypothesized that entropy is a complete isomorphism invariant for Bernoulli schemes. Ornstein later proved this conjecture [16]. However, prior to Ornstein's result, mathematicians began trying to construct isomorphisms between various independent processes with the same entropy. Meshalkin was one of these mathematicians and in 1959 he showed

that Bernoulli schemes with non-isomorphic state spaces can be isomorphic [14]. His results would later be expanded by Blum and Hanson [4]. Meshalkin, however, not only constructed an isomorphism, but a finitary isomorphism. Isomorphism and finitary isomorphism are formally defined as follows [8]. We will use the shorthand $x[m,n]$ to mean $x_m x_{m+1} \dots x_n$.

Definition 2.1. *Let $(X,\mathcal{U},\mu,T)$ and $(Y,\mathcal{V},\nu,S)$ be two processes. An isomorphism ϕ from $(X,\mathcal{U},\mu,T)$ to $(Y,\mathcal{V},\nu,S)$ is a bimeasurable equivariant map from a subset of X of measure one to a subset of Y of measure one which takes μ to ν. The isomorphism ϕ is finitary if for almost every $x \in X$ there exists integers $m \leqslant n$ such that the zero coordinates of $\phi(x)$ and $\phi(x')$ agree for almost all $x' \in X$ with $x[m,n] = x'[m,n]$, and similarly for ϕ^{-1}. If we drop the requirement that ϕ be invertible, we say ϕ is a finitary factor.*

Keane and Smorodinsky improved on Ornstein's result with the following theorem.

Theorem 2.2. *Entropy is a complete finitary isomorphism invariant for Bernoulli schemes [8].*

The theory of finitary isomorphisms found in [1], [7], [8], [9], [21], [22] and [28] has paralleled the theory of measure-theoretic isomorphisms as outlined by Ornstein and Sinai in [16], [17] and [27]. There are, however, results on the measure-theoretic side that are noticeably missing in the finitary theory. For instance, we know that all factors of Bernoulli schemes are measure-theoretically isomorphic to Bernoulli schemes [16]. The following finitary equivalent remains unresolved.

Conjecture 2.3. *Entropy is a complete finitary isomorphism invariant for finitary factors of Bernoulli schemes.*

The question of whether all finitary factors are finitarily isomorphic to a Bernoulli scheme appeared in [21]. It was later conjectured in [28]. The conjecture is reiterated in a recent survey of finitary codings [23]. This survey also serves as a nice introduction to the study of finitary isomorphisms.

In the proof of Theorem 2.2, Keane and Smorodinsky used what is commonly referred to as the marker and filler technique. In [25] and [26], this technique is described in detail. One might hope that the marker and filler methods could be extended to prove the above conjecture. In [24] and [25], it is shown that r-processes would play a role in any such extension.

Definition 2.4. *We say a renewal state $a \in A$ has n-Bernoulli distribution if for some nonnegative integer n, $P[X_{n'} = a | X_0 = a] = P[X_{n'} = a]$ for all $n' > n$. An r-process X is a renewal process such that a renewal state in X has n-Bernoulli distribution.*

A process X is m-dependent if the σ-algebras $\mathcal{U}(X_{m+1}, X_{m+2}, ...)$ and $\mathcal{U}(..., X_{-1}, X_0)$ are independent. Finite factors of Bernoulli schemes are m-dependent [28]. If a renewal process is m-dependent, then clearly any renewal state in that process has m-Bernoulli distribution. So, if a finite factor is a renewal process, it is an r-process. If a finite factor is finitarily Markovian, it is continuously isomorphic to to an r-process.

We will show through our work in Sections 3 and 4, that most finite factors of Bernoulli schemes are finitarily Markovian. We also give a sufficient condition on the factor map for the finite factor to be finitarily Markovian. This gives us a way to construct a large class of processes that are r-processes or are continuously isomorphic to r-processes. Since r-processes may be pivotal in resolving the above conjecture, this construction is useful.

2.2 Two examples

We now present two examples. Since it is nontrivial to find a finite factor of a Bernoulli scheme that is not finitarily Markovian, we begin with such an example.

Example 2.5. *Let $X = \{0,1\}^{\mathbb{Z}}$ with measure $\mu = (p_0, p_1)^{\mathbb{Z}}$ where $p_0 \neq p_1$. Let $f : X \to Y$ where $(f(x))_i = (x_i + x_{i+1})$ mod 2. Then Y is not finitarily Markovian.*

Proof. Let $y \in Y$. For any positive integer n let $R_n = y[-n,-1]$. There are two words of length $n+1$ in X that map to R_n under f. Let A_n and B_n be these two words. So, if $f(x)[-n,-1] = R_n$, then $0 < p(x[-n,0] = A_n) < 1$ and $0 < p(x[-n,0] = B_n) < 1$. We will refer to A_n and B_n as the possible pre-images of R_n. Notice that A_n and B_n are duals of each other. That is, A_n is obtained from B_n by replacing every 0 with a 1 and every 1 with a 0. Let $A_n(0) = \text{card}\{x_i | -n \leqslant i \leqslant 0, x_i = 0 \text{ and } x[-n,0] = A_n\}$. Let $B_n(0) = \text{card}\{x_i | -n \leqslant i \leqslant 0, x_i = 0 \text{ and } x[-n,0] = B_n\}$. Then $B_n(0) = n - A_n(0)$, and as $n \to \infty$, either

$$\text{(i) } \frac{A_n(0)}{n+1} \to p_0 \text{ and } \frac{B_n(0)}{n+1} \to p_1 \text{ or (ii) } \frac{B_n(0)}{n+1} \to p_0 \text{ and } \frac{A_n(0)}{n+1} \to p_1.$$

Since $X = \{0,1\}^{\mathbb{Z}}$ with measure $\mu = (p_0, p_1)^{\mathbb{Z}}$ and $p_0 \neq p_1$, in case (i), we obtain that A_n is the true pre-image of R_n and in case (ii), B_n is the true pre-image of R_n.

