Ornstein Isomophism and Algorithmic Randomness

Mrinalkanti Ghosh[1], Satyadev Nandakumar[2], and Atanu Pal[3]

[1]Toyota Technological Institute, Chicago
[2]Indian Institute of Technology Kanpur
[3]Strand Genomics, Bangalore

June 12, 2014

Introduction

Kolmogorov's Programme:

Kolmogorov's Programme:

"The application of probability theory can be put on a uniform basis. It is always a matter of hypotheses about the impossibility of reducing in one way or another the complexity of the description of objects in question."

Kolmogorov's Programme:

"The application of probability theory can be put on a uniform basis. It is always a matter of hypotheses about the impossibility of reducing in one way or another the complexity of the description of objects in question."

Consider theorems in Probability theory which hold "almost everywhere".

Kolmogorov's Programme:

"The application of probability theory can be put on a uniform basis. It is always a matter of hypotheses about the impossibility of reducing in one way or another the complexity of the description of objects in question."

Consider theorems in Probability theory which hold "almost everywhere". Can we show that if an object has maximum descriptional complexity, (i.e. is "random"), then it obeys the theorem?

Kolmogorov Theme

Computability Theory

↓ Compressibility

Information Theory

Kolmogorov Theme

Computability Theory

↓ Compressibility

Information Theory

↓ Entropy

Dynamical Systems

Dynamical Systems

Definition 1. Let $(X, \mathcal{F}, P)$ be a probability space.

Definition 1. Let $(X, \mathcal{F}, P)$ be a probability space.

A measurable transformation $T : X \to X$ is called *measure-preserving* if for every $A \in \mathcal{F}$, $P(T^{-1}A) = P(A)$.

Definition 1. Let $(X, \mathcal{F}, P)$ be a probability space.

A measurable transformation $T : X \to X$ is called *measure-preserving* if for every $A \in \mathcal{F}$, $P(T^{-1}A) = P(A)$.

A measure-preserving map T is *ergodic* if for all $A \in \mathcal{F}$, $TA = A$ only when $P(A) \in \{0, 1\}$.

Definition 1. Let $(X, \mathcal{F}, P)$ be a probability space.

A measurable transformation $T : X \to X$ is called *measure-preserving* if for every $A \in \mathcal{F}$, $P(T^{-1}A) = P(A)$.

A measure-preserving map T is *ergodic* if for all $A \in \mathcal{F}$, $TA = A$ only when $P(A) \in \{0, 1\}$.

Example. If X is a finite set with the uniform distribution on it, then every permutation is a measure-preserving transformation.

Definition 1. Let $(X, \mathcal{F}, P)$ be a probability space.

A measurable transformation $T : X \to X$ is called *measure-preserving* if for every $A \in \mathcal{F}$, $P(T^{-1}A) = P(A)$.

A measure-preserving map T is *ergodic* if for all $A \in \mathcal{F}$, $TA = A$ only when $P(A) \in \{0, 1\}$.

Example. If X is a finite set with the uniform distribution on it, then every permutation is a measure-preserving transformation.

Any permutation consisting of a single cycle is an ergodic transformation.

Definition 1. Let $(X, \mathcal{F}, P)$ be a probability space.

A measurable transformation $T : X \to X$ is called *measure-preserving* if for every $A \in \mathcal{F}$, $P(T^{-1}A) = P(A)$.

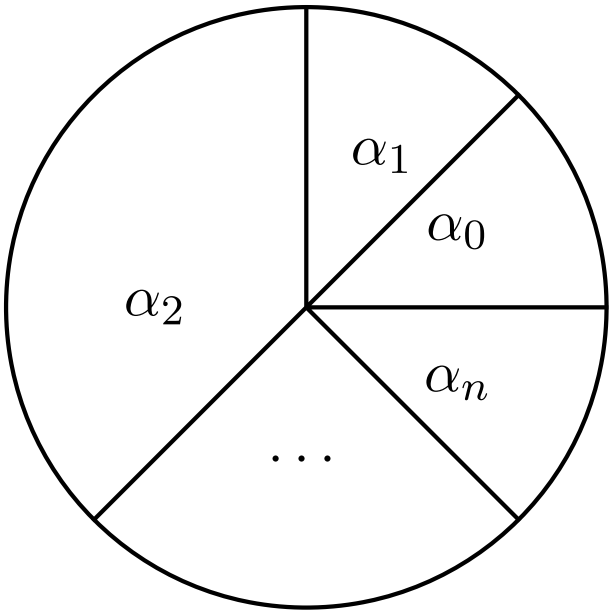
A measure-preserving map T is *ergodic* if for all $A \in \mathcal{F}$, $TA = A$ only when $P(A) \in \{0, 1\}$.

Example. If X is a finite set with the uniform distribution on it, then every permutation is a measure-preserving transformation.

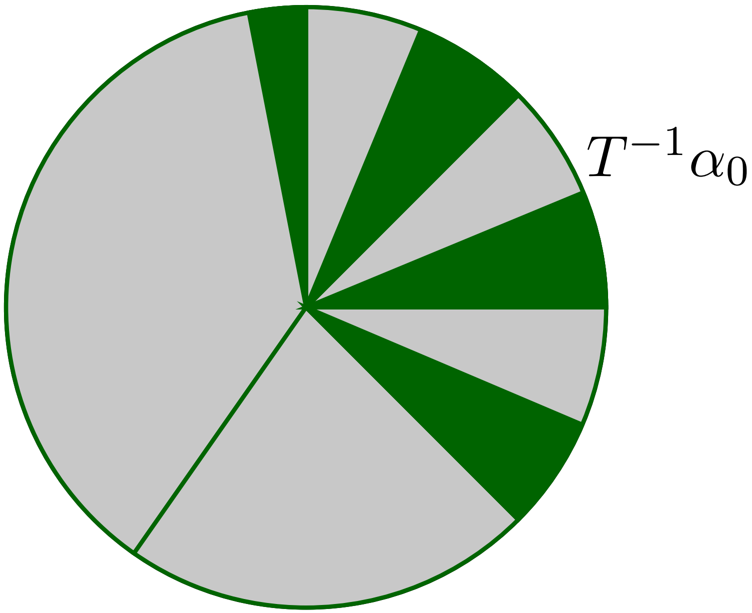
Any permutation consisting of a single cycle is an ergodic transformation.

Definition 2. A system $(X, \mathcal{F}, P, T)$ where $(X, \mathcal{F}, P)$ is a probability space and T is measure-preserving with respect to it, is called a *dynamical system*.

Partitions



Partitions



Kolmogorov-Sinai Entropy

The entropy of a partition $\alpha = (\alpha_0, \ldots, \alpha_{n-1})$ of X is

$$H(\alpha) = \sum_{i=0}^{n-1} P(\alpha_i) \log \left(\frac{1}{P(\alpha_i)} \right).$$

The entropy of a partition $\alpha = (\alpha_0, \ldots, \alpha_{n-1})$ of X is

$$H(\alpha) = \sum_{i=0}^{n-1} P(\alpha_i) \log \left(\frac{1}{P(\alpha_i)} \right).$$

The k-step entropy is

$$h_k(\alpha, T) = \frac{H\left(\alpha \vee \cdots \vee T^{-k+1}\alpha\right)}{k}.$$

Kolmogorov-Sinai Entropy

The entropy of a partition $\alpha = (\alpha_0, \ldots, \alpha_{n-1})$ of X is

$$H(\alpha) = \sum_{i=0}^{n-1} P(\alpha_i) \log \left(\frac{1}{P(\alpha_i)} \right).$$

The k-step entropy is

$$h_k(\alpha, T) = \frac{H\left(\alpha \vee \cdots \vee T^{-k+1}\alpha\right)}{k}.$$

The entropy of a transformation T wrt α is

$$h(\alpha, T) = \lim_{k \to \infty} h_k(\alpha, T).$$

The entropy of a partition $\alpha = (\alpha_0, \ldots, \alpha_{n-1})$ of X is

$$H(\alpha) = \sum_{i=0}^{n-1} P(\alpha_i) \log \left(\frac{1}{P(\alpha_i)} \right).$$

The k-step entropy is

$$h_k(\alpha, T) = \frac{H\left(\alpha \vee \cdots \vee T^{-k+1}\alpha\right)}{k}.$$

The entropy of a transformation T wrt α is

$$h(\alpha, T) = \lim_{k \to \infty} h_k(\alpha, T).$$

The entropy of a transformation T is

$$h(T) = \sup\{h(\alpha, T) \mid \alpha \text{ is a finite partition of } X\}.$$

Kolmogorov-Sinai Theorem

The partition α of X is called a *generator* if the σ-algebra on X is generated by $\cdots \vee T^{-2}\alpha \vee T^{-1}\alpha \vee \alpha \vee T\alpha \vee T^2\alpha \ldots$.

