Kolmogorov-Loveland Randomness and Stochasticity

Wolfgang Merkle[1] Joseph Miller[2] André Nies[3]
Jan Reimann[1] Frank Stephan[4]

[1]Institut für Informatik, Universität Heidelberg
[2]Department of Mathematics, Indiana University
[3]Department of Computer Science, University of Auckland
[4]Department of Mathematics, National University of Singapore

February 25, 2005

Overview

Overview

Overview

Overview

Overview

Approaches to Randomness

Question

How can we define a random binary sequence?

Possible approaches:

- Typicalness – a “large” set should contain a random element. [Martin-Löf; Schnorr]
- Incompressibility – a random sequence should be incompressible (by algorithmic means). [Kolmogorov; Levin; Chaitin]
- Unpredictability – no computable betting strategy should win against a random sequence. [Schnorr]
- Stochasticity – limit frequencies are preserved under selecting subsequences. [Von Mises]

Approaches to Randomness

- It turns out that randomness via incompressibility is algorithmically very strict.
- Incompressible sequences correspond to sequences against which no enumerable betting strategy wins.
- Schnorr's criticism: Notions of randomness should be based on computable objects, e.g. computable betting games.
- This talk: Is it possible to do this and still obtain a randomness concept as powerful as incompressibility?
- Key ingredient: non-monotonicity.

Betting games

Given: unknown infinite binary sequence A

A round in the game

- Start with a capital of 1.
- Select a position $k \in \mathbb{N}$ and specify a stake $v \in [0, 1]$.
- Predict the bit $A(k)$.
- If the prediction is correct, the capital is multiplied by $1 + v$. Otherwise the stake is lost.
- Continue: pick a new position not selected before.

A	?	?	?	?	?	?	?	?	?	?	?	?	?	?	?	?	?	?	?	?

Betting games

Given: unknown infinite binary sequence A

A round in the game

- Start with a capital of 1.
- Select a position $k \in \mathbb{N}$ and specify a stake $v \in [0, 1]$.
- Predict the bit $A(k)$.
- If the prediction is correct, the capital is multiplied by $1 + v$. Otherwise the stake is lost.
- Continue: pick a new position not selected before.

A | ? | ? | ? | ? | ? | ? | ? | ? | ? | ? | ? | ? | ? | ? | ? | ? | ? | ? | ? | ? | …

Capital: 1

Betting games

Given: unknown infinite binary sequence A

A round in the game

- Start with a capital of 1.
- Select a position $k \in \mathbb{N}$ and specify a stake $v \in [0, 1]$.
- Predict the bit $A(k)$.
- If the prediction is correct, the capital is multiplied by $1 + v$. Otherwise the stake is lost.
- Continue: pick a new position not selected before.

A																				
	?	?	?	?	?	?	?	?	?	?	?	?	?	?	?	?	?	?	?	?

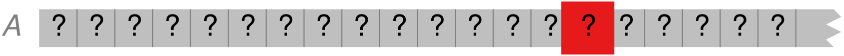
Capital: 1 Stake: 0.5

Betting games

Given: unknown infinite binary sequence A

A round in the game

- Start with a capital of 1.
- Select a position $k \in \mathbb{N}$ and specify a stake $v \in [0, 1]$.
- Predict the bit $A(k)$.
- If the prediction is correct, the capital is multiplied by $1 + v$. Otherwise the stake is lost.
- Continue: pick a new position not selected before.



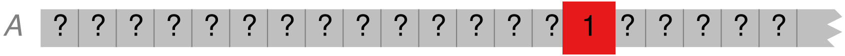
Capital: 1 Stake: 0.5 Prediction: 1

Betting games

Given: unknown infinite binary sequence A

A round in the game

- Start with a capital of 1.
- Select a position $k \in \mathbb{N}$ and specify a stake $v \in [0, 1]$.
- Predict the bit $A(k)$.
- If the prediction is correct, the capital is multiplied by $1 + v$. Otherwise the stake is lost.
- Continue: pick a new position not selected before.



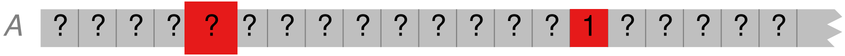
Capital: 1.5 Stake: 0.5 Prediction: 1

Betting games

Given: unknown infinite binary sequence A

A round in the game

- Start with a capital of 1.
- Select a position $k \in \mathbb{N}$ and specify a stake $v \in [0, 1]$.
- Predict the bit $A(k)$.
- If the prediction is correct, the capital is multiplied by $1 + v$. Otherwise the stake is lost.
- Continue: pick a new position not selected before.