Suppose $y_0 = 0$. Let x be such that $f(x) = y$. Then $x_0x_1 = 00$ or $x_0x_1 = 11$. Also, $y_1 = 0$ if and only if $x_0x_1x_2 = 000$ or $x_0x_1x_2 = 111$. Then we have $p[y_1 = 0 | x_0x_1 = 11] = p_1$ and $p[y_1 = 0 | x_0x_1 = 00] = p_0$. Since only one of B_n and A_n can be the true pre-image of R_n, either $\lim_{n\to\infty} p[y_1 = 0|R_n] = p_1$ or $\lim_{n\to\infty} p[y_1 = 0|R_n] = p_0$. Since for any finite n, both A_n and B_n map to R_n, the true probability that $y_1 = 0$ cannot be determined from R_n for any finite n. Therefore, Y is not finitarily Markovian. □

For our next example, again let $X = \{0,1\}^{\mathbb{Z}}$ with measure $\mu = (p_0, p_1)^{\mathbb{Z}}$, and let $f : X \to Y$ where $(f(x))_i = (x_i + x_{i+1})$ mod 2. Now suppose $p_0 = p_1 = \frac{1}{2}$. In this case, it is well-known that $X = Y$. If one follows the proof of Example 2.5, we see that the pre-images A_n and B_n still exist, but both converge to the correct distribution.

3 Distinguishing Labelings of Directed Graphs

3.1 Definition and a conjecture

Suppose you have a key ring with a finite number of seemingly indistinguishable keys. How can we label the keys so as to distinguish each from all of the others? Surely, labeling each key with a unique label is sufficient, but how many distinct labels are necessary? Motivated by this question, Albertson and Collins defined the r-distinguishing number of an undirected graph as follows [2]. Recall that a graph automorphism is a permutation of the vertices that preserves edge-connectivity.

Definition 3.1. *A labeling of the vertices of a graph G, $f : V(G) \to \{1, 2, \ldots, r\}$, is r-distinguishing if the only automorphism of the graph that preserves all of the vertex labels is the identity. The distinguishing number of a graph G, denoted by $D(G)$, is the minimum r such that G has an r-distinguishing labeling.*

While Albertson and Collins were primarily interested in undirected graphs, their definition applies to directed graphs (digraphs) as well. This was done in [11], for example. We will take our digraphs G to be defined by a finite set of vertices $V(G)$ and a finite set of directed edges $E(G)$. For a vertex $v \in V(G)$, we allow for $(v, v) \in E(G)$ (self-loops). For $u, v \in V(G)$, we allow for at most one edge $(u, v) \in E(G)$. We will assume our graph to be irreducible (i.e. for any u and v in $V(G)$ there exists a walk in G from u to v). For a more thorough introduction to the theory of graphs and directed graphs, we recommend [5].

In the first section we claim that the k-block factor of a process X can be defined by a labeling of the digraph G_k associated with $X^{(k)}$. To make the definition of G_k more formal, let X be a Bernoulli scheme on two symbols 0 and 1. For each $k \in \mathbb{Z}^+$, let G_k be the directed graph where $V(G)$ is the state space of $X^{(k)}$, and $(u, v) \in E(G)$ if and only if $p[X_1^{(k)} = v | X_0^{(k)} = u] > 0$. So, for $k = 1$ we have the graph in Figure 1. G_2 is the graph in Figure 2.

Figure 1

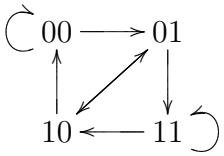

Figure 2

Let f be a labeling on some G_k. Then f defines a 1-block factor of $X^{(k)}$ (or a k-block factor of X). We will refer to the finite factor map defined by f with f as well, writing $f : X \to Y$. Then Y has measure $f(\mu)$ (where μ is the measure on X). We conjecture the following.

Conjecture 3.2. *Let $X = \{0, 1\}^{\mathbb{Z}}$ with measure $\mu = (p_0, p_1)^{\mathbb{Z}}$ where $p_0 \neq p_1$. Let $k \in \mathbb{Z}^+$, and let f be a labeling of G_k. Let $f : X \to Y$ denote the finite factor map defined by f. Then Y is finitarily Markovian if and only if f is r-distinguishing.*

3.2 The automorphism group of G_k

To understand when a labeling of G_k is distinguishing, we must first understand the automorphism group of G_k.

The adjacency matrix of a graph G with $V(G) = \{1, 2, \ldots, n\}$ is an $n \times n$ matrix A where entry $A_{ij} = 1$ if $(i, j) \in E(G)$ and $A_{ij} = 0$ otherwise. If p is a permutation of $V(G)$, the permutation matrix P corresponding to p is an $n \times n$ matrix where the entry $P_{ij} = 1$ if $p(i) = j$ and $P_{ij} = 0$ otherwise.

We will always list the vertices of G_k in numeric order when constructing A. For example, the adjacency matrix of G_3 is the following where vertex 1 is 000, vertex 2 is 001, etc. Note that vertex i is the binary representation of $i - 1$.

	000	001	010	011	100	101	110	111
000	1	1	0	0	0	0	0	0
001	0	0	1	1	0	0	0	0
010	0	0	0	0	1	1	0	0
011	0	0	0	0	0	0	1	1
100	1	1	0	0	0	0	0	0
101	0	0	1	1	0	0	0	0
110	0	0	0	0	1	1	0	0
111	0	0	0	0	0	0	1	1

In this section we determine the automorphisms of G_k. For ease of exposition we split the proof into smaller lemmas. We begin with the following well-known result [12].