Kolmogorov-Sinai Theorem

The partition α of X is called a *generator* if the σ-algebra on X is generated by $\cdots \vee T^{-2}\alpha \vee T^{-1}\alpha \vee \alpha \vee T\alpha \vee T^2\alpha \dots$.

Theorem 3. *If α is a generator, then $h(\alpha, T) = h(T)$.*

The partition α of X is called a *generator* if the σ-algebra on X is generated by $\cdots \vee T^{-2}\alpha \vee T^{-1}\alpha \vee \alpha \vee T\alpha \vee T^2\alpha \ldots$.

Theorem 3. *If α is a generator, then $h(\alpha, T) = h(T)$.*

(α is a "natural" partition induced by T.)

The partition α of X is called a *generator* if the σ-algebra on X is generated by $\cdots \vee T^{-2}\alpha \vee T^{-1}\alpha \vee \alpha \vee T\alpha \vee T^2\alpha \ldots$.

Theorem 3. *If α is a generator, then $h(\alpha, T) = h(T)$.*

(α is a "natural" partition induced by T.)

Definition 4. An *isomorphism* $\phi : A \to B$ is a function such that $\phi \circ T_A = T_B \circ \phi$.

The partition α of X is called a *generator* if the σ-algebra on X is generated by $\cdots \vee T^{-2}\alpha \vee T^{-1}\alpha \vee \alpha \vee T\alpha \vee T^2\alpha \ldots$.

Theorem 3. *If α is a generator, then $h(\alpha, T) = h(T)$.*

(α is a "natural" partition induced by T.)

Definition 4. An *isomorphism* $\phi : A \to B$ is a function such that $\phi \circ T_A = T_B \circ \phi$.

Theorem 5. *If two dynamical systems are isomorphic to each other, then they have the same Kolmogorov-Sinai entropy.*

Converse of the KS theorem

Let Σ_A and Σ_B be two finite alphabets.

Converse of the KS theorem

Let Σ_A and Σ_B be two finite alphabets.

Let $A = (\Sigma_A^\infty, \mathcal{B}(\Sigma_A^\infty), P_A, T_A)$ and $B = (\Sigma_B^\infty, \mathcal{B}(\Sigma_B^\infty), P_B, T_B)$ be two Bernoulli systems with the same KS entropy.

Converse of the KS theorem

Let Σ_A and Σ_B be two finite alphabets.

Let $A = (\Sigma_A^\infty, \mathcal{B}(\Sigma_A^\infty), P_A, T_A)$ and $B = (\Sigma_B^\infty, \mathcal{B}(\Sigma_B^\infty), P_B, T_B)$ be two Bernoulli systems with the same KS entropy.

Are the two systems necessarily isomorphic?

Converse of the KS theorem

Let Σ_A and Σ_B be two finite alphabets.

Let $A = (\Sigma_A^\infty, \mathcal{B}(\Sigma_A^\infty), P_A, T_A)$ and $B = (\Sigma_B^\infty, \mathcal{B}(\Sigma_B^\infty), P_B, T_B)$ be two Bernoulli systems with the same KS entropy.

Are the two systems necessarily isomorphic?

(Note: Σ_A and Σ_B need not have the same cardinality.)

Converse of the KS theorem

Let Σ_A and Σ_B be two finite alphabets.

Let $A = (\Sigma_A^\infty, \mathcal{B}(\Sigma_A^\infty), P_A, T_A)$ and $B = (\Sigma_B^\infty, \mathcal{B}(\Sigma_B^\infty), P_B, T_B)$ be two Bernoulli systems with the same KS entropy.

Are the two systems necessarily isomorphic?

(Note: Σ_A and Σ_B need not have the same cardinality.)

Answer: Yes [Orn70]. In fact, there is a finitary isomorphism between them [KS79].

The finite portions $x[\,-m\ldots 0\ldots m\,]$ of an infinite sequence x are the *cylinders* of x.

Setting

The finite portions $x[\,-m\ldots 0\ldots m\,]$ of an infinite sequence x are the *cylinders* of x.

A *finitary* map $\phi : A \to B$ is one where for every $x \in A$ such that $\phi(x)$ is defined, there is an N such that

Setting

The finite portions $x[\,-m \dots 0 \dots m\,]$ of an infinite sequence x are the *cylinders* of x.

A *finitary* map $\phi : A \to B$ is one where for every $x \in A$ such that $\phi(x)$ is defined, there is an N such that $\phi(x[\,-N \dots 0 \dots N\,])$ determines $\phi(x)[0]$.

Setting

The finite portions $x[\,-m\ldots 0\ldots m\,]$ of an infinite sequence x are the *cylinders* of x.

A *finitary* map $\phi : A \to B$ is one where for every $x \in A$ such that $\phi(x)$ is defined, there is an N such that $\phi(x[\,-N\ldots 0\ldots N\,])$ determines $\phi(x)[0]$.

This N, in general, depends on the x.

Setting

The finite portions $x[\,-m\ldots0\ldots m\,]$ of an infinite sequence x are the *cylinders* of x.

A *finitary* map $\phi : A \to B$ is one where for every $x \in A$ such that $\phi(x)$ is defined, there is an N such that $\phi(x[\,-N\ldots0\ldots N\,])$ determines $\phi(x)[0]$.

This N, in general, depends on the x.

Further, $\phi(x)$ may not be defined on some x.

The finite portions $x[\ -m \ldots 0 \ldots m\ ]$ of an infinite sequence x are the *cylinders* of x.

A *finitary* map $\phi : A \to B$ is one where for every $x \in A$ such that $\phi(x)$ is defined, there is an N such that $\phi(x[\ -N \ldots 0 \ldots N\ ])$ determines $\phi(x)[0]$.

This N, in general, depends on the x.

Further, $\phi(x)$ may not be defined on some x.

