Outline of Lecture 5

Randomness in Fractal Geometry and Dynamical Systems

- Randomness for Hausdorff Measures.
- Kolmogorov Complexity and Hausdorff Dimension.
- Frostman's Lemma.
- Extracting Randomness.
- Selection Rules, Lowness, and Triviality

Hausdorff Measures

- For $s \geq 0$, define set function

$$\mathcal{H}_\delta^s(E) = \inf \left\{ \sum_i d(U_i)^s : E \subseteq \bigcup_i U_i,\ d(U_i) \leq \delta \right\}.$$

- Letting $\delta \to 0$ yields an outer measure.
- The s-dimensional Hausdorff measure $\mathcal{H}^s$ is defined as

$$\mathcal{H}^s(E) = \lim_{\delta \to 0} \mathcal{H}_\delta^h(E)$$

Properties of Hausdorff Measures

- $\mathcal{H}^s$ is Borel regular:
  all Borel sets B are measurable, i.e.

$$(\forall A \subseteq X)\, \mathcal{H}^s(A) = \mathcal{H}^s(A \cap B) + \mathcal{H}^s(A \setminus B),$$

  and for all $A \subseteq X$ there is a Borel set $B \subseteq A$ such that

$$\mathcal{H}^s(B) = \mathcal{H}^s(A).$$

- For $X = \mathbb{R}^n$ (Euclidean) and $s = n$, $\mathcal{H}^n$ yields the usual Lebesgue measure λ (up to a multiplicative constant).

From Measure to Dimension

- Important property: For $0 \leq s < t < \infty$ und $E \subseteq X$,

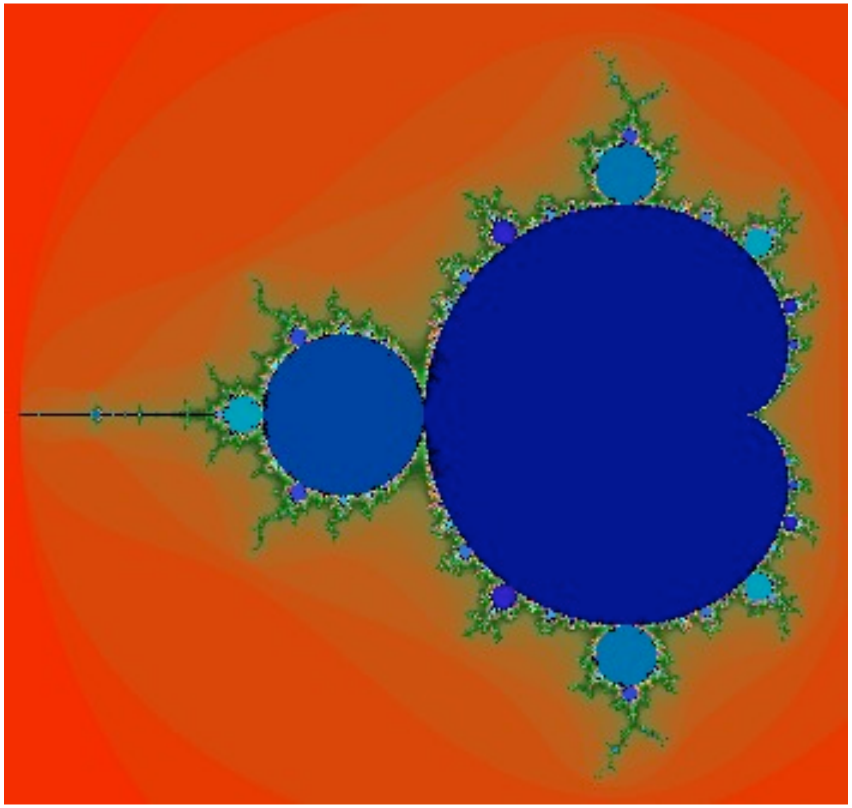
$$\begin{aligned} \mathcal{H}^s(E) < \infty \quad &\Rightarrow \quad \mathcal{H}^t(E) = 0, \\ \mathcal{H}^t(E) > 0 \quad &\Rightarrow \quad \mathcal{H}^s(E) = \infty. \end{aligned}$$