Lemma 3.3. *Let G be an irreducible digraph. Let p be a permutation of $V(G)$. Let P be the permutation matrix of p. Let A be the adjacency matrix of G. Then p is an automorphism of G if and only if $PA = AP$.*

Each symbol in $X^{(k)}$ (vertex of G_k) is a word of length k in X. Let $a = a_1 a_2 \ldots a_k$ be a word in X. We define the dual of a, denoted $\hat{a} = \hat{a}_1 \hat{a}_2 \ldots \hat{a}_k$ to be the word in X such that $a_i \neq \hat{a}_i$ for $1 \leqslant i \leqslant k$. For example, the dual of 001 is 110. We will call the permutation p where $p(a) = \hat{a}$ for all $a \in G_k$, the dual permutation.

Lemma 3.4. *Let A be the adjacency matrix of G_k. For all i where $1 \leqslant i \leqslant 2^k$, vertex i is the dual of vertex $2^k + 1 - i$.*

Proof. Recall that i is the binary representation of $i - 1$. Then $2^k + 1 - i$ is the binary representation of $2^k - i$. We will use induction.

When $i = 1$, $i = 0^k$ and $2^k + 1 - i = 1^k$. These are duals. We now have to show that if (in binary) $a = a_1 a_2 \ldots a_k$ is the dual of $\hat{a} = \hat{a}_1 \hat{a}_2 \ldots \hat{a}_k$, that $a_1 a_2 \ldots a_k + 1$ is the dual of $\hat{a}_1 \hat{a}_2 \ldots \hat{a}_k - 1$.

Let $L = \{l' | 1 \leqslant l' \leqslant k \text{ and } a_{l'} = 0\}$. Let $l = \max\{l' \in L\}$. Then $a + 1 = a_1 a_2 \ldots a_{l-1} \hat{a}_l \hat{a}_{l+1} \ldots \hat{a}_k$. Let $M = \{m' | 1 \leqslant m' \leqslant k \text{ and } \hat{a}_{m'} = 1\}$. Since a is the dual of $\hat{a}$, $l = \max\{m' \in M\}$. Then $\hat{a} - 1 = \hat{a}_1 \hat{a}_2 \ldots \hat{a}_{l-1} a_l a_{l+1} \ldots a_k$. Thus, $a + 1$ is the dual of $\hat{a} - 1$.

By induction, $\hat{i} = 2^k + 1 - i$ for $1 \leqslant i \leqslant n$. □

The counteridentity matrix is 1 on the antidiagonal, and 0 elsewhere.

Lemma 3.5. *Let A be the adjacency matrix of G_k. Let Q be the permutation matrix of the identity permutation. Let Q' be the permutation matrix of the dual permutation. Then Q is the identity matrix and Q' is the counteridentity matrix.*

Proof. Clearly, Q is the identity matrix. Note that $Q'_{ij} = 1$ if and only if $j = \hat{i}$. The graph G_k has 2^k vertices. If C is the counteridentity matrix, then $C_{ij} = 1$ if and only if $j = 2^k + 1 - i$. Then $C = Q'$ if and only if $\hat{i} = 2^k + 1 - i$. This follows from Lemma 3.4. Therefore, Q' is the counteridentity. □

Lemma 3.6. *Let A be the adjacency matrix of G_k. Let I be the identity matrix, and let C be the counteridentity. Then I and C commute with A.*

Proof. Clearly I commutes. Note that C is its own inverse. Then $AC = CA$ if and only if $A = CAC^{-1} = CAC$. For each i and j where $1 \leqslant i \leqslant 2^k$ and $1 \leqslant j \leqslant 2^k$, let $l = 2^k + 1 - i$ and let $m = 2^k + 1 - j$. It follows that $A = CAC$ if and only if $A_{ij} = A_{lm}$ for all i and j. This is because CAC swaps columns i and $2^k + 1 - i$ and rows j and $2^k + 1 - j$ in A.

We need to show $(i, j) \in E(G_k)$ if and only if $(2^k+1-i, 2^k+1-j) \in E(G)$. Suppose $(i, j) \in E(G_k)$. Let $i = a = a_1a_2 \dots a_k$ and $j = b = b_1b_2 \dots b_k$. Since $(i, j) \in E(G_k)$, $a_{t+1} = b_t$ for $1 \leqslant t \leqslant k-1$. So, $j = a_2a_3 \dots a_kb_k$. By Lemma 3.4, i is the dual of $2^k + 1 - i$ and j is the dual of $2^k + 1 - j$. So, $2^k + 1 - i = \hat{a}_1\hat{a}_2 \dots \hat{a}_k$ and $2^k + 1 - j = \hat{a}_2\hat{a}_3 \dots \hat{a}_k\hat{b}_k$. Therefore, $(2^k + 1 - i, 2^k + 1 - j) \in E(G_k)$. A similar argument works to show that if $(2^k + 1 - i, 2^k + 1 - j) \in E(G_k)$, then $(i, j) \in E(G_k)$.

□

A walk in a directed graph G between distinct vertices u and v that traverses l edges is minimal if there does not exist a walk in G from u to v that traverses with less than l edges.

