Dimension 1 sequences are close to randoms

Linda Brown Westrick
University of Connecticut
Joint with N. Greenberg, J. Miller and A. Shen

October 23, 2016
Midwest Computability Seminar
in honor of Carl Jockusch's 75th birthday
University of Chicago

Outline

The *Besicovitch pseudo-distance between* two sequences $X, Y \in 2^\omega$ is the upper density of their symmetric difference.

$$d(X,Y) = \limsup_{n\to\infty} \frac{|X \restriction_n \Delta Y \restriction_n|}{n}$$

A sequence $X \in 2^\omega$ is

- (Martin-Löf) *random* if $K(X \restriction_n) \geq n - O(1)$
- of (effective) *dimension* at least s if $K(\restriction_n) \geq sn - o(n)$
- *weakly s-random* if $K(X \restriction_n) \geq sn - O(1)$

In this talk we discuss how far (and near) the following kinds of sequences can and must be from each other.

- Dimension 1 sequences and randoms
- Dimension s sequences and randoms
- Dimension s sequences and weakly s-randoms

Dimension 1 sequences and randoms

If $d(X, Y) = 0$, then X and Y are said to be *coarsely similar*.

Theorem 1

A sequence X has dimension 1 if and only if it is coarsely similar to a random.

Proof of ($\Leftarrow$) direction:
Suppose Y is random and $d(X, Y) = 0$. Then for each $\varepsilon > 0$, for sufficiently large n we can code $Y \upharpoonright n$ by providing $X \upharpoonright n$ and a list of the locations of at most εn bits to be flipped. Therefore,

$$K(Y \upharpoonright_n) \leq K(X \upharpoonright n) + O(\varepsilon n \log n).$$

Since this holds for arbitrary ε, X has dimension 1.

Harper's Theorem

For $\sigma, \tau \in 2^n$, let

$$d(\sigma, \tau) = \frac{|\sigma \Delta \tau|}{n}$$

For $A, B \subseteq 2^n$, let

$$d(A, B) = \min(d(\sigma, \tau) : \sigma \in A, \tau \in B).$$

The *Hamming ball of radius* r *centered at* σ is

$$B_r(\sigma) := \{\tau : d(\sigma, \tau) \leq r\}.$$

A *Hamming sphere* centered at σ is a set S such that for some r,

$$B_r(\sigma) \subseteq S \subseteq B_{r+1/n}(\sigma).$$

Harper's Theorem

For any $A, B \subseteq 2^n$, there are Hamming spheres A' and B', centered at 0^n and 1^n respectively, such that $|A| = |A'|$, $|B| = |B'|$, and $d(A', B') \geq d(A, B)$.

Entropy and density

The *Shannon entropy function* is

$$H(p) = p \log p + (1-p)\log(1-p)$$

where $0 \log 0 = 0$ by convention.

This function relates the size of a Hamming ball to its radius.

- For any $\sigma \in 2^n$,

$$H(p)n - o(n) \leq \log |B_p(\sigma)| \leq H(p)n.$$

- If X has asymptotic density p, then its dimension is at most $H(p)$.
- The Bernoulli p-random sequences have dimension exactly $H(p)$.

A finite application

Theorem (essentially Buhrman, Fortnow, Newman& Vereshchagin)

For every $\epsilon > 0$ there is a $q < 1$ such that for sufficiently large n, if $\tau \in 2^n$ with $K(\tau) > qn$, then there is $\sigma \in B_\epsilon(\tau)$ with $K(\sigma) > n$.

Proof. Given ϵ, let q be larger than $H(1/2 - \varepsilon)$.

- Let A be the set of random strings ($K(\sigma) \geq n$).
- Let B be $\{\rho : d(\rho, A) > \epsilon\}$, so $d(A, B) > \epsilon$.
- Harper's Theorem provides A', B', Hamming spheres.
- $B_{1/2}(0^n) \subseteq A'$, because A contains at least half the strings.
- So $B' \subseteq B_{1/2-\varepsilon}(1^n)$.
- So $\log |B| \leq \log |B_{1/2-\varepsilon}(1^n)| \leq H(1/2 - \varepsilon)n < qn$.
- And B is c.e. So for sufficiently large n, if $\rho \in B$, $K(\rho) < qn$.
- Given τ with $K(\tau) \geq qn$, $\tau \notin B$, so $d(\tau, A) \leq \varepsilon$.