Overview of the Proof

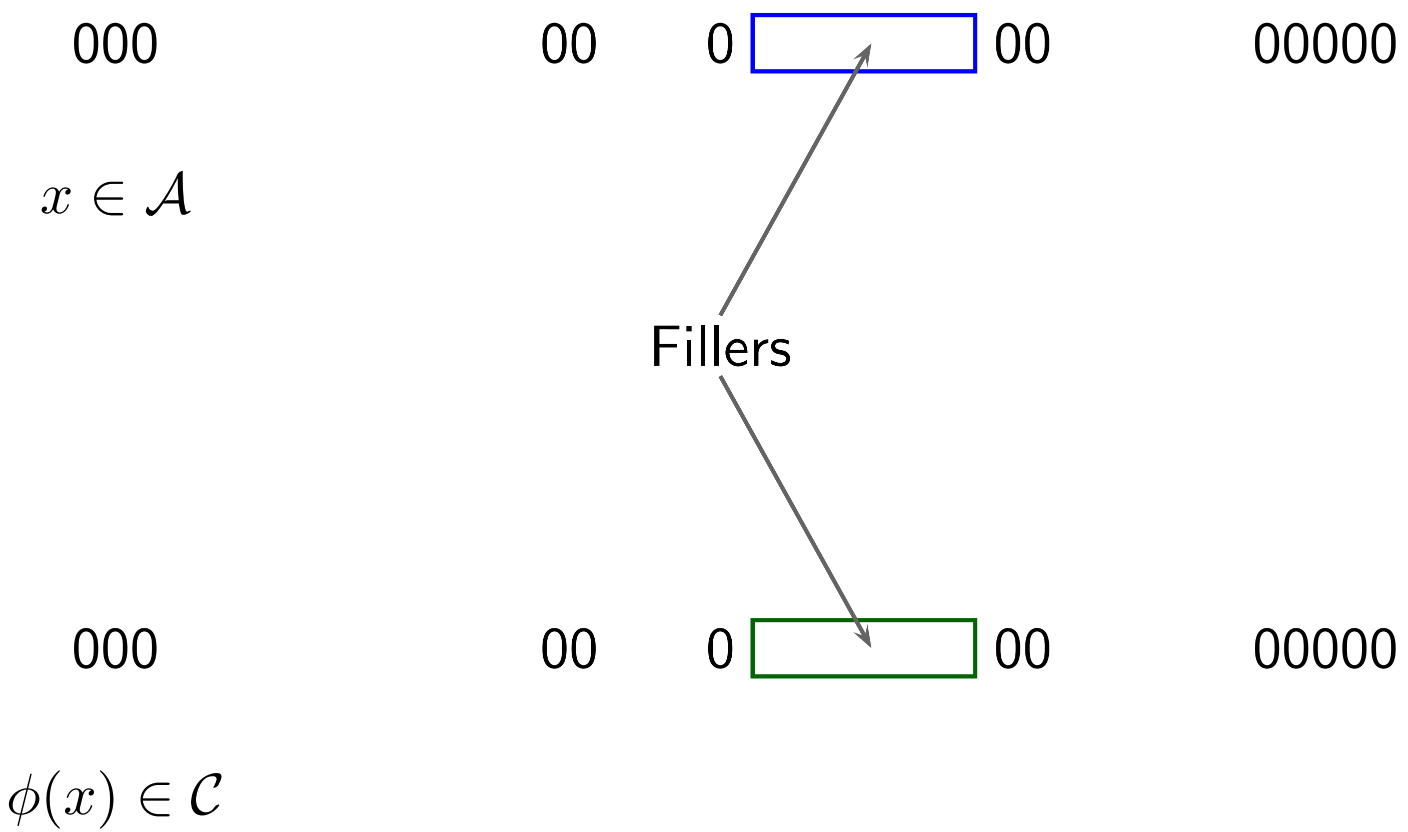
000 00 0 00 00000

$x \in \mathcal{A}$

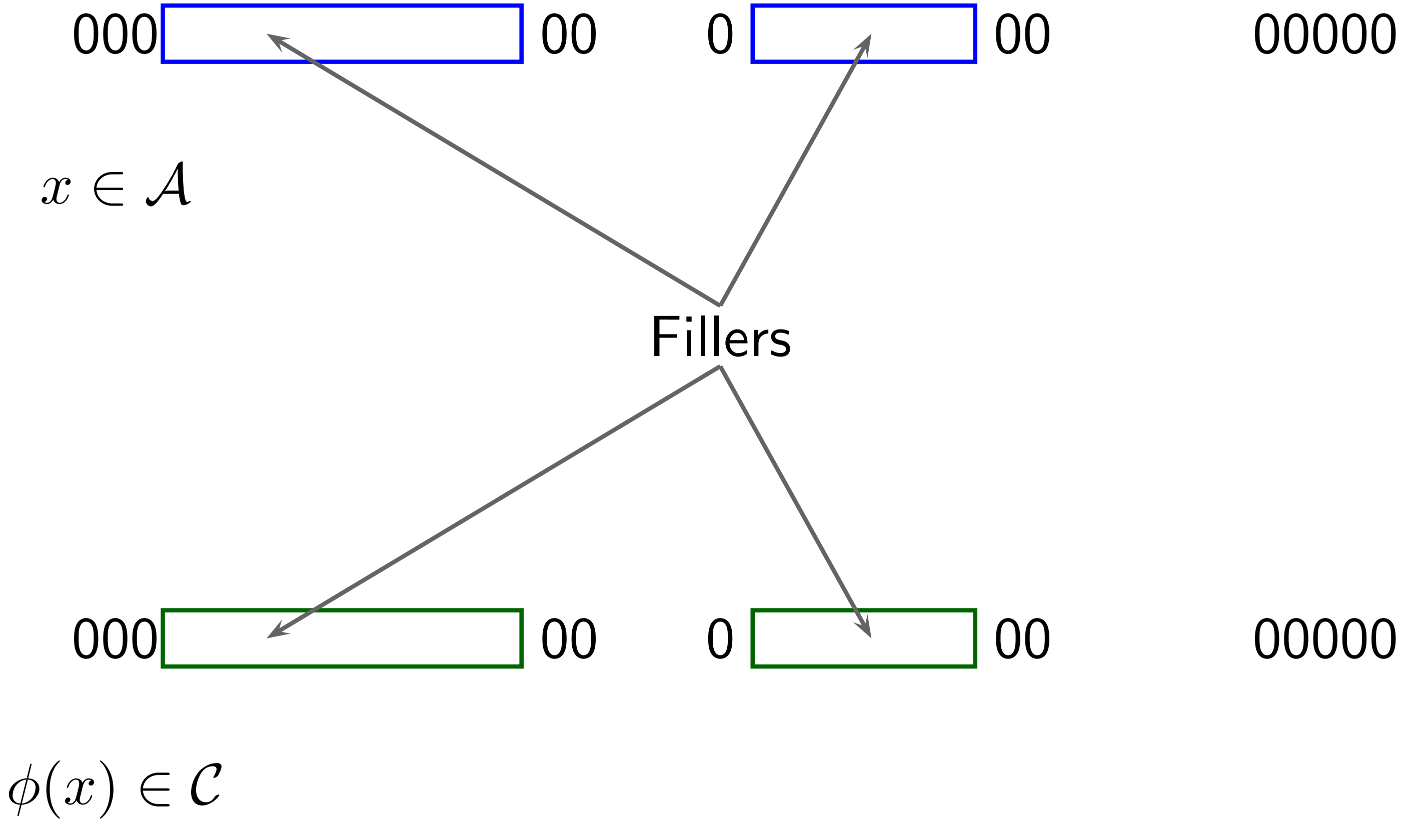
000 00 0 00 00000

$\phi(x) \in \mathcal{C}$

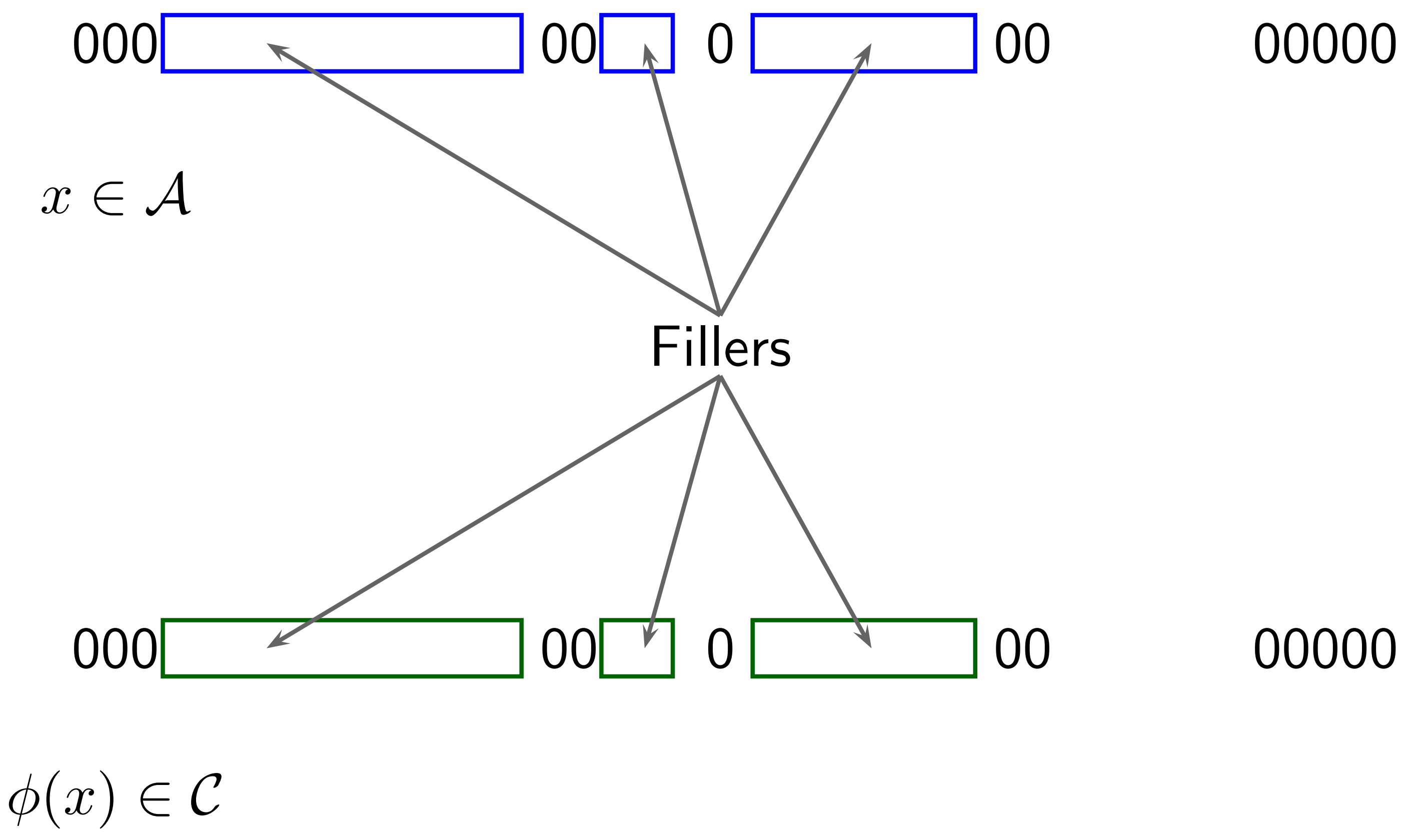
Overview of the Proof



Overview of the Proof



Overview of the Proof



Definition 6. A dynamical system $A = (\Sigma_A^\infty, \mathcal{B}(\Sigma_A^\infty), P_A, T_A)$ is called *computable* if $P_A : \Sigma_A^* \to [0, 1]$ is computable, and $T_A : \Sigma_A^* \to \Sigma_A^*$ is a computable monotone transformation.

Computability of ϕ

Definition 6. A dynamical system $A = (\Sigma_A^\infty, \mathcal{B}(\Sigma_A^\infty), P_A, T_A)$ is called *computable* if $P_A : \Sigma_A^* \to [0,1]$ is computable, and $T_A : \Sigma_A^* \to \Sigma_A^*$ is a computable monotone transformation.

Let us assume that A and B are computable systems.

Definition 6. A dynamical system $A = (\Sigma_A^\infty, \mathcal{B}(\Sigma_A^\infty), P_A, T_A)$ is called *computable* if $P_A : \Sigma_A^* \to [0, 1]$ is computable, and $T_A : \Sigma_A^* \to \Sigma_A^*$ is a computable monotone transformation.

Let us assume that A and B are computable systems.

Does this make ϕ computable?

Definition 6. A dynamical system $A = (\Sigma_A^\infty, \mathcal{B}(\Sigma_A^\infty), P_A, T_A)$ is called *computable* if $P_A : \Sigma_A^* \to [0,1]$ is computable, and $T_A : \Sigma_A^* \to \Sigma_A^*$ is a computable monotone transformation.

Let us assume that A and B are computable systems.

Does this make ϕ computable?

No! ϕ is undefined at several points - it is defined on some measure 1 proper subset, but may be undefined on a measure 0, nonempty set.

Computability of ϕ

Definition 6. A dynamical system $A = (\Sigma_A^\infty, \mathcal{B}(\Sigma_A^\infty), P_A, T_A)$ is called *computable* if $P_A : \Sigma_A^* \to [0,1]$ is computable, and $T_A : \Sigma_A^* \to \Sigma_A^*$ is a computable monotone transformation.

Let us assume that A and B are computable systems.

Does this make ϕ computable?

No! ϕ is undefined at several points - it is defined on some measure 1 proper subset, but may be undefined on a measure 0, nonempty set.

Where exactly is the isomorphism well-defined?

Definition 6. A dynamical system $A = (\Sigma_A^\infty, \mathcal{B}(\Sigma_A^\infty), P_A, T_A)$ is called *computable* if $P_A : \Sigma_A^* \to [0,1]$ is computable, and $T_A : \Sigma_A^* \to \Sigma_A^*$ is a computable monotone transformation.