- The Hausdorff dimension of a set E is defined as

$$\begin{aligned} \dim_H(E) \quad &= \quad \inf\{s \geq 0 : \mathcal{H}^s(E) = 0\} \\ &= \quad \sup\{t \geq 0 : \mathcal{H}^t(E) = \infty\} \end{aligned}$$

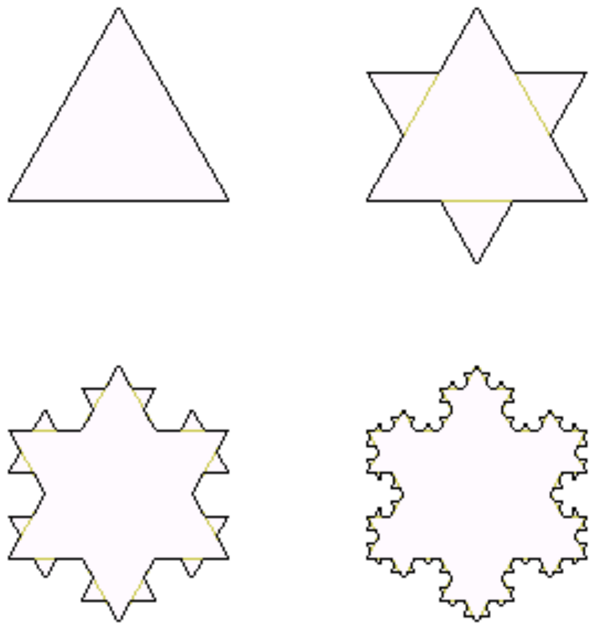
Hausdorff Dimension – Famous examples

Mandelbrot set – $\dim_H = 2$



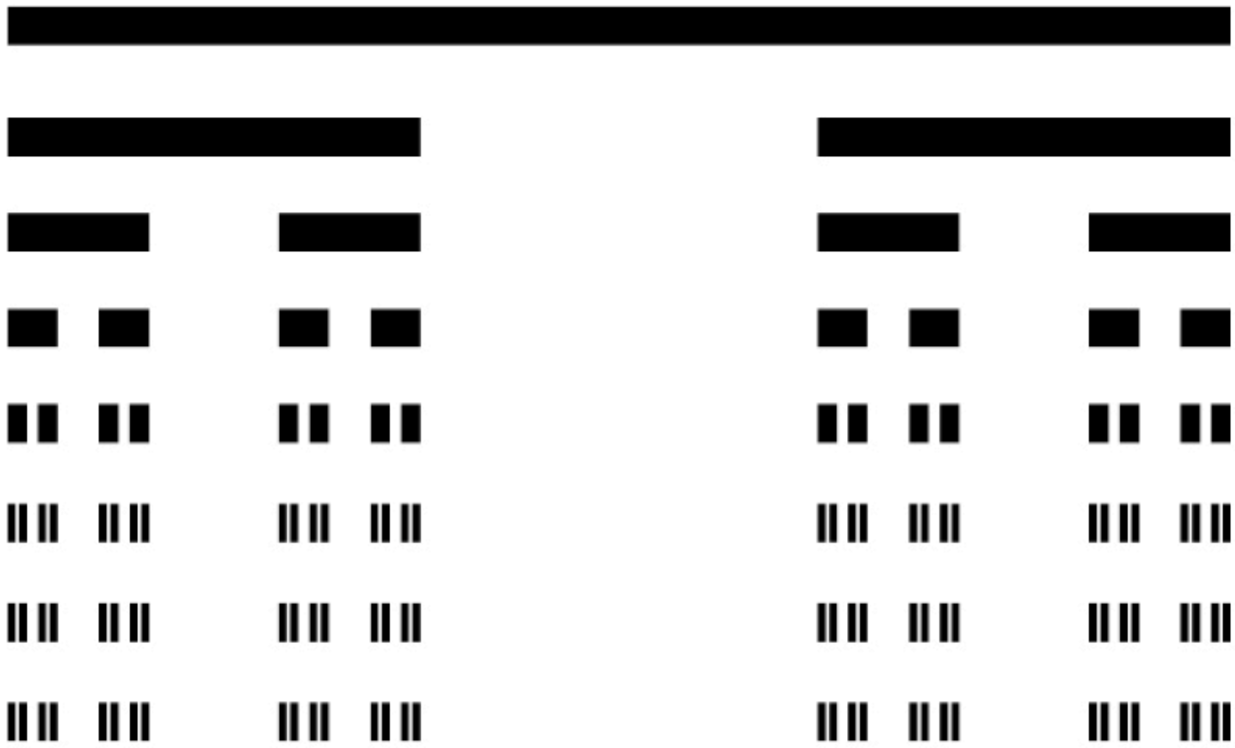
Hausdorff Dimension – Famous examples

Koch snowflake – $\dim_H = \log 4 / \log 3$



Hausdorff Dimension – Famous examples

Cantor set – $\dim_H = \log 2 / \log 3$



Hausdorff Dimension – Famous examples

Frequency sets – For $0 \le p \le 1$, let

$$A_p = \left\{ X \in 2^{\mathbb{N}} : \lim_n \frac{|\{i < n : X(i) = 1\}|}{n} = p \right\}.$$

Then $\dim_H A_p = H(p) = -[p \log p + (1-p)\log(1-p)]$ [Eggleston].

Properties of Hausdorff Dimension

- **Lebesgue measure:** $\lambda(A) > 0$ implies $\dim_H(A) = 1$.
- **Monotony:** $A \subseteq B$ implies $\dim_H(A) \leq \dim_H(B)$.
- **Stability:** For $A_1, A_2, \dots \subseteq 2^{\mathbb{N}}$ it holds that

$$\dim_H(\bigcup A_i) = \sup\{\dim_H(A_i)\}.$$

- Important **geometric properties:**
  - If F is **Hölder continuous**, i.e. if there are constants $c, r > 0$ for which

    $$(\forall x, y)\ d(F(x), F(y)) \leq c\, d(x, y)^r,$$

    then

    $$\dim_H F(A) \leq (1/r) \dim_H(A).$$

  - For $r = 1$, F is **Lipschitz continuous**. If F is bi-Lipschitz, then

    $$\dim_H h(A) = \dim_H(A).$$

Hausdorff Dimension and Martingales

Hausdorff dimension can be expressed in terms of martingales.

- Given $s \geq 0$, a martingale F is called s-successful on a real $X \in 2^{\mathbb{N}}$ if