Capital: 1.5 Stake: 0.5 Prediction: 1

Betting strategies

Definition

- A betting strategy is a partial function that, given the outcomes of the previous rounds, determines
  - the position on which to bet next,
  - the stake to bet,
  - the value predicted.
- Formally, a betting strategy is a partial mapping

$$b : (\mathbb{N} \times \{0, 1\})^* \to \mathbb{N} \times [0, 1] \times \{0, 1\}.$$

- A betting strategy is monotone if the positions to bet on are chosen in an increasing order.
- A betting strategy b succeeds on sequence S if the capital grows unbounded when playing against S according to b.

Randomness as Unpredictability

Definition

- A sequence is Kolmogorov-Loveland random (KL-random) if no (partial) computable betting strategy succeeds on it.
- A sequence is computably random if no computable monotone betting strategy succeeds on it.

Randomness as Typicalness

Definition

- A Martin-Löf test (ML-test) is a uniformly computable sequence $(V_n)_{n\in\mathbb{N}}$ of c.e. sets of strings such that for all n,

$$\sum_{\sigma\in V_n} 2^{-|\sigma|} \leq 2^{-n}.$$

- An ML-test (V_n) covers a sequence A if $(\forall n)(\exists\sigma \in V_n)\ \sigma \sqsubset A$.
- A sequence is Martin-Löf random (ML-random) if it is not covered by an ML-test.
- A Schnorr test is an ML-test (V_n) such that the real number $\sum_{\sigma\in V_n} 2^{-|\sigma|}$ is uniformly computable. A sequence is Schnorr random if it not covered by a Schnorr test.

Relations between Randomness Notions

Fact

ML-random $\subseteq$ *KL-random* $\subsetneq$ *computably random* $\subsetneq$ *Schnorr random*

Open Question (Muchnik, Semenov, and Uspensky, 1998)

Is KL-randomness equivalent to ML-randomness?

Resource-Bounded Betting Games

Buhrman, van Melkebeek, Regan, Sivakumar, and Strauss (2000) studied resource-bounded betting strategies.

Some Results

- If pseudorandom generators computable in exponential time (E, EXP) with exponential security exist, then every betting strategy computable in exponential time can be simulated by an exponential time monotone betting strategy.
- If exponential time betting strategies have the finite union property, then BPP $\neq$ EXP.

Resource-Bounded Betting Games

Buhrman, van Melkebeek, Regan, Sivakumar, and Strauss (2000) studied resource-bounded betting strategies.

Some Results

- If pseudorandom generators computable in exponential time (E, EXP) with exponential security exist, then every betting strategy computable in exponential time can be simulated by an exponential time monotone betting strategy.
- If exponential time betting strategies have the finite union property, then BPP $\neq$ EXP.

Example: Computably Enumerable Sets

Example

No partial computable non-monotonic betting strategy can succeed on all computably enumerable sets.

Example: Computably Enumerable Sets

Example

No partial computable non-monotonic betting strategy can succeed on all computably enumerable sets.

Given a computable betting strategy b, define a c.e. set W such that b does not succeed on W by enumerating elements into W according to the places selected by b.

Example: Computably Enumerable Sets

Example

No partial computable non-monotonic betting strategy can succeed on all computably enumerable sets.

Fact

There exist computable non-monotonic betting strategies b_0 and b_1 such that for every c.e. set W, at least one of b_0 and b_1 will succeed on W.

b_0 succeeds on all rather sparse sets, whereas b_1 succeeds if a lot of elements are enumerated into the set W.

Failure of finite union property

Computable betting strategies do not have the finite union property.

Kolmogorov Complexity

Definition

Let U be a universal Turing machine. For a string σ define the Kolmogorov complexity C of a string as

$$\mathrm{C}(\sigma) = \mathrm{C}_U(\sigma) = \min\{|p| : \ p \in \{0,1\}^*, \ U(p) = \sigma\},$$

i.e. $\mathrm{C}(\sigma)$ is the length of the shortest U-program for σ.

Fact (Kolmogorov; Solomonoff)

C *is independent of the choice of U, up to an additive constant.*

Variant

Prefix-free complexity K. Based on prefix-free Turing-machines – no two converging inputs are prefixes of one another.

