Fractal Dimensions in Recursion Theory

Jan Reimann

April 21, 2006

Overview

- Von Mises tried to base probability on individual objects. Probabilities could be assigned by studying a single instance in a Collective (Kollektiv). ("*First the collective, then the probability.*")
- Von Mises ideas ("admissible selection rules") led to the theory of stochasticity. (Wald, Church, Loveland, ...)
- The modern theory of probability follows Kolmogorov's approach: measure theoretic, random variables instead of individual random objects.

- Martin-Löf proposed a definition of randomness by combining measure theory and recursion theory (effective nullsets).
- This approach is supported by the fact that it coincides with the definition of randomness as "incompressibility" (Kolmogorov complexity).

Measures on Cantor Space

Outer measures from premeasures

- A premeasure is a function $\rho : 2^{<\omega} \to \mathbb{R}_0^+ \cup \{\infty\}$.
- One can obtain an outer measure μ_ρ from ρ by letting

$$\mu_\rho(X) = \inf_{C \subseteq 2^{<\omega}} \left\{ \sum_{\sigma \in C} \rho(\sigma) : \bigcup_{\sigma \in C} N_\sigma \supseteq X \right\},$$

  where N_σ is the basic open cylinder induced by σ.
  (Set $\mu_\rho(\emptyset) = 0$.)
- $\mu = \mu_\rho$ is a countably subadditive, monotone set function.

- The way we constructed outer measures, $\mu(A) = 0$ is equivalent to the existence of a sequence $(C_n)_{n\in\omega}$, $C_n \subseteq 2^{<\omega}$, such that for all n,

$$A \subseteq \bigcup_{C_n} N_\sigma \quad \text{and} \quad \sum_{C_n} \rho(\sigma) \leqslant 2^{-n}.$$

- Thus, every nullset is contained in a G_δ nullset.

— By requiring that the covering nullset is effectively G_δ (in a presentation of ρ), we obtain a notion of effective nullsets.

Definition Let μ ($= \mu_\rho$) be an outer measure. A set A is effectively μ-null if there exists a function f recursive in ρ such that for all n,

$$A \subseteq \bigcup_{W_{f(n)}} N_\sigma \quad \text{and} \quad \sum_{W_{f(n)}} \rho(\sigma) \leqslant 2^{-n}.$$

Definition A real $X \in 2^\omega$ is μ-random iff $\{X\}$ is not μ-null.

- One can obtain stronger (or weaker) versions of randomness by relaxing the effectiveness condition:
- Stronger:
  - f recursive in $\emptyset^{(n)}$,
  - f arithmetical,
  - replace uniformly r.e. by Π^1_1 [Hjorth and Nies],
- Weaker:
  - replace uniformly r.e. by unif. recursive [Schnorr],
  - instead with covers work with martingales and impose subrecursive rescource bounds [Lutz].

— There seem to be two directions of study:

— From reals to measures:
Given a real (or a set of reals), study the measures with respect to which this is random, and for which level of randomness. [Reimann and Slaman]

— From measures to reals:
Given a measure (usually the uniform distribution $\rho(\sigma) = 2^{-|\sigma|}$), study the corresponding random reals. Reals random with respect to the uniform distribution are usually called Martin-Löf random. (Algorithmic randomness, a lot of progress over the last decade.)

This talk Hausdorff measures

Hausdorff Measures

— (Generalized) Hausdorff measures $\mathcal{H}^h$ correspond to premeasures of the type

$$\rho(\sigma) = h(|\sigma|),$$

where h is a decreasing function with $\lim_n h(n) = 0$.

— Note: ρ depends only on the length of σ, that is, the diameter of the accordant open set.

— Usually: $h(n) = 2^{-ns}$, s a nonnegative real number. In this case, we simply write $\mathcal{H}^s$ and call it the s-dimensional Hausdorff measure.