$$\limsup \frac{F(X \upharpoonright_n)}{2^{(1-s)n}} = \infty.$$

- Note that the usual success-notion for martingales is just being 1-successful.

Theorem [Lutz]

For any set $A \subseteq 2^{\mathbb{N}}$,

$$\dim_H A = \inf\{s : \exists \text{ martingale } F \ s\text{-successful on all } X \in A\}.$$

Packing Dimension

Lutz' martingale characterization allows for an easy characterization of another dimension concept, packing dimension, which can be seen as a dual to Hausdorff dimension.

- Instead of "covering" a set with open balls, "pack" it with disjoint balls.

Given $0 < s \leq 1$, a martingale F is strongly s-successful on a real X if

$$\liminf \frac{F(X\upharpoonright_n)}{2^{(1-s)n}} \to \infty.$$

Theorem [Athreya, Hitchcock, Lutz, and Mayordomo]

For any set $A \subseteq 2^{\mathbb{N}}$,

$$\dim_P A = \inf\{s : \exists\, F \text{ strongly } s\text{-successful on all } X \in A\}.$$

Effective Hausdorff Dimension

The effective Hausdorff dimension, or constructive dimension), of $A \subseteq 2^{\mathbb{N}}$ is defined as

$$\dim_H^1 A = \inf \{s \in \mathbb{Q}_0^+ : A \text{ is effectively } \mathcal{H}^s\text{-null}\}.$$

- Effective dimension has an important stability property [Lutz]:

$$\dim_H^1 \mathcal{A} = \sup \{\dim_H^1\{X\} : X \in \mathcal{A}\}.$$

- For a single real $X \in 2^{\mathbb{N}}$, we put $\dim_H^1 X = \dim_H^1\{X\}$. There are single reals of non-zero dimension: every λ-random real has dimension one.

Effective Dimension and Kolmogorov Complexity

Effective Hausdorff dimension can be interpreted as a degree of incompressibility.