Lemma 3.7. *Let u and v be distinct vertices in G_k. There exists a unique minimal walk in G_k from u to v.*

Proof. Let $u = u_1u_2 \dots u_k$ and let $v = v_1v_2 \dots v_k$. We begin with the "no overlap" case, meaning there is no word at the end of u that begins v. That is, there does not exist an $i' \in [1, k-1]$ such that $u_{i'+j} = v_j$ for all $1 \leqslant j \leqslant k - i'$. Then (using parenthesis for readability)

$$u = (u_1u_2 \dots u_k), (u_2u_3 \dots u_kv_1), \dots$$
$$(u_kv_1v_2 \dots v_{k-1}), (v_1v_2 \dots v_k) = v$$

is the unique minimal walk. Otherwise, let $I = \{i' \in [1, k-1] | u_{i'+j} = v_j \text{ for all } 1 \leqslant j \leqslant k - i'\}$ and let $i = \min\{i' \in I\}$. Then (using parenthesis for readability)

$$u = (u_1u_2 \dots u_k), (u_2u_3 \dots u_kv_{i+1}), \dots$$
$$(u_iu_{i+1} \dots u_kv_{i+1}v_{i+2} \dots v_{k-1}), (u_{i+1}u_{i+2} \dots u_kv_{i+1}v_{i+2} \dots v_k) = v$$

is the unique minimal walk. □

Lemma 3.8. *Let p be an automorphism of G_k. Let u and v be distinct vertices of G_k. Let $w = u, a_1, a_2, \ldots, a_n, v$ be the unique minimal walk from u to v. If $p(u) = u$ and $p(v) = v$, then $p(a_i) = a_i$ for all i where $1 \leqslant i \leqslant n$.*

Proof. Since p preserves edges in G_k, p must preserve walks in G_k. If p fixes u and v, then any walk from u to v on l edges must be mapped to a walk from u to v on l edges. Suppose w is a walk on l edges. Since w is the only walk from u to v on l edges, $p(w) = w$. □

Lemma 3.9. *Let p be an automorphism of G_k. Let u and v be distinct vertices of G_k. Let $w = u, a_1, a_2, \ldots, a_n, v$ be the unique minimal walk from u to v. If $p(u) = \hat{u}$ and $p(v) = \hat{v}$, then $p(a_i) = \hat{a}_i$ for all i where $1 \leqslant i \leqslant n$.*

Proof. By Lemma 3.7, there exists a unique minimal walk from u to v and a unique walk from $\hat{u}$ to $\hat{v}$.

Let $w = u, a_1, a_2, \ldots, a_n, v$ be the unique minimal walk from u to v. The dual of w is $\hat{w} = \hat{u}, \hat{a}_1, \hat{a}_2, \ldots, \hat{a}_n, \hat{v}$. This is a walk from $\hat{u}$ to $\hat{v}$. Suppose there exists a walk from $\hat{u}$ to $\hat{v}$ on less than $n+1$ edges. Then for some n' where $n' < n$, there exists a walk $w' = \hat{u}, b_1, b_2, \ldots, b_{n'}, \hat{v}$. The dual of w' is $u, \hat{b}_1, \hat{b}_2, \ldots, \hat{b}_{n'}, v$. This walk from u to v is on less than $n+1$ edges. This contradicts the minimality of w. Therefore, w and $\hat{w}$ are the unique minimal walks from u to v and $\hat{u}$ to $\hat{v}$ respectively, and they are both on $n+1$ edges.

Since p preserves walks in G_k, p must map w to a walk on $n+1$ edges from $\hat{u}$ to $\hat{v}$. Since $\hat{w}$ is the unique minimal walk on $n+1$ edges from $\hat{u}$ to $\hat{v}$, $p(w) = \hat{w}$. Since $\hat{w}$ is the dual of w, $p(a_i) = \hat{a}_i$ for all i where $1 \leqslant i \leqslant n$. □

Lemma 3.10. *If p is an automorphism of G_k such that $p(0^k) = 0^k$ and $p(1^k) = 1^k$, then p is the identity.*

Proof. Suppose $p(0^k) = 0^k$ and $p(1^k) = 1^k$. By Lemma 3.7, there exists a unique minimal walk from 0^k to 1^k and from 1^k to 0^k. By Lemma 3.8, $p(a) = a$ for all vertices a on these two walks.

Let $W_1 = \{v = v_1 v_2 \ldots v_k \in V(G_k) | v_1 = 1\}$. For each $v \in V(G_k)$, let $t(v) = \text{card}\{i, i+1 | 1 \leqslant i \leqslant k-1 \text{ and } v_i v_{i+1} = 01 \text{ or } v_i v_{i+1} = 10\}$. Here "card" means cardinality. We will prove by induction on t, that p fixes all vertices $v \in W_1$.

Let $v \in W_1$ such that $t(v) = 1$. Then $v = 1^{m_0} 0^{m_1}$ where $m_0 \geqslant 1$, $m_1 \geqslant 1$, and $m_0 + m_1 = k$. Then v is on the unique minimal walk from 1^k to 0^k. By Lemma 3.8, $p(v) = v$.

Now let n be an integer such that $1 \leqslant n \leqslant k-1$. Suppose for all $v \in W_1$ where $t(v) = n$, $p(v) = v$. Now let $v \in W_1$ where $t(v) = n+1$. We consider two cases, when $n+1$ is even and when $n+1$ is odd.

Suppose $n+1$ is even. Then $v = 1^{m_0} 0^{m_1} 1^{m_2} \ldots 0^{m_n} 1^{m_{n+1}}$ where $\sum_{i=0}^{n+1} m_i = k$. Let $v' = 1^{m_{n+1}} 1^{m_0} 0^{m_1} 1^{m_2} \ldots 0^{m_n}$. Then $t(v') = n$. By the induction hypothesis, $p(v') = v'$.

We know that $p(1^k) = (1^k)$. By Lemma 3.7, there exists a unique minimal walk in G_k from v' to 1^k. By Lemma 3.8, $p(a) = a$ for all vertices a on this walk. Since v is on this walk, $p(v) = v$.