— $\mathcal{H}^1$ corresponds to the uniform distribution, i.e. Lebesgue measure λ on 2^ω.

Hausdorff Measures

Short remark

- The actual definition of the Hausdorff measure $\mathcal{H}^h$ is a little more involved. (One wants to ensure that for the resulting measures, all Borel sets are measurable.)
- We are primarily concerned with nullsets. For nullsets the more involved definition coincides with the one given here.

Hausdorff Dimension

— It is not hard to see that if $s < t$, then

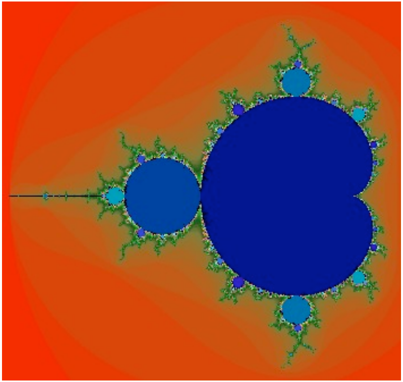
$$\mathcal{H}^s(A) = 0 \;\Rightarrow\; \mathcal{H}^t(A) = 0.$$

— Hausdorff dimension "picks" the "right" scaling factor for a set.

Definition The Hausdorff dimension of A is defined as

$$\dim_{\mathrm{H}} A = \inf\{s \geqslant 0 : \mathcal{H}^s(A) = 0\}.$$

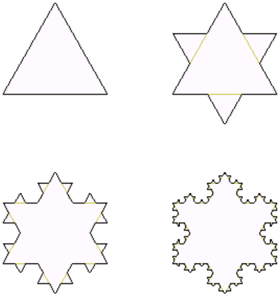
— Mandelbrot set – $\dim_H = 2$



Hausdorff Dimension

Famous examples

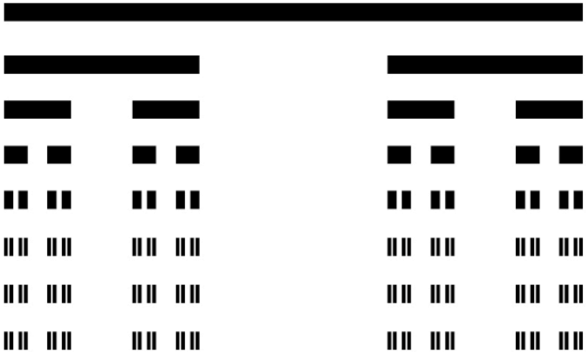
- Koch snowflake – $\dim_H = \log 4 / \log 3$



Hausdorff Dimension

Famous examples

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



— Frequency sets – For $0 \leqslant p \leqslant 1$, let

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

Then $\dim_{\mathrm{H}} A_p = H(p) = -[p \log p + (1-p)\log(1-p)]$ [Eggleston].

Properties of Hausdorff Dimension

- Lebesgue measure: $\lambda(A) > 0$ implies $\dim_H(A) = 1$.
- Monotony: $A \subseteq B$ implies $\dim_H(A) \leqslant \dim_H(B)$.
- Stability: For $A_1, A_2, \dots \subseteq 2^\omega$ it holds that

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

(Immediately implies that all countable sets have dimension 0.)

Properties of Hausdorff Dimension

Geometric properties

— Geometric transformations: If h is Hölder continuous, i.e. if there are constants $c, r > 0$ for which

$$(\forall x, y)\ d(h(x), h(y)) \leqslant c d(x, y)^r,$$

then

$$\dim_{\mathrm{H}} h(A) \leqslant (1/r) \dim_{\mathrm{H}}(A).$$

— For $r = 1$, h is Lipschitz continuous. If h is bi-Lipschitz, then

$$\dim_{\mathrm{H}} h(A) = \dim_{\mathrm{H}}(A).$$

— Fractal geometry $\hat{=}$ study of properties invariant under bi-Lipschitz transformations.