Let us assume that A and B are computable systems.

Does this make ϕ computable?

No! ϕ is undefined at several points - it is defined on some measure 1 proper subset, but may be undefined on a measure 0, nonempty set.

Where exactly is the isomorphism well-defined?

Definition 6. A dynamical system $A = (\Sigma_A^\infty, \mathcal{B}(\Sigma_A^\infty), P_A, T_A)$ is called *computable* if $P_A : \Sigma_A^* \to [0, 1]$ is computable, and $T_A : \Sigma_A^* \to \Sigma_A^*$ is a computable monotone transformation.

Let us assume that A and B are computable systems.

Does this make ϕ computable?

No! ϕ is undefined at several points - it is defined on some measure 1 proper subset, but may be undefined on a measure 0, nonempty set.

Where exactly is the isomorphism well-defined?

Answer: (At least) over the Martin-Löf random points in the systems.

Layerwise Computability

Most strings in Σ^n are random.

Layerwise Computability

Most strings in Σ^n are random. Similarly, a measure 1 subset of any computable space consist of "random" objects.

Layerwise Computability

Most strings in Σ^n are random. Similarly, a measure 1 subset of any computable space consist of "random" objects.

Let $U_1, U_2, \ldots$ be some computable enumeration of open intervals with rational endpoints, in the space.

Most strings in Σ^n are random. Similarly, a measure 1 subset of any computable space consist of “random” objects.

Let $U_1, U_2, \ldots$ be some computable enumeration of open intervals with rational endpoints, in the space. A constructive measure 0 set is one which can be expressed as

$$\bigcap_{m>0} \bigcup_{n=1}^{\infty} U_{i_n,m},$$

where for each m, we have that the open cover $\bigcup_{n=1}^{\infty} U_{i_n,m}$ has probability less than $\frac{1}{2^m}$.

Layerwise computability

Since there is a universal Turing machine, there is a largest constructive measure 0 set.

Layerwise computability

Since there is a universal Turing machine, there is a largest constructive measure 0 set.

The complement of this set is the smallest co-constructive measure 1 set, which is called the set of Martin-Löf random objects.

Layerwise computability

Since there is a universal Turing machine, there is a largest constructive measure 0 set.

The complement of this set is the smallest co-constructive measure 1 set, which is called the set of Martin-Löf random objects.

Let us denote $K_m = U_m^c$. The sequence $\langle K_m \rangle_{m=1}^{\infty}$ is called a *layering* of X.