Suppose $n+1$ is odd. Then $v = 1^{m_0}0^{m_1}1^{m_2}0^{m_3}\dots 1^{m_n}0^{m_{n+1}}$ where $\sum_{i=0}^{n+1} m_i = k$. Let $v' = 1^{m_{n+1}}1^{m_0}0^{m_1}1^{m_2}0^{m_3}\dots 1^{m_n}$. Then $t(v') = n$. By the induction hypothesis, $p(v') = v'$. We know that $p(0^k) = (0^k)$. By Lemma 3.7, there exists a unique minimal walk in G_k from v' to 0^k. By Lemma 3.8, $p(a) = a$ for all vertices a on this walk. Since v is on this walk, $p(v) = v$.

By induction, $p(v) = v$ for all vertices $v \in W_1$.

Let $W_0 = \{v = v_1v_2\dots v_k \in V(G_k) | v_1 = 0\}$. An analogous argument shows that $p(v) = v$ for all $v \in W_0$.

□

Lemma 3.11. *If p is an automorphism of G_k such that $p(0^k) = 1^k$ and $p(1^k) = 0^k$, then p is the dual permutation.*

Proof. The proof of this lemma is similar to the proof of Lemma 3.10, except that we will use Lemma 3.9 in place of Lemma 3.8.

Let p be an automorphism of G_k such that $p(0^k) = 1^k$ and $p(1^k) = 0^k$. By Lemma 3.7, there exists a unique minimal walk from 0^k to 1^k and from 1^k to 0^k. By Lemma 3.9, $p(a) = \hat{a}$ for all vertices a on these two walks.

Let W_1 and $t(v)$ be as defined in the proof of Lemma 3.10. We will prove by induction on t, that $p(v) = \hat{v}$ for all vertices $v \in W_1$.

Let $v \in W_1$ such that $t(v) = 1$. Then $v = 1^{m_0}0^{m_1}$ where $m_0 \geqslant 1$, $m_1 \geqslant 1$, and $m_0 + m_1 = k$. Then v is on the unique minimal walk from 1^k to 0^k. By Lemma 3.9, $p(v) = \hat{v}$.

Now let n be an integer such that $1 \leqslant n \leqslant k-1$. Suppose for all $v \in W_1$ where $t(v) = n$, $p(v) = \hat{v}$. Now let $v \in W_1$ where $t(v) = n+1$. We consider two cases, when $n+1$ is even and when $n+1$ is odd.

Suppose $n+1$ is even. Then $v = 1^{m_0}0^{m_1}1^{m_2}\dots 0^{m_n}1^{m_{n+1}}$ where $\sum_{i=0}^{n+1} m_i = k$. Let $v' = 1^{m_{n+1}}1^{m_0}0^{m_1}1^{m_2}\dots 0^{m_n}$. Then $t(v') = n$. By the induction hypothesis, $p(v') = \hat{v'}$. We know that $p(1^k) = (0^k)$. By Lemma 3.7, there exists a unique minimal walk in G_k from v' to 1^k. By Lemma 3.9, $p(a) = \hat{a}$ for all vertices a on this walk. Since v is on this walk, $p(v) = \hat{v}$.

Suppose $n+1$ is odd. Then $v = 1^{m_0}0^{m_1}1^{m_2}0^{m_3}\dots 1^{m_n}0^{m_{n+1}}$ where $\sum_{i=0}^{n+1} m_i = k$. Let $v' = 1^{m_{n+1}}1^{m_0}0^{m_1}1^{m_2}0^{m_3}\dots 1^{m_n}$. Then $t(v') = n$. By the induction hypothesis, $p(v') = \hat{v'}$. We know that $p(0^k) = (1^k)$. By Lemma 3.7, there exists a unique minimal walk in G_k from v' to 0^k. By Lemma 3.9, $p(a) = \hat{a}$ for all vertices a on this walk. Since v is on this walk, $p(v) = \hat{v}$.

By induction, $p(v) = \hat{v}$ for all vertices $v \in W_1$.

Let $W_0 = \{v = v_1v_2\dots v_k \in V(G_k) | v_1 = 0\}$. An analogous argument shows that $p(v) = \hat{v}$ for all $v \in W_0$.

□

Theorem 3.12. *A permutation p of $V(G_k)$ is an automorphism of G_k if and only if p is the identity or p is the dual permutation.*

Proof. The identity is always an automorphism. Let p be the dual permutation. Let Q' be the permutation matrix of p. Let A be the adjacency matrix of G_k. By Lemma 3.3, p is an automorphism of G_k if and only if $Q'A = AQ'$. By Lemma 3.5, $Q' = C$ where C is the counteridentity matrix. By Lemma 3.6, C commutes with A. Therefore, by Lemma 3.3, p is an automorphism of G_k.

Since 0^k and 1^k are the only vertices in G_k with self-loops, any automorphism p of G_k must fix 0^k and 1^k or swap them. If an automorphism p fixes 0^k and 1^k, then by Lemma 3.10, p is the identity. If an automorphism p is such that $p(0^k) = 1^k$ and $p(1^k) = 0^k$, then, by Lemma 3.11, p is the dual permutation.

Therefore, p is an automorphism of G_k if and only if p is the identity or p is the dual permutation. □

We say a labeling f of G_k pairs duals if $f(a) = f(\hat{a})$ for all $a \in V(G_k)$.