Effective Hausdorff Dimension

Definition The effective Hausdorff dimension of $A \subseteq 2^\omega$ is defined as

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

[Lutz 2000]

— There are single reals of non-zero dimension: every λ-random real has dimension one.

— Effective dimension has an important stability property:

$$\dim_{\mathrm{H}}^1 A = \sup\{\dim_{\mathrm{H}}^1\{X\} : \ X \in A\}.$$

[Lutz 2000]

Effective Dimension and Algorithmic Entropy

- Kolmogorov complexity: U a universal Turing-machine. Define

$$C(\sigma) = C_U(\sigma) = \min\{|p| : p \in 2^{<\omega}, U(p) = \sigma\},$$

i.e. $C(\sigma)$ is the length of the shortest program (for U) that outputs σ.

- Kolmogorov's invariance theorem: C is optimal (up to an additive constant).

- A prefix-free Turing machine is a TM with prefix-free domain. The prefix-free version of C (use universal prefix free TM) is denoted by K.

— A fundamental theorem of algorithmic randomness establishes that randomness is incompressibility:

$$\alpha \ \lambda\text{-random} \quad \Leftrightarrow \quad (\exists c)\,(\forall n)\, K(\alpha\restriction_n) \geqslant n - c.$$

[Schnorr 1971]

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

Theorem For every real X,

$$\dim^1_{\mathrm{H}} X = \liminf_{n\to\infty} \frac{\mathrm{K}(X\restriction_n)}{n}.$$

[Ryabko 1984; Mayordomo 2002]

— 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 $\dim^1_{\mathrm{H}} X_r = r$.

— Geometry: Hölder transformation of Cantor set

Information theory: Insert redundancy

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

$$\dim_{\mathrm{H}}^{1} B = H(\mu_p) := -[p \log p + p \log(1-p)].$$

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

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

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

has effective dimension s [Tadaki 2002].

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

- 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, Levin and Kautz have shown that any real which is random with respect to a recursive measure computes a Martin-Löf random real.
- $\Omega^{(s)}$ 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

Are there "genuine" reals of non-integral dimension?

— 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?

— This is an open problem. 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.

Upper Cones

Upper cones always have maximal dimension

- For upper cones, the situation is quite clear.
- The Turing upper cone of a real has Lebesgue measure zero unless the real is recursive [Sacks 1963].

Theorem For any real X, the many-one upper cone of X has (classical) Hausdorff dimension 1.

Lower Cones and Degrees

- The dimension of a lower cone and a degree coincide.
- This follows from the sparse coding technique: Given two reals $A \leqslant_r B$, choose a recursive set R of density $\lim_n |R \cap \{0, \ldots, n-1\}|/n = 1$, and let C equal A on R and B on the complement of R.
- C will be r-equivalent to B and be of the same dimension as A. It follows that the dimension of the degree and the lower cone of a set coincide.

Symmetry of Information

— An important tool: Symmetry of algorithmic information.

$$\mathsf{K}(\langle x, y\rangle) \stackrel{+}{=} \mathsf{K}(x) + \mathsf{K}(y|x, \mathsf{K}(x))$$

Many-One Reducibility

Theorem Let μ_p be a computable Bernoulli measure with bias p. If A is μ_p-random, then

$$B \leqslant_{\mathrm{m}} A \;\Rightarrow\; \dim^1_{\mathrm{H}} B \leqslant H(\mu_p).$$

[Reimann and Terwijn 2004]

— Proof. Given an m-reduction f, define $F = \{n : (\forall m < n) f(m) \neq f(n)\}$, so F is the set of all positions of B, where an instance of A is queried for the first time.

— F induces a Kolmogorov-Loveland place selection rule. If A is μ_p-random, this selection rule will yield a new sequence with the same limit frequency as A.