The Complexity of Martin-Löf Random Sequences

Theorem (Schnorr)

Given a sequence A, if there exists a function

$$h : \mathbb{N} \to \mathbb{N}$$

such that for all n,

$$K(A\restriction_{h(n)}) \leq h(n) - n,$$

then A is not ML-random.

The Complexity of Martin-Löf Random Sequences

Theorem (Schnorr)

Given a sequence A, if there exists a function

$$h : \mathbb{N} \to \mathbb{N}$$

such that for all n,

$$K(A\restriction_{h(n)}) \leq h(n) - n,$$

then A is not ML-random.

The Complexity of KL-Random Sequences

Theorem (Muchnik)

Given a sequence A, if there exists a computable function

$$h : \mathbb{N} \to \mathbb{N}$$

such that for all n,

$$K(A\restriction_{h(n)}) \leq h(n) - n,$$

then A is not KL-random.

The Complexity of KL-Random Sequences

Note that this is fails for computably random sequences. In fact, there can be computably random sequences of very low complexity. [Muchnik; Merkle]

Extracting Subsequences

Let Z be an infinite, co-infinite subset of $\mathbb{N}$.

Definition

Given a sequence A, the Z-subsequence of A, $A\restriction_Z$, is defined as

$$A\restriction_Z(n) = 1 \iff A(p_Z(n)) = 1,$$

where $p_z(n)$ is the $n+1$st element of Z.

Extracting Subsequences

Let Z be an infinite, co-infinite subset of $\mathbb{N}$.

Definition

Given a sequence A, the *Z-subsequence* of A, $A\restriction_Z$, is defined as

$$A\restriction_Z(n) = 1 \;\Leftrightarrow\; A(p_Z(n)) = 1,$$

where $p_z(n)$ is the $n+1$st element of Z.

Extracting Subsequences

Let Z be an infinite, co-infinite subset of $\mathbb{N}$.

Definition

Given a sequence A, the Z-subsequence of A, $A\upharpoonright_Z$, is defined as

$$A\upharpoonright_Z(n) = 1 \;\Leftrightarrow\; A(p_Z(n)) = 1,$$

where $p_z(n)$ is the $n+1$st element of Z.

A	1	1	1	0	0	1	0	1	1	0	1	0	1	0	1	1	1	0	0	1

Extracting Subsequences

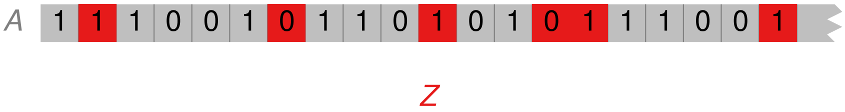
Let Z be an infinite, co-infinite subset of $\mathbb{N}$.

Definition

Given a sequence A, the Z-subsequence of A, $A\restriction_Z$, is defined as

$$A\restriction_Z(n) = 1 \;\Leftrightarrow\; A(p_Z(n)) = 1,$$

where $p_z(n)$ is the $n+1$st element of Z.



Extracting Subsequences

Let Z be an infinite, co-infinite subset of $\mathbb{N}$.

Definition

Given a sequence A, the Z-subsequence of A, $A\restriction_Z$, is defined as

$$A\restriction_Z(n) = 1 \iff A(p_Z(n)) = 1,$$

where $p_z(n)$ is the $n+1$st element of Z.

A	1	1	1	0	0	1	0	1	1	0	1	0	1	0	1	1	1	0	0	1
$A\restriction_Z$	1	0	1	0	1	1														

Generalized Joins

Definition

The *Z-join* of two sequences A_0, A_1,

$$A_0 \oplus_Z A_1,$$

is defined as the unique sequence A such that

$$A \restriction_{\overline{Z}} = A_0 \quad \text{and} \quad A \restriction_Z = A_1.$$

Generalized Joins

Definition

The *Z*-join of two sequences A_0, A_1,

$$A_0 \oplus_Z A_1,$$

is defined as the unique sequence A such that

$$A \upharpoonright_{\overline{Z}} = A_0 \quad \text{and} \quad A \upharpoonright_Z = A_1 .$$

A

Generalized Joins

Definition

The Z-join of two sequences A_0, A_1,

$$A_0 \oplus_Z A_1,$$

is defined as the unique sequence A such that

$$A\restriction_{\overline{Z}} = A_0 \quad \text{and} \quad A\restriction_Z = A_1 .$$