Corollary 3.13. *A labeling f of G_k is r-distinguishing if and only if f does not pair duals.*

Proof. A labeling of f of G_k is r-distinguishing if and only if the only automorphism of G_k that preserves the labeling is the identity. By Theorem 3.12, the only automorphisms of G_k are the identity and the dual permutation. So, a labeling f is r-distinguishing if and only if the dual permutation does not preserve the labeling. The dual permutation preserves the labeling if and only if $f(a) = f(\hat{a})$ for all $a \in V(G_k)$. Therefore, a labeling f of G_k is r-distinguishing if and only if f does not pair duals. □

4 Finitarily Markovian Factors

The k-block presentation of a Bernoulli scheme is a Markov process, but of course, not all Markov processes are k-block presentations of a Bernoulli scheme. We will begin with a definition and a lemma that apply to the more general case. Let X be a Markov process defined by some irreducible directed graph G and measure μ. Let $f : V(G) \to \{1, 2, \ldots, r\}$ be a labeling of the vertices of G. The labeling f then describes a factor map on X. Let f denote this factor map as well. Let $f : X \to Y$. A word in X is a finite length walk on the vertices of G. We will use the shorthand $x[1, k]$ to mean $x_1 x_2 \ldots x_k$.

Definition 4.1. *If there exists a word $w = w_1, w_2, \ldots, w_k$ in Y and a symbol a in the alphabet of X such that for any $x \in X$ where $f(x)[1, k] = w$, there exists an i in $[1, k]$ where $x_i = a$, then we say f is r-demarcating.*

The importance of this definition is demonstrated by the following lemma.

Lemma 4.2. *If f is r-demarcating, then Y is finitarily Markovian.*

Proof. Let Y have σ-algebra $\mathcal{V}$ and X have σ-algebra $\mathcal{U}$. If $Y[1,k] = w$, then $X_i = a$. Since X is Markov, $\mathcal{U}(\dots X_{-1}, X_0)$ and $\mathcal{U}(X_{k+1}, X_{k+2}, \dots)$ are independent given $X_i = a$. The σ-algebras $f^{-1}\mathcal{V}(\dots Y_{-1}, Y_0)$ and $f^{-1}\mathcal{V}(Y_{k+1}, Y_{k+2}, \dots)$ are sub-algebras of, respectively, $\mathcal{U}(\dots X_{-1}, X_0)$ and $\mathcal{U}(X_{k+1}, X_{k+2}, \dots)$. Then, $\mathcal{V}(\dots Y_{-1}, Y_0)$ and $\mathcal{V}(Y_{k+1}, Y_{k+2}, \dots)$ are independent given $Y[1,k] = w$. So, w is a renewal state in $Y^{(k)}$. This implies $Y^{(k)}$ is a renewal process. Therefore, Y is finitarily Markovian. □

With Lemma 4.2, one may find examples of labelings that define finitarily Markovian factors. Let X be a Bernoulli scheme. Let G be the digraph associated with $X^{(k)}$. Let f be a labeling of the vertices of G such that there exists a $v \in V(G)$ where $f(v) \neq f(u)$ for all $u \in V(G)$ where $u \neq v$. In other words, let f be a labeling so that one symbol in the image has a unique pre-image. In this case $f(v)$ is the w and v is the a in Definition 4.1. The labeling f is clearly r-demarcating. If we let $f : X^{(k)} \to Y$, then by Lemma 4.2, Y is finitarily Markovian. In Section 6, we will show that most labelings of $X^{(k)}$ have a symbol in the image that has a unique pre-image. Thus, most finite factors of a Bernoulli scheme are finitarily Markovian.

5 r-demarcating Implies r-distinguishing

Lemma 4.2 implies that if a labeling of a k-block presentation of a Bernoulli scheme defines a factor that is not finitarily Markovian, that labeling must not be r-demarcating.

Lemma 5.1. *Let $X = \{0,1\}^{\mathbb{Z}}$ be a Bernoulli scheme, let $k \in \mathbb{Z}^+$, and let G_k be the digraph associated to $X^{(k)}$. If a labeling f of G_k pairs duals then f is not r-demarcating.*

Proof. For every $x \in X$, there exists a $\hat{x} \in X$ such that $x_i \neq \hat{x}_i$ for all $i \in \mathbb{Z}$. Let f be a labeling of G_k that pairs duals. If $f : X \to Y$ also denotes the finite factor map defined by f, then $f(x) = f(\hat{x})$. If there existed a word $w = w_1, w_2, \dots, w_k$ in Y and a symbol a in the alphabet of X such that for any $x \in X$ where $f(x)[1,k] = w$, there exists an i in $[1,k]$ where $x_i = a$, then $f(\hat{x}) \neq f(x)$. Thus, by this contradiction, f must not be r-demarcating. □

Theorem 5.2. *If a labeling f of G_k is r-demarcating, then f is r-distinguishing.*

Proof. Suppose f is r-demarcating. Then by Lemma 5.1, f does not pair duals. By Corollary 3.13, f is r-distinguishing. □

We do not know if the converse holds in this case. The converse does not always hold in the more general setting where $X^{(k)}$ is replaced by any Markov process. Consider the directed graph in Figure 4. Since 5 is the only vertex in our graph with three incoming edges, any automorphism of this graph must fix 5. Then 0, 2, and 4 must be fixed. Finally, this implies 1 and 3 must be fixed. Thus, the only automorphism of this graph is the identity. In this instance, any labeling of the vertices would be r-distinguishing. In particular, we could label the graph as was done in Figure 5. It is clear that this labeling is not r-demarcating. We do not yet know for what class of digraphs, a labeling is r-demarcating if and only if it is r-distinguishing.