Layerwise computability

Since there is a universal Turing machine, there is a largest constructive measure 0 set.

The complement of this set is the smallest co-constructive measure 1 set, which is called the set of Martin-Löf random objects.

Let us denote $K_m = U_m^c$. The sequence $\langle K_m \rangle_{m=1}^{\infty}$ is called a *layering* of X. If a function $\phi(K_m)$ is computable uniformly in m, we say that it is *layerwise computable*. (Such computations converge on all random points.)

Layerwise computability

Since there is a universal Turing machine, there is a largest constructive measure 0 set.

The complement of this set is the smallest co-constructive measure 1 set, which is called the set of Martin-Löf random objects.

Let us denote $K_m = U_m^c$. The sequence $\langle K_m \rangle_{m=1}^{\infty}$ is called a *layering* of X. If a function $\phi(K_m)$ is computable uniformly in m, we say that it is *layerwise computable*. (Such computations converge on all random points.)

Structure of the Proof

We will construct a layerwise computable isomorphism which will take Martin-Löf random points in A to those in B and conversely.

1. The Marker Lemma
2. The Skeleton Lemma
3. The Filler Lemma
4. The Marriage Lemma
5. The Assignment Lemma

Marker Lemma

Reduce the problem to the following: construct an isomorphism between two mixing Markov systems with the same entropy **and**

Marker Lemma

Reduce the problem to the following: construct an isomorphism between two mixing Markov systems with the same entropy **and** having some symbol with equal probability.

Reduce the problem to the following: construct an isomorphism between two mixing Markov systems with the same entropy **and** having some symbol with equal probability.

Sort Σ_A and Σ_B in decreasing order of probability.[1] Designate the symbol with the highest probability in Σ_A as 0, and that with the least probability in Σ_B as 1.

Marker Lemma

Reduce the problem to the following: construct an isomorphism between two mixing Markov systems with the same entropy **and** having some symbol with equal probability.

Sort Σ_A and Σ_B in decreasing order of probability.[1] Designate the symbol with the highest probability in Σ_A as 0, and that with the least probability in Σ_B as 1.

Construct a mixing Markov system C with approximately the same entropy as A and B, with $P_C(0) = P_A(0)$ and $P_C(1) = P_B(1)$.

Marker Lemma

Reduce the problem to the following: construct an isomorphism between two mixing Markov systems with the same entropy **and** having some symbol with equal probability.

Sort Σ_A and Σ_B in decreasing order of probability.[1] Designate the symbol with the highest probability in Σ_A as 0, and that with the least probability in Σ_B as 1.

Construct a mixing Markov system C with approximately the same entropy as A and B, with $P_C(0) = P_A(0)$ and $P_C(1) = P_B(1)$.
Fix an alphabet size c large enough that the entropy of the partition Σ_C is greater than that of Σ_A and Σ_B.

Reduce the problem to the following: construct an isomorphism between two mixing Markov systems with the same entropy **and** having some symbol with equal probability.

Sort Σ_A and Σ_B in decreasing order of probability.[1] Designate the symbol with the highest probability in Σ_A as 0, and that with the least probability in Σ_B as 1.

Construct a mixing Markov system C with approximately the same entropy as A and B, with $P_C(0) = P_A(0)$ and $P_C(1) = P_B(1)$.
Fix an alphabet size c large enough that the entropy of the partition Σ_C is greater than that of Σ_A and Σ_B.

Now, we need an algorithm to define the probabilities of strings $x \in \Sigma_C^*$.

[1] We work with $(1 \pm \epsilon_n)$ approximations of probability.

Marker Lemma - Algorithm

Let m be the memory of the Markov systems.

1. Input: a string $x \in \Sigma^*$ and $n \in \mathbb{N}$ where we require $|H_A - H_C|, |H_B - H_C| < \frac{1}{2^n}$.

Marker Lemma - Algorithm

Let m be the memory of the Markov systems.

1. Input: a string $x \in \Sigma^*$ and $n \in \mathbb{N}$ where we require $|H_A - H_C|, |H_B - H_C| < \frac{1}{2^n}$.

2. If $x \in \{0\}^*$, then output $P_A(0^*, n)$. If $x \in \{1\}^*$, then output $P_B(1^*, n)$.

Marker Lemma - Algorithm

Let m be the memory of the Markov systems.

1. Input: a string $x \in \Sigma^*$ and $n \in \mathbb{N}$ where we require $|H_A - H_C|, |H_B - H_C| < \frac{1}{2^n}$.

2. If $x \in \{0\}^*$, then output $P_A(0^*, n)$. If $x \in \{1\}^*$, then output $P_B(1^*, n)$.

3. If $|x| < m + 1$, then adjust the probabilities of the alphabet symbols in C such that the entropy condition holds.

Marker Lemma - Algorithm

Let m be the memory of the Markov systems.

1. Input: a string $x \in \Sigma^*$ and $n \in \mathbb{N}$ where we require $|H_A - H_C|, |H_B - H_C| < \frac{1}{2^n}$.

2. If $x \in \{0\}^*$, then output $P_A(0^*, n)$. If $x \in \{1\}^*$, then output $P_B(1^*, n)$.

3. If $|x| < m + 1$, then adjust the probabilities of the alphabet symbols in C such that the entropy condition holds.

4. If $|x| > m + 1$, then compute $P(c|z)$ for all $c \in \Sigma_C$, and $z \in \Sigma^m$, and compute $P_C(x)$.

Skeleton Lemma

Now, $P_A(0) = P_C(0)$. We will identify finite strings from Σ_A^* which can be mapped to finite strings in Σ_C^*.

Skeleton Lemma

Now, $P_A(0) = P_C(0)$. We will identify finite strings from Σ_A^* which can be mapped to finite strings in Σ_C^*.

Idea: We will potentially match strings in A and C if their patterns of 0s is the same.

Skeleton Lemma

Now, $P_A(0) = P_C(0)$. We will identify finite strings from Σ_A^* which can be mapped to finite strings in Σ_C^*.

Idea: We will potentially match strings in A and C if their patterns of 0s is the same.

Let $N_0 < N_1 < \ldots$ be a sequence of positive integers.

Now, $P_A(0) = P_C(0)$. We will identify finite strings from Σ_A^* which can be mapped to finite strings in Σ_C^*.

Idea: We will potentially match strings in A and C if their patterns of 0s is the same.

Let $N_0 < N_1 < \ldots$ be a sequence of positive integers.

Map all non-zero symbols in a sequence (from A or C) to ␣. A *skeleton* of rank r at position i in x, denoted $S(x, r, i)$ is defined as the shortest string enclosing $x[i]$ and delimited by N_r many zeroes on either end.

Skeleton Lemma

Now, $P_A(0) = P_C(0)$. We will identify finite strings from Σ_A^* which can be mapped to finite strings in Σ_C^*.

Idea: We will potentially match strings in A and C if their patterns of 0s is the same.

Let $N_0 < N_1 < \dots$ be a sequence of positive integers.

Map all non-zero symbols in a sequence (from A or C) to ␣. A *skeleton* of rank r at position i in x, denoted $S(x, r, i)$ is defined as the shortest string enclosing $x[i]$ and delimited by N_r many zeroes on either end.

A skeleton $S(x, r, i)$ can be decomposed uniquely into skeletons of rank $r - 1$.