Theorem (Ryabko; Mayordomo)

For every real X,

$$\dim_H^1 X = \liminf_{n \to \infty} \frac{K(X \upharpoonright_n)}{n}.$$

Effective Dimension and Kolmogorov Complexity

Effective packing dimension (constructive strong dimension) can be effectivized using the martingale characterization by Athreya et al.

Theorem (Athreya et al)

For every real X,

$$\dim^1_P X = \limsup_{n\to\infty} \frac{K(X\restriction_n)}{n}.$$

The three basic examples

Let $0 < r < 1$ rational. Given a Martin-Löf random set X, define X_r by

$$X_r(m) = \begin{cases} X(n) & \text{if } m = \lfloor n/r \rfloor, \\ 0 & \text{otherwise.} \end{cases}$$

Then $\dim_H^1 X_r = r$.

- Geometry: Hölder transformation of Cantor set
- Information theory: Insert redundancy

The three basic examples

Let μ_p be a Bernoulli ("coin-toss") measure with bias $p \in \mathbb{Q} \cap [0, 1]$, and let X be random with respect to μ_p.
Then

$$\dim_H^1 X = H(\mu_p) := -[p \log p + p \log(1 - p)].$$

[Lutz; Eggleston]

- Kolmogorov complexity can be seen as an effective version of entropy.

The three basic examples

Let $\mathbb{U}$ be a universal, prefix-free machine. Given a computable real number $0 < s \leq 1$, the binary expansion of the real number

$$\Omega^{(s)} = \sum_{\sigma \in \mathrm{dom}(\mathbb{U})} 2^{-\frac{|\sigma|}{s}}$$

has effective dimension s [Tadaki].

- Note that $\Omega^{(1)}$ is just Chaitin's Ω.

Randomness Extraction

Each of the three examples actually computes a Martin-Löf random real.

- This is obvious for the "diluted" sequence.
- For recursive Bernoulli measures, one may use Von-Neumann's trick to turn a biased random real into a uniformly distributed random real.
  More generally, any real which is random with respect to a recursive measure computes a Martin-Löf random real [Levin; Kautz].
- $\Omega^{(s)}$ computes a fixed-point free function. It is of r.e. degree, and hence it follows from the Arslanov completeness criterion that $\Omega^{(s)}$ is Turing complete (and thus T-equivalent to a Martin-Löf random real).

The Dimension Problem

The stability property implies that the Turing lower cone of each of the three examples has effective dimension 1.

Question

Are there any Turing lower cones of non-integral dimension?

- Any such lower cone would come from a real of non-integral dimension for which it is not possible to extract some content of higher degree of randomness effectively.

The Dimension Problem

Construction of reals of positive dimension for which randomness cannot be extracted:

- For $\leq_{\mathrm{m}}$ [Reimann and Terwijn, 2004]
- For $\leq_{\mathrm{wtt}}$ [Reimann and Nies, 2007]

Finally, J. Miller [2010] constructed a real with $\dim^1_{\mathrm{H}} = 1/2$ that does not compute any real of dimension $> 1/2$.

Hausdorff Dimension

Mass Distribution Principle

Support of a probability measure

$\mathrm{supp}(\mu)$ is the smallest closed set F such that $\mu(2^{\mathbb{N}} \setminus F) = 0$.

$A \subseteq 2^{\mathbb{N}}$ **supports** a measure μ if $\mathrm{supp}(\mu) \subseteq A$.

Mass Distribution Principle

If A supports a probability measure μ such that for all σ,

$$\mu(\sigma) \leq c2^{-|\sigma|s},$$

then $\dim_H A \geq s$.

Hausdorff Dimension

Frostman's Lemma

A fundamental result due to Frostman (1935) asserts that the converse holds, too, as long as A is not too complicated.

Frostman's Lemma

If A is analytic and $\dim_H A > s > 0$, then there exists a probability measure μ such that $\operatorname{supp}(\mu) \subseteq A$ and for some $c > 0$,

$$(\forall \sigma)\ \mu(\sigma) \leq c2^{-|\sigma|s}.$$

(Call such a measure s-bounded.)