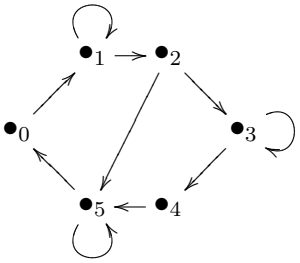

Figure 4

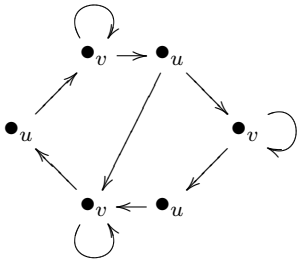

Figure 5

6 Counting Cases with the Bell Numbers

If X is an m-state Bernoulli scheme, then the number of k-block factors of X is always $B_{m^k} - 1$ where B_n is the nth Bell number. For a thorough introduction to the Bell numbers, see [6]. A Bell number B_n can be described as the number of ways to partition n numbers. We find that $B_0 = 0$, $B_1 = 1$, $B_2 = 2$, and $B_3 = 5$. There is the following recurrence relation for the Bell numbers.

$$B_n = \sum_k \binom{n-1}{k} B_k$$

In addition, B_n is described as the number of ways to put n labeled balls in n indistinguishable urns. Here, our balls are k blocks in X. Since X has m states, there are m^k k-blocks in X. The urns are the states in the factor. They are indistinguishable since we want to consider two factors that are 1-block codings of each other as the same factor. We take $B_{m^k} - 1$ because we are not interested in the factor that equates to putting all of our balls in the same urn.

In summary, we have the following simple lemma.

Lemma 6.1. *There are $B_{m^k} - 1$ distinct k-block factors of X.*

For example, the number of distinct 3-block factors of a 2-state Bernoulli scheme is $B_{2^3} - 1 = 4139$. The number of distinct 4-block factors of a 2-state Bernoulli scheme is $B_{2^4} - 1 = 10480142147$. The following theorem shows that most finite factors are finitarily Markovian.

Theorem 6.2. *Let X be a Bernoulli scheme with $m \geqslant 2$ states. For each $k \in \mathbb{Z}^+$, define $N(k)$ to be the number of distinct k-block factors of X, and define $R(k)$ be the number of distinct k-block factors of X that are finitarily Markovian. Then,*

$$\lim_{k\to\infty} \frac{R(k)}{N(k)} = 1.$$

Proof. For each $k \in \mathbb{Z}^+$, let $R'(k)$ be the number of k-block factors of X such that one state in the factor has a unique pre-image. Then, by Lemma 4.2, $R'(k) \leqslant R(k)$. So to show

$$\lim_{k\to\infty} \frac{R(k)}{N(k)} = 1.$$

it suffices to show

$$\lim_{k\to\infty} \frac{R'(k)}{N(k)} = 1.$$

We know that $N(k) = B_{m^k} - 1$. Let C_n be the number of ways to put n labeled balls in n indistinguishable bins such that every bin has > 1 or 0 balls. Then $B_n = C_n + C_{n+1}$. If $D_n = B_n - C_n = C_{n+1}$, then $D_n = B_n - D_{n-1}$. The number of ways to factor so that one symbol in the image has a unique pre-image is the total number of ways to factor minus the number of factors where every symbol in the image has at least 2 pre-images. So, $D_{m^k} = R'(k)$. Then,

$$\lim_{k\to\infty} \frac{R'(k)}{N(k)} = \lim_{k\to\infty} \frac{D_{m^k}}{B_{m^k} - 1} = \lim_{k\to\infty} \frac{B_{m^k} - D_{m^k-1}}{B_{m^k}}.$$

Since $D_n \leqslant B_n$ for all n,

$$\lim_{k\to\infty} \frac{B_{m^k} - D_{m^k-1}}{B_{m^k}} \geqslant 1 - \lim_{k\to\infty} \frac{B_{m^k-1}}{B_{m^k}}.$$

All that is left to show is that

$$\lim_{k\to\infty} \frac{B_{m^k-1}}{B_{m^k}} = 0.$$

We thank [13] for providing us with the following simple argument. Dobinski's formula tells us that B_n is the nth moment of a Posson distribution with expected value 1 [18]. Then for any real number r, there exists $n \in \mathbb{Z}^+$ such that $r^n \leqslant B_n$. By Jensen's inequality, $B_n^{\frac{n+1}{n}} \leqslant B_{n+1}$. Then we have

$$1 \leqslant \frac{B_{n+1}}{\sqrt[n]{B_n^{n+1}}} \text{ which implies } \sqrt[n]{B_n} \leqslant \frac{B_{n+1}}{B_n}.$$

So for any $r \in \mathbb{R}$, there exists an $n \in \mathbb{Z}^+$ such that $r \leqslant \sqrt[n]{B_n} \leqslant \frac{B_{n+1}}{B_n}$. Therefore,

$$\lim_{k\to\infty} \frac{B_{m^k-1}}{B_{m^k}} = 0.$$

□

Let $X = \{0,1\}^{\mathbb{Z}}$ with measure $\mu = (p_0, p_1)^{\mathbb{Z}}$ where $p_0 \neq p_1$. Let $k \in \mathbb{Z}$ and let f be a labeling of G_k (the digraph associated with $X^{(k)}$). We conjecture that $f(X)$ is finitarily Markovian if and only if f is r-distinguishing. By our work in Section 3, we know that f is r-distinguishing if and only if f does not pair duals. So, we conjecture that the number of finite factors that are not finitarily Markovian is the number of factors that pair duals. We can use the Bell numbers to count these factors as well. There are 2^{k-1} pairs of duals in X. So, there are $B_{2^{(k-1)}} - 1$ distinct k-block factors of X that pair duals.

7 Remarks and Future Directions

7.1 Summary of implications

Let $X = \{0,1\}^{\mathbb{Z}}$ with measure $\mu = (p_0, p_1)^{\mathbb{Z}}$ where $p_0 \neq p_1$. Let $k \in \mathbb{Z}^+$, and let f be a labeling of G_k. Let $f : X \to Y$ denote the finite factor map defined by f. In Section 3, we conjectured that Y is finitarily Markovian if and only if f is r-distinguishing.