Skeletons

000 00 0 00 00000

$x \in \mathcal{A}$

000 00 0 00 00000

$\phi(x) \in \mathcal{C}$

Skeletons

000 00 0 00 00000

$x \in \mathcal{A}$

Skeletons

000 00 0 00 00000

$\phi(x) \in \mathcal{C}$

Skeletons

000 00 0 00 00000

S_0

$x \in \mathcal{A}$

Skeletons

000 00 0 00 00000

$\phi(x) \in \mathcal{C}$

Skeletons

000 00 0 00 00000

S_0

$x \in \mathcal{A}$ $\longleftrightarrow$

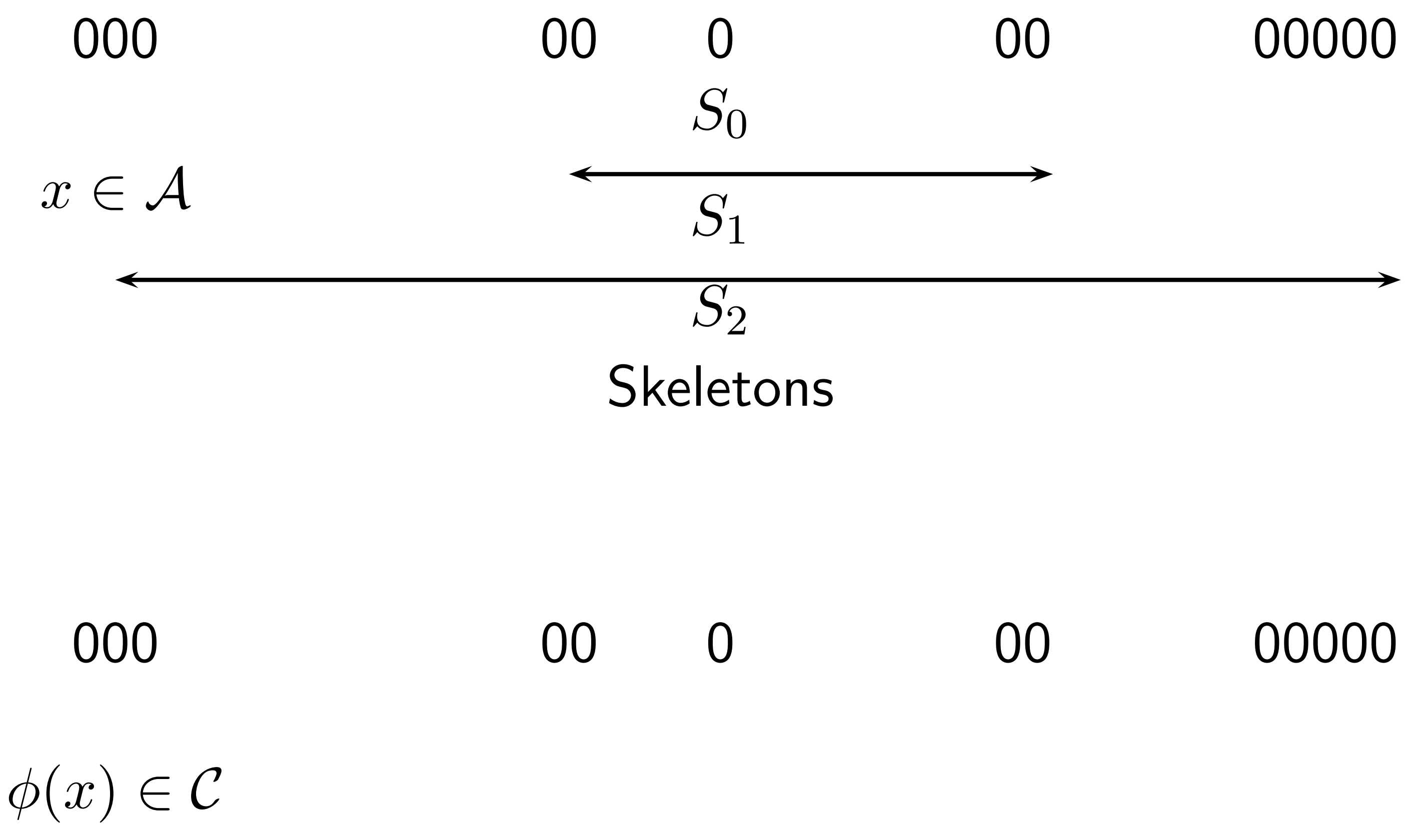
S_1

Skeletons

000 00 0 00 00000

$\phi(x) \in \mathcal{C}$

Skeletons



000
00
0
00
00000
S_0
$x \in \mathcal{A}$
S_1
S_2
Skeletons
000
00
0
00
00000
$\phi(x) \in \mathcal{C}$

Skeleton Lemma

Lemma Let $\langle L_r \rangle_{r=1}^{\infty}$ be an increasing sequence of positive integers. Then there is a layering $\langle K'_r \rangle_{r=1}^{\infty}$ of A and an increasing sequence of positive integers $\langle N_r \rangle_{r=0}^{\infty}$ uniformly computable in r such that for every $r \in \mathbb{N}$ and every $x \in K'_r$, the following hold.

Skeleton Lemma

Lemma Let $\langle L_r \rangle_{r=1}^{\infty}$ be an increasing sequence of positive integers. Then there is a layering $\langle K'_r \rangle_{r=1}^{\infty}$ of A and an increasing sequence of positive integers $\langle N_r \rangle_{r=0}^{\infty}$ uniformly computable in r such that for every $r \in \mathbb{N}$ and every $x \in K'_r$, the following hold.

- □ There is a skeleton centered at $x[0]$ delimited by N_r many zeroes. (denoted $S(x, r, 0)$.)

Skeleton Lemma

Lemma Let $\langle L_r \rangle_{r=1}^{\infty}$ be an increasing sequence of positive integers. Then there is a layering $\langle K'_r \rangle_{r=1}^{\infty}$ of A and an increasing sequence of positive integers $\langle N_r \rangle_{r=0}^{\infty}$ uniformly computable in r such that for every $r \in \mathbb{N}$ and every $x \in K'_r$, the following hold.

- ☐ There is a skeleton centered at $x[0]$ delimited by N_r many zeroes. (denoted $S(x, r, 0)$.)
- ☐ $S(x, r, 0)$ has at least L_r many gaps.

Skeleton Lemma

Lemma Let $\langle L_r \rangle_{r=1}^{\infty}$ be an increasing sequence of positive integers. Then there is a layering $\langle K'_r \rangle_{r=1}^{\infty}$ of A and an increasing sequence of positive integers $\langle N_r \rangle_{r=0}^{\infty}$ uniformly computable in r such that for every $r \in \mathbb{N}$ and every $x \in K'_r$, the following hold.

- □ There is a skeleton centered at $x[0]$ delimited by N_r many zeroes. (denoted $S(x, r, 0)$.)
- □ $S(x, r, 0)$ has at least L_r many gaps.

Proof Idea: If such skeletons occur only finitely often on x, we can form a layerwise computable integrable test that will attain ∞ on x.

Lemma Let $\langle L_r \rangle_{r=1}^{\infty}$ be an increasing sequence of positive integers. Then there is a layering $\langle K'_r \rangle_{r=1}^{\infty}$ of A and an increasing sequence of positive integers $\langle N_r \rangle_{r=0}^{\infty}$ uniformly computable in r such that for every $r \in \mathbb{N}$ and every $x \in K'_r$, the following hold.

- □ There is a skeleton centered at $x[0]$ delimited by N_r many zeroes. (denoted $S(x, r, 0)$.)
- □ $S(x, r, 0)$ has at least L_r many gaps.

Proof Idea: If such skeletons occur only finitely often on x, we can form a layerwise computable integrable test that will attain ∞ on x.
Then $x \notin$ MLR.

Filler Lemma

We have identified potential matches between elements in A and C based on identical skeletons. We have to decide what goes in the gaps.

Filler Lemma

We have identified potential matches between elements in A and C based on identical skeletons. We have to decide what goes in the gaps.

Let η_r and θ_r denote the minimum and the maximum conditional probabilites of symbols in A and C at precision r. Fix a sequence of numbers L_r, $r = 1, 2, \ldots$ such that

$$\lim_{r \to \infty} \frac{1}{\eta_r} 2^{-L_r(1/2^r)} = 0.$$

Filler Lemma

We have identified potential matches between elements in A and C based on identical skeletons. We have to decide what goes in the gaps.

