Mutual Theories of Algorithmic Information

Jan Reimann

February 3, 2008

Selection Rules

Von Mises revisited

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 011010100 . . . , 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 $T : 2^{\mathbb{N}} \to 2^{\mathbb{N}}$ 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?

Selection Rules

Normality and finite automata

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.

Martin-Löf Randomness

From stochasticity to randomness

The second approach to Kollektives, fixing a set of admissible selection rules first, drew criticism in the 1930's, culminating in Ville's construction.

- Ville (1939) was able to show that for any countable set of (monotone) selections rules, the resulting Kollektiv contains a sequences violating the law of the iterated logarithm.

In spite of the widespread acceptance of the axiomatic foundation of probability given by Kolmogorov (1933), Church suggested the use of computable selection rules. This, and variants thereof (Kolmogorov, Loveland), is now known as the theory of stochastic sequences.

Martin-Löf Randomness

From stochasticity to randomness

Martin-Löf (1966) defined random sequences as sequences not belonging to any effective nullset, which turned out to be a very robust definition.

Definition

Let μ be an (outer) measure on $2^{\mathbb{N}}$.

- A μ-test relative to $z \in 2^{\mathbb{N}}$ is a set $W \subseteq \mathbb{N} \times 2^{<\mathbb{N}}$ which is r.e. (Σ^0_1) in z such that

$$\sum_{\sigma \in W_n} \mu(N_\sigma) \leq 2^{-n},$$

  where $W_n = \{\sigma : (n, \sigma) \in W\}$.
- A real x passes a test W if $x \notin \bigcap_n N(W_n)$.
- A real x is μ-random relative to z if it passes all μ-tests relative to $z \oplus \pi_\mu$, where π_μ is a suitable representation of μ.

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 $\mathcal{L}$.

Question

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

Martin-Löf Randomness

Lowness for randomness

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?

Kolmogorov complexity

U universal Turing-machine. Define

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

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.

Algorithmic Entropy

Entropy and randomness

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

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

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:

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

Algorithmic Entropy

Lowness for randomness = K-trivial

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.

Algorithmic Entropy

K-trivial = low information content

There is another characterization in terms of mutual information.

Mutual information for finite strings - Kolmogorov, Levin

$$I(\sigma : \tau) = K(\sigma) + K(\tau) - K(\sigma, \tau).$$

This can be extended to infinite sequences in various ways, e.g. Levin (1984).

Then we can characterize K-triviality as having no information about other sequences (Hirschfeldt and R.).

Theorem
A real y is K-trivial if and only if for all x, $I(y, x) < \infty$.

Positive Entropy

Can we extract information/randomness?

Sinai's Theorem

Any ergodic system with positive entropy has a Bernoulli factor.

Say a real x has positive entropy if $\liminf_n K(x\upharpoonright_n)/n > 0$.

Question

Given a real of positive entropy, can we effectively compute a Martin-Löf random real from it?

- Note that it would be enough to compute a μ-random real, where μ is a Bernoulli measure, since every Bernoulli random real can be effectively transformed into a $\mathcal{L}$-random real (e.g. using von Neumann's trick).

Positive Entropy

The standard examples

Three examples – all compute a Martin-Löf random real

- Let $0 < r < 1$ rational. Given a Martin-Löf random real x, define x_r by inserting 0 with density 0. Then $\liminf_n K(x_r \upharpoonright_n)/n = r$.
- If μ_p is a Bernoulli measure with bias $p \in \mathbb{Q} \cap [0,1]$, then for any μ_p-random real x, $\liminf_n K(x_r \upharpoonright_n)/n = h(\mu_p)$.
- Let U be a universal, prefix-free machine. Given a computable real number $0 < s \leq 1$, the binary expansion of the real number

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

  satisfies $\liminf_n K(x_r \upharpoonright_n)/n = s$ (Tadaki, 2002)

Positive Entropy

Hausdorff measures and dimension

Let $h : \mathbb{N} \to \mathbb{N}$ be recursive, non-decreasing, unbounded.

Theorem

For any $x \in 2^{\mathbb{N}}$, the following are equivalent.

(i) *For all n, $K(x\restriction_n) \geq h(n)$.*

(ii) *x is $\mathcal{H}^h$-random.*

This yields that a real has positive entropy if and only if it has positive effective Hausdorff dimension, in fact (Ryabko, Mayordomo)

$$\dim^1_{\mathrm{H}} x = \liminf_n \frac{K(x\restriction_n)}{n}.$$

Relation to Kamae entropy (Brudno, 1984)

$$\limsup_n \frac{K(x\restriction_n)}{n} \leq h(x).$$

Positive Entropy

Entropy implies randomness

Using techniques from geometric measure theory (capacities, Frostman's Lemma) along with methods from recursion theory (basis results for Π^0_1 sets), one can show that positive entropy implies randomness.

Theorem

If $\dim^1_H x > s > 0$, then there exists a probability measure μ such that x is μ-random and for all σ,

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

However, this yields no progress on the extractability problem.

Positive Entropy

Strong reducibilities

Many-one reducibility

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

$$y \leq_{\mathrm{m}} x \;\Rightarrow\; \dim_{\mathrm{H}}^1 y \leq h(\mu_p).$$

R. and Terwijn (2004)

Weak truth-table reducibility

For each rational α, $0 \leq \alpha \leq 1$, there is a real $x \leq_{\mathrm{wtt}} \emptyset'$ such that

$$\dim_{\mathrm{H}}^1 x = \alpha \quad \text{and} \quad (\forall y \leq_{\mathrm{wtt}} x)\ \dim_{\mathrm{H}}^1 y \leq \alpha.$$

Nies and R. (2006)

Definition

A function f is diagonally nonrecursive (dnr) if for all n, $f(n) \neq \varphi_n(n)$.

Kjos-Hanssen, Merkle, and Stephan (2006) revealed an interesting connection between entropy and dnr functions.

Theorem

- *A real x tt-computes a dnr function iff x is $\mathcal{H}^h$-random for a recursive, non-decreasing, unbounded h.*
- *A real x T-computes a dnr function iff x is $\mathcal{H}^h$-random for some $h \leq_T x$.*

Positive Entropy

A minimal degree of positive entropy

Theorem

There exists a minimal dnr degree. (Kumabe)

Kumabe's construction uses bushy trees in Baire space.

Recently, Greenberg and Miller were combine such trees with the Kjos-Hanssen-Merkle-Stephan correspondance to prove a non-extractability result.

Theorem

There exists a minimal Turing degree that contains a real of entropy 1*.*

- Note that no Martin-Löf random real can have this property, since every recursive splitting yields two relatively random (and hence T-incomparable) halves.

Positive Entropy