The theorem can be interpreted in the framework of capacity theory. Define the capacitary dimension of A to be

$$\dim_c(A) = \sup\{s : \ A \text{ supports an } s\text{-bounded prob. measure}\}.$$

Then we have for analytic sets, $\dim_c = \dim_H$.

A pointwise version of Frostman's Lemma

We will prove a pointwise version of Frostman's Lemma.

The connection with Kolmogorov complexity

An order is a nondecreasing, unbounded function $h : \mathbb{N} \to \mathbb{N}$. h is called convex if for all n, $h(n+1) \leq h(n) + 1$.

Kjos-Hanssen et al called a real complex if for a computable order h

$$(\forall n)\ [K(x \upharpoonright_n) \geq h(n)],$$

where K denotes prefix-free Kolmogorov complexity.

If x is complex via h, then we call x h-complex. Reimann showed that x is h-complex if and only if it is 2^{-h}-random.

A pointwise version of Frostman's Lemma

Randomness and complexity

We need to replace K by another type of Kolmogorov complexity.

A (continuous) semimeasure is a function $\eta : 2^{<\mathbb{N}} \to [0, 1]$ such that

$$\forall \sigma\, [\eta(\sigma) \geq \eta(\sigma 0) + \eta(\sigma 1)].$$

There exists an optimal enumerable semimeasure $\overline{M}$ that dominates (up to a multiplicative constant) any other enumerable semimeasure (Levin).

The a priori complexity of a string σ is defined as $-\log \overline{M}(\sigma)$.

Given a computable order h, we say a real $x \in 2^{\mathbb{N}}$ is strongly h-complex ($-\log \overline{M}(\sigma) \leq K(\sigma)$ up to an additive constant) if

$$(\forall n)\, [-\log \overline{M}(x \upharpoonright_n) \geq h(n)],$$

A pointwise version of Frostman's Lemma

The main result

Given an order h, we say x is h-capacitable if there exists an h-bounded probability measure μ such that x is μ-random.

Effective Capacitability Theorem

Suppose $x \in 2^{\mathbb{N}}$ is strongly h-complex, where h is a computable, convex order function. Then x is h-capacitable.

Effective Dimension and Continuous Randomness

Proving the effective capacitability theorem

- By the Kucera-Gacs Theorem, there exists a λ-random real y such that $y \geq_{wtt} x$ via some reduction Φ.
- For every $\sigma \in 2^{<\mathbb{N}}$ we define

$$\mathrm{Pre}(\sigma) = \{\tau : \ \Phi(\tau) \supseteq \sigma \,\&\, \forall \tau' \subset \tau (\Phi(\tau') \not\supseteq \sigma)\}.$$

  $\lambda(\mathrm{Pre}(.))$ is an enumerable semimeasure.
- It follows that $\lambda(\mathrm{Pre}(.))$ is multiplicatively dominated by $\overline{M}$.
- Since x is strongly $\mathcal{H}^h$-complex, there exists a constants c' and c'' such that for all n,

$$\lambda(\mathrm{Pre}(x\restriction_n)) \leq c\overline{M}(x\restriction_n) \leq c''2^{-h(n)}$$

- x is an infinite path through the co-r.e. tree

$$T = \{\sigma \in 2^{<\mathbb{N}} : \text{for all } n \leq |\sigma|, \, \lambda(\mathrm{Pre}(\sigma\restriction_n)) \leq c'' 2^{-h(n)}\}.$$

Effective Dimension and Continuous Randomness

Proving the effective capacitability theorem

- We want to define $\mu(\sigma)$, $\sigma \in 2^{<\mathbb{N}}$. We have to satisfy two requirements:
  - Preserve randomness of R when transforming with Φ. Require that

    $$(\forall \sigma \in T)\ [\lambda(\mathrm{Pre}(\sigma)) \leq \mu(\sigma)].$$

    This way, a possible μ-test covering x would "lift" to a λ-test covering R.
  - Observe the h-bound:

    $$\mu(\sigma) \leq \gamma 2^{-h(|\sigma|)},$$

    for some constant γ.
- This singles out suitable completions of the semimeasure induced by Φ.