Let η_r and θ_r denote the minimum and the maximum conditional probabilites of symbols in A and C at precision r. Fix a sequence of numbers L_r, $r = 1, 2, \ldots$ such that

$$\lim_{r\to\infty} \frac{1}{\eta_r} 2^{-L_r(1/2^r)} = 0.$$

Let $S(x, r, i)$ have ℓ blanks in positions $s_1, s_2, \ldots, s_\ell$. We fix the filler alphabet as Σ_A^ℓ and Σ_C^ℓ in A and C respectively.

Filler Lemma

We have identified potential matches between elements in A and C based on identical skeletons. We have to decide what goes in the gaps.

Let η_r and θ_r denote the minimum and the maximum conditional probabilites of symbols in A and C at precision r. Fix a sequence of numbers L_r, $r = 1, 2, \ldots$ such that

$$\lim_{r\to\infty} \frac{1}{\eta_r} 2^{-L_r(1/2^r)} = 0.$$

Let $S(x, r, i)$ have ℓ blanks in positions $s_1, s_2, \ldots, s_\ell$. We fix the filler alphabet as Σ_A^ℓ and Σ_C^ℓ in A and C respectively.

For a filler F, let $J(F, n) \subseteq \{s_1, \ldots, s_\ell\}$ be an index set of the positions in S filled by F.

Filler Lemma

We have identified potential matches between elements in A and C based on identical skeletons. We have to decide what goes in the gaps.

Let η_r and θ_r denote the minimum and the maximum conditional probabilites of symbols in A and C at precision r. Fix a sequence of numbers L_r, $r = 1, 2, \ldots$ such that

$$\lim_{r \to \infty} \frac{1}{\eta_r} 2^{-L_r(1/2^r)} = 0.$$

Let $S(x, r, i)$ have ℓ blanks in positions $s_1, s_2, \ldots, s_\ell$. We fix the filler alphabet as Σ_A^ℓ and Σ_C^ℓ in A and C respectively.

For a filler F, let $J(F, n) \subseteq \{s_1, \ldots, s_\ell\}$ be an index set of the positions in S filled by F.

Define an equivalence relation $\sim_n$: $F \sim_n F'$ if $J(F, n) = J(F', n)$ and F agrees with F' on $J(F, n)$.

Constructing $J(F, n)$

The equivalence classes are constructed inductively on the rank.

Constructing $J(F, n)$

The equivalence classes are constructed inductively on the rank.

For a rank 1 skeleton, set $J(F, n)$ to be the largest subset P of $\{s_1, \ldots, s_\ell\}$ such that the probability of the cylinder specified by P is at least $\frac{3}{2\eta_1} 2^{-L_1(H-\varepsilon_1)}$.

Constructing $J(F, n)$

The equivalence classes are constructed inductively on the rank.

For a rank 1 skeleton, set $J(F, n)$ to be the largest subset P of $\{s_1, \ldots, s_\ell\}$ such that the probability of the cylinder specified by P is at least $\frac{3}{2\eta_1} 2^{-L_1(H-\varepsilon_1)}$.

Let $P_r = \{s_{j_1}, \ldots, s_{j_k}\}$ be fixed by skeletons of rank $\leq r$.

The equivalence classes are constructed inductively on the rank.

For a rank 1 skeleton, set $J(F,n)$ to be the largest subset P of $\{s_1,\ldots,s_\ell\}$ such that the probability of the cylinder specified by P is at least $\frac{3}{2\eta_1}2^{-L_1(H-\varepsilon_1)}$.

Let $P_r=\{s_{j_1},\ldots,s_{j_k}\}$ be fixed by skeletons of rank $\leq r$.

For a skeleton of rank $r+1$, pick the largest subset of P of $\{s_1,\ldots,s_\ell\}-P_r$ so that the probability of the cylinder specified by $P_r\cup P$ is at least

$$\frac{(1+\varepsilon_r)}{\eta_r}2^{-L_r(H-\varepsilon_r)}.$$

The equivalence classes are constructed inductively on the rank.

For a rank 1 skeleton, set $J(F,n)$ to be the largest subset P of $\{s_1,\ldots,s_\ell\}$ such that the probability of the cylinder specified by P is at least $\frac{3}{2\eta_1}2^{-L_1(H-\varepsilon_1)}$.

Let $P_r=\{s_{j_1},\ldots,s_{j_k}\}$ be fixed by skeletons of rank $\leq r$.

For a skeleton of rank $r+1$, pick the largest subset of P of $\{s_1,\ldots,s_\ell\}-P_r$ so that the probability of the cylinder specified by $P_r\cup P$ is at least

$$\frac{(1+\varepsilon_r)}{\eta_r}2^{-L_r(H-\varepsilon_r)}.$$

Then $J(F,n)$ for a rank $r+1$ skeleton is $P_r\cup P$.

Filler Lemma

Lemma There is a layering $\langle K''_p \rangle_{p=1}^{\infty}$ such that or every n, there is a large enough r such that for every skeleton S of rank r and length ℓ corresponding to $x \in K''_r$, we have:

Filler Lemma

Lemma There is a layering $\langle K''_p \rangle_{p=1}^{\infty}$ such that or every n, there is a large enough r such that for every skeleton S of rank r and length ℓ corresponding to $x \in K''_r$, we have:

1. For all $F \in \mathcal{F}(S)$,
$-\log_{\mathbf{2}} \mathbf{P_A}\left(\tilde{\mathbf{F}}_{\mathbf{r}}, \mathbf{n}\right) + \log_2(1 - \varepsilon_n) \leq \mathbf{L}(\mathbf{H} - \varepsilon_{\mathbf{r}})$

Filler Lemma

Lemma There is a layering $\langle K''_p \rangle_{p=1}^{\infty}$ such that or every n, there is a large enough r such that for every skeleton S of rank r and length ℓ corresponding to $x \in K''_r$, we have:

1. For all $F \in \mathcal{F}(S)$,
   $-\log_{\mathbf{2}} \mathbf{P_A}\left(\tilde{\mathbf{F}}_{\mathbf{r}}, \mathbf{n}\right) + \log_2(1-\varepsilon_n) \leq \mathbf{L}(\mathbf{H} - \varepsilon_{\mathbf{r}})$
2. For all $F \in \mathcal{F}(S)$ except maybe on a set of measure ε_n:
   (a) $-\log_{\mathbf{2}} \mathbf{P_A}(\tilde{\mathbf{F}}_{\mathbf{r}}, \mathbf{n}) + \log_2 \frac{(1-\varepsilon_r)\eta_r}{(1+\varepsilon_n)^2} > \mathbf{L}(\mathbf{H} - \varepsilon_{\mathbf{r}})$

Filler Lemma

Lemma There is a layering $\langle K''_p \rangle_{p=1}^{\infty}$ such that or every n, there is a large enough r such that for every skeleton S of rank r and length ℓ corresponding to $x \in K''_r$, we have:

1. For all $F \in \mathcal{F}(S)$,
   $-\log_{\mathbf{2}} \mathbf{P_A}\left(\tilde{\mathbf{F}}_{\mathbf{r}}, \mathbf{n}\right) + \log_2(1-\varepsilon_n) \leq \mathbf{L}(\mathbf{H} - \varepsilon_{\mathbf{r}})$
2. For all $F \in \mathcal{F}(S)$ except maybe on a set of measure ε_n:
   (a) $-\log_{\mathbf{2}} \mathbf{P_A}(\tilde{\mathbf{F}}_{\mathbf{r}}, \mathbf{n}) + \log_2 \frac{(1-\varepsilon_r)\eta_r}{(1+\varepsilon_n)^2} > \mathbf{L}(\mathbf{H} - \varepsilon_{\mathbf{r}})$
   (b) $\frac{1}{L}|J(F,r)| > 1 - \frac{3}{|\log_2 \theta_r|}\varepsilon_r$

where $L = \ell + |Z_S|$.