Weaker Reducibilities

- This technique does not extend to weaker reducibilities, since for Bernoulli measures the Levin-Kautz result holds for a total Turing reduction.
- Stephan [2005] was able to construct wtt-lower cone of non-integral effective dimension in a relativized world:

  There is a real A and an oracle B such that

$$1/3 \leqslant \dim_{\mathrm{H}}^{B}\{D : D \leqslant_{\mathrm{wtt}}^{B} A\}) \leqslant 1/2.$$

Theorem For each rational α, $0 \leqslant \alpha \leqslant 1$, there is a real $X \leqslant_{\mathrm{wtt}} \emptyset'$ such that

$$\dim^1_{\mathrm{H}} X = \alpha \quad \text{and} \quad (\forall Z \leqslant_{\mathrm{wtt}} X)\ \dim^1_{\mathrm{H}} Z \leqslant \alpha.$$

[Nies and Reimann 2006]

- Requirements:

$$R_{\langle e,j\rangle} : Z = \Psi_e(X) \;\Rightarrow\; \exists(k \geqslant j)\, K(Z\restriction_k) \leqslant^+ (\alpha + 2^{-j})k$$

  where (Ψ_e) is a uniform listing of wtt reduction procedures.

- We can assume each Ψ_e also has a certain (non-trivial) lower bound on the use g_e, because otherwise the reduction would decrease complexity anyway.

- We construct X inside the Π^0_1 class

$$P = \{Y : (\forall n \geqslant n_0)\, K(Y\restriction_n) \geqslant \lfloor \alpha n \rfloor\}$$

(This ensures X has dimension at least α.)

- P is given as an effective approximation through clopen sets P_s.

- We approximate longer and longer initial segments σ_j of X, where σ_j is a string of length m_j, both σ_j, m_j controlled by R_j.

- Define a length k_j where we intend to compress Z, and let $m_j = g_e(k_j)$.
- Define σ_j of length m_j in a way that, if $x = \Psi_e^{\sigma_j}$ is defined then we compress it down to $(\alpha + 2^{-b_j})k_j$, by constructing an appropriate nullset L.
- The opponent's answer could be to remove σ_j from P. (σ_j is not of high dimension.)
- In this case, the capital he spent for this removal exceeds what we spent for our request, so we can account our capital against his.
- Of course, usually σ_j is much longer than x. So we will only compress x when the measure of oracle strings computing it is large.

- In the course of the construction, some R_j might have to pick a new σ_j.
- In this case we have to initialize all R_n of lower priority ($n > j$).
- We have to make sure that this does not make us enumerate too much measure into L.
- We therefore have to assign a new length k_n to the strategies R_n.
- For this, it is important to know the use of the reduction related to R_j.

The Turing Case

— The Turing case appears to be much harder. Currently, the best known result is the following.

Theorem There exists recursive, non-decreasing, unbounded function h and a real X such that for all n,

$$K(X\restriction_n) \geqslant h(n) \qquad (*)$$

and X does not compute a Martin-Löf random set.

[Kjos-Hanssen, Merkle, and Stephan 2004; Reimann and Slaman 2004]

— The condition $(*)$ can be interpreted in terms of (generalized) Hausdorff measures. Reals satisfying $(*)$ are called complex.

Complex Reals and DNR Functions

- The proof by Kjos-Hanssen, Merkle, and Stephan reveals an interesting connection between entropy and diagonally nonrecursive functions.
- A function f is diagonally nonrecursive (dnr) if for all n, $f(n) \neq \varphi_n(n)$.
- Kjos-Hanssen, Merkle, and Stephan showed that a real is complex iff it truth-table computes a dnr function.
- Ambos-Spies, Kjos-Hanssen, Lempp, and Slaman [2004] showed that there exists a dnr function that does not compute a dnr function whose values are bounded by a recursive function.
- It is known that every Martin-Löf random real computes a 0-1 valued dnr function.

Further Rescources

- For papers, preprints, and my thesis on effective dimension:

  http://math.uni-heidelberg.de/logic/reimann