Effective Dimension and Continuous Randomness

Proving the effective capacitability theorem

- It can be shown that there exists a non-empty Π^0_1-class of suitable completions.
  - For this, the set of probability measures on $2^{\mathbb{N}}$ has to be topologized in an effective way.
- Note that if (V_n) were a μ-test covering x, then $\Phi^{-1}(V_n)$ would be a λ-test relative to μ covering y.
  - So, what we need to show is that y is λ-random relative to μ for some $\mu \in M$.

Theorem

If $B \subseteq 2^{\mathbb{N}}$ is nonempty and Π^0_1, then, for every y which is λ-random there is $z \in B$ such that y is λ-random relative to z.

(Downey, Hirschfeldt, Miller, and Nies; Reimann and Slaman)

Applications of Effective Capacitability

- A new proof of Frostman's Lemma.
- A new characterization of effective dimension.
- Comparison of randomness notions.

Applications of Effective Capacitability

A new proof of Frostman's Lemma

We obtain a new proof of Frostman's Lemma for the base case of closed sets.

- Let $A \subseteq 2^{\mathbb{N}}$ be closed with $\mathcal{H}^s(A) > 0$.
- A is $\Pi^0_1(z)$ relative to some $z \in 2^{\mathbb{N}}$.
- Since $\mathcal{H}^s(A) > 0$, there exists an $x \in A$ that is strongly $\mathcal{H}^s$-complex relative to $s \oplus z$.
- A relativized version of the effective capacitability theorem yields the existence of a μ such that x is μ-random relative to $s \oplus z$ and μ is s-bounded with constant γ.
- A is $\Pi^0_1(z)$ and contains a μ-z-random real, it follows that $\mu(A) > 0$.
- Restrict μ to A and normalize.

Applications of Effective Capacitability

A new proof of Frostman's Lemma

The new proof is of a profoundly effective nature.

- Kucera-Gacs Theorem (does not have a classical counterpart)
- Compactness is used in the form of a basis result for Π^0_1 classes.
- The problem of assigning non-trivial measure to A is solved by making an element of A random.

Kjos-Hanssen observed that strong randomness is the precise effective level for which a pointwise Frostman Lemma holds.

Theorem

If x is not strongly $\mathcal{H}^h$-random then x is not effectively h-capacitable.

We also obtain a new characterization of effective dimension.

Theorem

For any real $x \in 2^{\mathbb{N}}$,

$$\dim_H^1 x = \sup\{s \in \mathbb{Q} : x \text{ is h-capacitable for } h(n) = sn\}.$$

Selection Rules

Von Mises (1919) – Grundlagen der Wahrscheinlichkeitsrechung

Kollektives – Probabilities from a single sequence of outcomes

(1) "The relative frequencies of the attributes must possess limiting values."

(2) "... these limiting values must remain the same in all partial sequences which may be selected from the original one in an arbitrary way... The only essential condition is that the question whether or not a certain member of the original sequence belongs to the selected partial sequence should be settled independently of the result of the corresponding observation."

Selection Rules

Von Mises revisited

Admissible selection rules

How should the notion of a selection rule be formalized? What does "independently of" mean?

- Admissible: Select all even/odd/prime/... positions.
- Not admissible: Given a sequence $01101010 0\dots$, select all positions where 0 occurs.

Two alternatives

(1) Fix the Kollektiv. Then try to find out what the admissible selection rules are.

(2) Fix the admissible selection rules. Then investigate the Kollektiv obtained.

Selection Rules

The Kollektive of normal numbers

Normal numbers

In a normal sequence every finite binary string σ occurs with limiting frequency $2^{-|\sigma|}$.

Normal numbers as Kollektives – the modern view

Let $T : 2^{\mathbb{N}} \longrightarrow 2^{\mathbb{N}}$ be the shift map, and given $x \in 2^{\mathbb{N}}$, let δ_x be the Dirac measure residing on x. Then, if x is normal, any limit point (in the weak topology) of the measures

$$\mu_n = \frac{1}{n} \sum_{i=0}^{n-1} \delta_{T^i(x)}$$

is the uniform $(1/2, 1/2)$-Bernoulli measure.