Lemma There is a layering $\langle K''_p \rangle_{p=1}^{\infty}$ such that or every n, there is a large enough r such that for every skeleton S of rank r and length ℓ corresponding to $x \in K''_r$, we have:

1. For all $F \in \mathcal{F}(S)$,
   $-\log_{\mathbf{2}} \mathbf{P_A}\left(\tilde{\mathbf{F}}_{\mathbf{r}}, \mathbf{n}\right) + \log_2(1-\varepsilon_n) \leq \mathbf{L}(\mathbf{H} - \varepsilon_{\mathbf{r}})$
2. For all $F \in \mathcal{F}(S)$ except maybe on a set of measure ε_n:
   (a) $-\log_{\mathbf{2}} \mathbf{P_A}(\tilde{\mathbf{F}}_{\mathbf{r}}, \mathbf{n}) + \log_2 \frac{(1-\varepsilon_r)\eta_r}{(1+\varepsilon_n)^2} > \mathbf{L}(\mathbf{H} - \varepsilon_{\mathbf{r}})$
   (b) $\frac{1}{L}|J(F,r)| > 1 - \frac{3}{|\log_2 \theta_r|}\varepsilon_r$

where $L = \ell + |Z_S|$.

Proof Idea: Estimates follow from the effective Shannon-McMillan-Breiman theorem [Hoc09], [Hoy12].

Marriage Lemma

We have a bipartite graph G with left set : $\sim_n$-equivalence classes of fillers for A, and right set : $\sim_n$-equivalence classes of fillers for B. Each vertex $\tilde{F}$ on the left has probability $P_A(\tilde{F})$ and each vertex $\tilde{G}$ on the right has probability $P_C(\tilde{G})$.

Marriage Lemma

We have a bipartite graph G with left set : $\sim_n$-equivalence classes of fillers for A, and right set : $\sim_n$-equivalence classes of fillers for B. Each vertex $\tilde{F}$ on the left has probability $P_A(\tilde{F})$ and each vertex $\tilde{G}$ on the right has probability $P_C(\tilde{G})$.

A *society* f is a relation so that for every subset S of left vertices, $P_A(S) \leq P_C(f(S))$.

Marriage Lemma

We have a bipartite graph G with left set : $\sim_n$-equivalence classes of fillers for A, and right set : $\sim_n$-equivalence classes of fillers for B. Each vertex $\tilde{F}$ on the left has probability $P_A(\tilde{F})$ and each vertex $\tilde{G}$ on the right has probability $P_C(\tilde{G})$.

A *society* f is a relation so that for every subset S of left vertices, $P_A(S) \leq P_C(f(S))$.

This implies that for every subset T of right vertices, $P_C(T) \leq P_A(f^{-1}T)$. (*i.e.* The "dual" graph also defines a society.)

Marriage Lemma

We have a bipartite graph G with left set : $\sim_n$-equivalence classes of fillers for A, and right set : $\sim_n$-equivalence classes of fillers for B. Each vertex $\tilde{F}$ on the left has probability $P_A(\tilde{F})$ and each vertex $\tilde{G}$ on the right has probability $P_C(\tilde{G})$.

A *society* f is a relation so that for every subset S of left vertices, $P_A(S) \leq P_C(f(S))$.

This implies that for every subset T of right vertices, $P_C(T) \leq P_A(f^{-1}T)$. (*i.e.* The "dual" graph also defines a society.)

A *minimal society* is a society where the removal of any edge violates the condition for a society.

Marriage Lemma

Every society has a minimal subsociety which is produced by a joining - that is, a joint distribution on $L \times R$ with marginals P_A and P_C.

Marriage Lemma

Every society has a minimal subsociety which is produced by a joining - that is, a joint distribution on $L \times R$ with marginals P_A and P_C.

In a minimal subsociety, there is at least one vertex on the right which knows at most one left vertex.

Marriage Lemma

Every society has a minimal subsociety which is produced by a joining - that is, a joint distribution on $L \times R$ with marginals P_A and P_C.

In a minimal subsociety, there is at least one vertex on the right which knows at most one left vertex.

Our modification: a society is called ϵ-robust if for every left set S, $P_A(S)(1+\epsilon) \quad \leq \quad P_B(f(S))(1-\epsilon)$, and for every right set T, $P_B(1-\epsilon) \quad \leq \quad P_A(f^{-1}(T))(1+\epsilon)$,

Assignment Lemma

Lemma 7 (Assignment Lemma). *If $x \in A$ such that $x \in K_{r'} \cap K'_{r'}$, with $x[0]$ not contained in a block of 0 longer than m, then there is an even r, computable from r', such that*

1. *With respect to the society $R_{S_r(x)} : \bar{\mathcal{G}}(S_r(x)) \rightsquigarrow \tilde{\mathcal{F}}(S_r(x))$, $R^{-1}_{S_r(x)}(\tilde{\mathcal{F}}_r(x))$ is a singleton, say, $\bar{\mathcal{G}}_r(x)$.*
2. *$i_r(x) \in J_0(\bar{\mathcal{G}}_r(x))$.*

Lemma 7 (Assignment Lemma). *If $x \in A$ such that $x \in K_{r'} \cap K'_{r'}$, with $x[0]$ not contained in a block of 0 longer than m, then there is an even r, computable from r', such that*

1. *With respect to the society $R_{S_r(x)} : \bar{\mathcal{G}}(S_r(x)) \rightsquigarrow \tilde{\mathcal{F}}(S_r(x))$, $R^{-1}_{S_r(x)}(\tilde{\mathcal{F}}_r(x))$ is a singleton, say, $\bar{\mathcal{G}}_r(x)$.*
2. *$i_r(x) \in J_0(\bar{\mathcal{G}}_r(x))$.*

Intuitively, this lemma says that $\phi(x)[0]$ is determined from some long enough central cylinder of x.

Assignment Lemma

Lemma 7 (Assignment Lemma). *If $x \in A$ such that $x \in K_{r'} \cap K'_{r'}$, with $x[0]$ not contained in a block of 0 longer than m, then there is an even r, computable from r', such that*

1. *With respect to the society $R_{S_r(x)} : \bar{\mathcal{G}}(S_r(x)) \rightsquigarrow \tilde{\mathcal{F}}(S_r(x))$, $R^{-1}_{S_r(x)}(\tilde{\mathcal{F}}_r(x))$ is a singleton, say, $\bar{\mathcal{G}}_r(x)$.*
2. *$i_r(x) \in J_0(\bar{\mathcal{G}}_r(x))$.*

Intuitively, this lemma says that $\phi(x)[0]$ is determined from some long enough central cylinder of x.

ϕ commutes with T_A and T_C. The image of Martin-Löf points under measure-preserving transformations is Martin-Löf random. Hence for $x \in \mathrm{MLR}_A$, every co-ordinate of $\phi(x)$ is determined in a layerwise computable way.

Thank You.

[Hoc09] Michael Hochman. Upcrossing inequalities for stationary sequences and applications. *Annals of Probability*, 37(6):2135–2149, 2009.

[Hoy12] Mathieu Hoyrup. The dimension of ergodic random sequences. In *Symposium on Theoretical Aspects of Computer Science*, pages 567–576, 2012.

[KS79] Michael Keane and Meier Smorodinsky. Bernoulli schemes of the same entropy are finitarily isomorphic. *Annals of Mathematics*, 109(2):397–406, 1979.

[Orn70] Donald Ornstein. Bernoulli shifts with the same entropy are isomorphic. *Advances in Mathematics*, 4(3):337 – 352, 1970.