Theorem 1 proof sketch

A finite extension construction that fails:

Given X of dimension 1, we want a random Y with $d(X,Y) = 0$. Let P be a Π^0_1 class of randoms.

- Build $\tau_0\tau_1\tau_2 \cdots \prec Y$.
- (Corresponding to $\sigma_0\sigma_1\sigma_2 \cdots \prec \ldots X$.)
- Maintain $\tau_0 \ldots \tau_n$ extendible in P while waiting for the dimension of X to rise for good above some $q = H(1/2 - \varepsilon)$.
- Since (roughly) $K(\sigma_{n+1}|\sigma_0 \ldots \sigma_n) \geq q|\sigma_{n+1}|$, adapt the finite case to find a random extension τ_{n+1} which is ε-close to σ_{n+1}.

Problem: $\tau_0 \ldots \tau_n$ has more information than $\sigma_0 \ldots \sigma_n$. The opponent can copy the extra information to σ_{n+1}. So $K(\sigma_{n+1}|\tau_0 \ldots \tau_n) < q|\sigma_{n+1}|$.

Solution: At each stage, consider not only one $\tau_0 \ldots \tau_n$, but all $\tau_0 \ldots \tau_n$ which are in P and close to $\sigma_0 \ldots \sigma_n$, and extend one which the opponent did not copy information from. Then apply compactness.

Dimension s sequences to randoms

Theorem 2

For every sequence X of dimension s, there is a random Y such that $d(X, Y) \leq \frac{1}{2} - H^{-1}(s)$.

- Here H^{-1} picks out the smaller of two possible values.
- The result is optimal. If X is a Bernoulli p-random sequence, its dimension is $s = H(p)$, and its distance to a random is at least $1/2 - p$.
- To prove it, modify the previous construction.
- If $s_n = K(\sigma_{n+1} | \sigma_0 \ldots \sigma_n)$, choose $\tau_0 \ldots \tau_{n+1}$ in the tree of randoms to satisfy $d(\tau_i, \sigma_i) \leq \frac{1}{2} - H^{-1}(s_i)$.
- Apply concavity of $\frac{1}{2} - H^{-1}(s_i)$.

Randoms to dimension s sequences

Proposition 3

If $d(X, Y) \leq d$, then $\dim(Y) \leq \dim(X) + H(d)$.

Proof. To give a code for $Y \upharpoonright_n$, provide $X \upharpoonright_n$ and a description of the dn changes.

$$K(Y \upharpoonright_n) \leq K(X \upharpoonright_n) + \log |B_d(0^n)| + O(1).$$

Recall that $\log |B_d(0^n)| \leq H(d)n$.

Corollary: If Y is random and $\dim(X) = s$, then $d(X, Y) \geq H^{-1}(1 - s)$.

Theorem 3

If Y is random, then there is a sequence X of dimension s with $d(X, Y) = H^{-1}(1 - s)$.

Follows from finite version.

Dimension s sequences and weakly s-randoms

Theorem 4

A sequence X has dimension s if and only if it is coarsely similar to a weakly s-random.

Finite version (via Harper's Theorem): For all $s < 1$ and ε, there is a δ such that for all sufficiently large n and all $\sigma \in 2^n$ with $K(\sigma) = sn$, there is $\tau \in B_\varepsilon(\sigma)$ with $K(\tau) \geq (s + \delta)n$.

Proof idea: Using density ε of changes, start building Y as if it were a dimension $(s + \delta)$ sequence, so that a buffer of extra information is built up: $K(Y \restriction_n) > (s + \delta)n$. Use compactness to keep the opponent from eating into the buffer. When the buffer is large enough, safely decrease ε.

Dimension s sequences and dimension t sequences

Distance from an arbitrary A to the nearest B, where $1 > t > s$.

	$\dim(B) = 1$	$\dim(B) = t$	$\dim(B) = s$
$\dim(A) = 1$	0	$H^{-1}(1-t)$	$H^{-1}(1-s)$
$\dim(A) = t$	$\frac{1}{2} - H^{-1}(t)$	0	strictly $> H^{-1}(t-s)$
$\dim(A) = s$	$\frac{1}{2} - H^{-1}(s)$	at least $H^{-1}(t) - H^{-1}(s)$	0

Question

For every $s < t < 1$ and every X of dimension s, is there a Y of dimension t within distance $H^{-1}(t) - H^{-1}(s)$ of X?