Selection Rules

Types

Two types of selection rules

- **Oblivious selection rule:** sequence $S \in 2^{\mathbb{N}}$.
  Subsequence $y = x/_S$ obtained: all the bits $x(i)$ with $S(i) = 1$.
- (General) **Selection rule:** set $L \subseteq 2^{<\mathbb{N}}$.
  Subsequence $y = x/_L$ obtained: the bits $x(i)$ such that the prefix $x(0) \ldots x(i-1)$ is in L.

Question

Which general selection rules preserve normality?

Fundamental result by Agafonoff (1968), Schnorr and Stimm (1972), and Kamae and Weiss (1975).

Theorem

If L is recognized by a finite automata, then L preserves normality.

Selection Rules

Normality and automata

Kamae and Weiss asked if normality is preserved by larger classes of languages, too (e.g. context-free languages).

By generalizing Champernowne's construction Merkle and R. (2006) gave two counterexamples:

Theorem

There exist

- a normal sequence not preserved by a deterministic one-counter language (accepted by a deterministic pushdown automata with unary stack alphabet);
- a normal sequence not preserved by a linear language (slightly more complicated).

Selection Rules

Oblivious selection rules – the role of entropy

For oblivious selection rules, Kamae (1973) gave a complete characterization in terms of entropy of measures generated by sequences under shift map.

Invariant measures for the shift map

If T denotes the shift map on $2^{\mathbb{N}}$ and $x \in 2^{\mathbb{N}}$, then any limit point of the measures

$$\mu_n^x = \frac{1}{n} \sum_{i=0}^{n-1} \delta_{T^i(x)}$$

is shift invariant.

Selection Rules

Kamae's Theorem

To any shift-invariant measure μ is assigned an entropy $h(\mu)$.

Kamae-entropy

For $x \in 2^{\mathbb{N}}$, define $h(x) = \sup\{h(\mu) \colon \mu \text{ is a limit point of } \{\mu_n^x\}\}$.

Theorem

If $S \in 2^{\mathbb{N}}$ *has positive lower density, i.e.* $\liminf_n 1/n \sum_k S(k) > 0$, *then the following are equivalent.*

(i) S *preserves normality*

(ii) $h(S) = 0$

The proof uses Furstenberg's notion of disjointness: Every process of Kamae entropy 0 is disjoint from a process of completely positive entropy.

Lowness for randomness

Van Lambalgen (1987) studied reals that preserve Martin-Loef randomness in the following sense:

- If x is μ-random, then it is also μ-random relative to z.

(The sequence z provides no useful information to prove any μ-random real non-random.) Call such reals low for mu-random.

In the following we restrict ourselves to Lebesgue measure λ.

Question

Are there non-computable reals that are low for random?

Kucera and Terwijn (1999) showed that such reals exist. They constructed a simple r.e. set that is low for random.

- The construction was the first example of a cost function construction.

Questions

- What is the recursion theoretic nature of such reals?
- Is there a connection to entropy as in Kamae's result?

Schnorr's Theorem (1973)

A real x is Martin-Löf random if and only if

$$\exists c \, \forall n \; K(x \upharpoonright_n) \geq n - c.$$

Pointwise Shannon-McMillan-Breiman Theorem (Levin,Brudno)

If μ is a computable Bernoulli measure, then for any μ-random x

$$\lim_{n \to \infty} \frac{K(x \upharpoonright_n)}{n} = h(\mu) = -[p \log p + (1-p) \log(1-p)].$$

Algorithmic Entropy

Reals of low information content

K-triviality

- Chaitin (1976) considered trivial reals:

  $$\exists c\, \forall n\, C(A\restriction_n) \leq C(n) + c$$

  He showed that a real is C-trivial if and only if it is recursive.
- Solovay (1975) constructed non-recursive K-trivial reals. Chaitin showed that all K-trivial reals are Δ^0_2.

Low for K

Muchnik (1999) introduced reals that are low for K:

$$\exists c\, \forall \sigma\, C^x(\sigma) \geq C(\sigma) - c$$

Work mainly by Nies (2005) showed that all notions coincide.

Theorem

A real x *is low for random iff it is low for* K *iff it is* K*-trivial.*

K-triviality hence provides a robust notion of low information content.

Computational properties

The K trivial reals form a Σ^0_3 ideal in the Turing degrees.