Let $r - dem$ (resp. $r - dis$) mean f is r-demarcating (resp. r-distinguishing). Let $DNPD$ mean f does not pairs duals, and let FM mean Y is finitarily Markovian. We can summarize the main results of this paper with the following implications:

$$r - dis \iff DNPD \impliedby r - dem \implies FM.$$

7.2 Hidden Markov models

Hidden Markov models (or probabilistic functions of Markov chains) [3] are 1-block factors of Markov processes. Let X be a Markov process and let $f : X \to Y$ be a 1-block factor. The question of whether or not Y is Markovian has been well-studied (see [20], for example). However, not much has been written on whether or not a given hidden Markov model is finitarily Markovian. Some of our work here was done in this more general setting. Specifically, in Lemma 4.2, we have shown that r-demarcating is always sufficient for the factor to be finitarily Markovian.

References

[1] M. A. Akcoglu, A. Del Junco, and M. Rahe (1979). Finitary Codes between Markov Processes. Wahrsch. Verw. Gebiete **47** 305-314. MR0525312

[2] M. O. Albertson and K. L. Collins (1996) *Symmetry Breaking in Graphs,* Electron. J. Combin. **3** #R18. MR1394549

[3] L. Baum and T. Petrie (1966) *Statistical inference for probabilistic functions of finite state Markov chains*, Ann. Math. Statist. **37** 1554-1563. MR0202264

[4] J.R. Blum and D. Hanson (1963) *On isomorphism problems for Bernoulli schemes*, Bull.A.M.S. 69, 221-223 MR0143862

[5] G. Chartand and L. Lesniak (2005), *Graphs and Digraphs*, Fourth Edition. Chapman and Hall/CRC, Boca Raton, London, New York, Washington, D.C. MR2107429

[6] J. Harris, J. Hirst, and M. Mossinghoff (2008) *Combinatorics and Graph Theory, 2nd Ed.* Undergraduate Texts in Mathematics. Springer, New York. MR2440898

[7] M. Keane and M. Smorodinsky (1977). *A class of finitary codes*, Israel J. Math. 26 nos. 3-4, 352-371. MR0450514

[8] M. Keane and M. Smorodinsky (1979). *Bernoulli Schemes of the same entropy are finitarily isomorphic*, Ann. of Math. (2) 109, 2, 397-406. MR528969

[9] M. Keane and M. Smorodisky (1979). *Finitary isomorphisms of irreducible Markov shifts*, Isreal J. Math. 34, 4, 281-286 (1980). MR570887

[10] A. N. Kolmogorov (1958) *A new metric invariant of transitive dynamic systems and automorphism in Lebesgue spaces*, Dokl, Akad. Nauk. SSSR 119, 861-864 MR0103254

[11] C. Laflamme, L. Nguyen Van The, and N. Sauer (2010) *Distinguishing number of countable homogeneous relational structures*, Electon J. Combin. 17, #R20. MR2587751

[12] J. Lauri and R. Scapellato (2003) *Topics in Graph Automorphisms and Reconstruction.* Cambridge University Press. MR1971819

[13] Mathoverflow.net. On the Bell Numbers. Asked March 4, 2010. http://mathoverflow.net/questions/17091/on-the-bell-numbers.

[14] L. D. Meshalkin (1962) *A case of isomorphism of Bernoulli schemes*, Dokl, Akad, Nauk. SSSR 147, 797-800, and Soviet Math. Dokl. 3, 1725-1729. MR0110782

[15] G. Morvai and B. Weiss (2007) *On estimating the memory for finitarily Markovian processes*, Ann. I.H. Poincare-PR, vol. 43, pp 15-30. MR2288267

[16] D. Ornstein, Ergodic Theory, Randomness, and Dynamical Systems, Yale University Press, 1974. MR0447525

[17] D. Ornstein (1970) *Factors of Bernouli Shifts are Bernoulli Shifts*, Advances in Math. 5, 349-364. MR0274717

[18] J. Pitman (2006) *Combinatorial Stochastic Processes*, in: Lecture Notes Math., vol. 1875, Springer. MR2245368

[19] J. Rissanen(1983) *A universal data compression system*. IEEE Trans. Inform. **29** 656-664. MR0730903

[20] G. Rubino and B. Sericola (1991) *A finite characterization of weak lumpable Markov processes. Part I: The discrete time case*, Stoch. Proc. Appl. 38 195-204. MR1119981

[21] D. Rudolph, (1981). *A characterization of those processes finitarily isomorphic to a Bernoulli shift*. In Ergodic Theory and Dynamical systems, I (College Park, Md. 1979-80). Progr. Math., Vol 10. Birkhauer Boston, Mass. 1-64. MR633760

[22] D. Rudolph, (1982) *A mixing Markov chain with exponentially decaying return times is finitarily Bernoulli* Ergodic Theory, Dynam. Systems 2 85-97. MR0684246

[23] J. Serafin, *Finitary codes, a short survey*, Dynamics & stochastics, 262273, IMS Lecture Notes Monogr. Ser., 48, Inst. Math. Statist., Beachwood, OH, 2006. MR2306207

[24] S. Shea (2010) *A note on r-processes*, Braz. J. Probab. Stat. Vol 24, Num 3, 502-508. MR2719699

[25] S. Shea (2009) *Finitary isomorphisms of some renewal processes to Bernoulli schemes*, Indagationes Mathematcae, Vol 20, Iss 3, 463-476. MR2639984

[26] S. Shea, *On the marker method for constructing finitary isomorphisms*, accepted for publication in the Rocky Mountain Journal of Mathematics.

[27] Ya. G. Sinai, *A Weak Isomorphism of Transformations Having an Invariant Measure*, Dokl. Akad. Nauk, SSSR 147 (1962), 797-800. MR0161960

[28] M. Smorodinsky (1992) *Finitary isomorphism of m-dependent processes*, symbolic dynamics and its applications, Contemp. Math. 135, 373-376. MR1853803