Generalized Joins

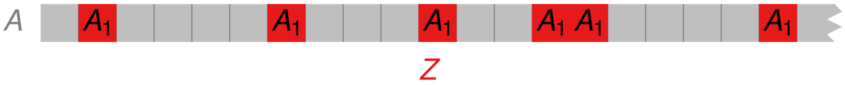
Definition

The Z-join of two sequences A_0, A_1,

$$A_0 \oplus_Z A_1,$$

is defined as the unique sequence A such that

$$A \restriction_{\overline{Z}} = A_0 \quad \text{and} \quad A \restriction_Z = A_1 .$$



Generalized Joins

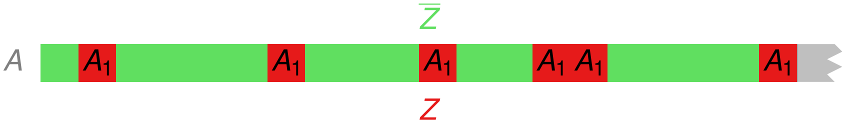
Definition

The Z-join of two sequences A_0, A_1,

$$A_0 \oplus_Z A_1,$$

is defined as the unique sequence A such that

$$A \restriction_{\overline{Z}} = A_0 \quad \text{and} \quad A \restriction_Z = A_1 .$$



Generalized Joins

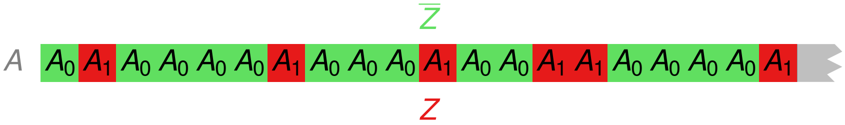
Definition

The Z-join of two sequences A_0, A_1,

$$A_0 \oplus_Z A_1,$$

is defined as the unique sequence A such that

$$A\restriction_{\overline{Z}} = A_0 \quad \text{and} \quad A\restriction_{Z} = A_1 .$$



Splitting Properties

If we split a KL-random sequence effectively, both subsequences obtained must be KL-random relative to each other.

Observation

Let Z be a computable, infinite and co-infinite set of natural numbers, and let $A = A_0 \oplus_Z A_1$. A is KL-random if and only if

$$A_0 \text{ is } \mathrm{KL}^{A_1}\text{-random} \quad \text{and} \quad A_1 \text{ is } \mathrm{KL}^{A_0}\text{-random.}$$

Splitting Properties

If we split a KL-random sequence effectively, both subsequences obtained must be KL-random relative to each other.

Observation

Let Z be a computable, infinite and co-infinite set of natural numbers, and let $A = A_0 \oplus_Z A_1$. A is KL-random if and only if

$$A_0 \text{ is } \mathrm{KL}^{A_1}\text{-random} \quad \text{and} \quad A_1 \text{ is } \mathrm{KL}^{A_0}\text{-random.}$$

Splitting Properties

Observation

Let Z be a computable, infinite and co-infinite set of natural numbers, and let $A = A_0 \oplus_Z A_1$. A is KL-random if and only if

$$A_0 \text{ is } \mathrm{KL}^{A_1}\text{-random} \quad \text{and} \quad A_1 \text{ is } \mathrm{KL}^{A_0}\text{-random.}$$

Proof of "$\Rightarrow$":

- Suppose b_1 computable in A_1 succeeds on A_0.
- Devise betting strategy successful on A:
  - Scan the Z-positions of the sequence (corresponding to the places where A_1 is coded).
  - Find a new initial segment which allows to compute a new value of b_1.
  - Bet on the $\overline{Z}$-positions of the sequence according to b_1.

Splitting Properties

We can use this observation to show that one “half” of a KL-random sequence must always be ML-random.

Theorem

Let Z be a computable, infinite and co-infinite set of natural numbers. If the sequence $A = A_0 \oplus_Z A_1$ is KL-random, then at least one of A_0, A_1 is Martin-Löf random.

Splitting Properties

We can use this observation to show that one “half” of a KL-random sequence must always be ML-random.

Theorem

Let Z be a computable, infinite and co-infinite set of natural numbers. If the sequence $A = A_0 \oplus_Z A_1$ is KL-random, then at least one of A_0, A_1 is Martin-Löf random.

Splitting Properties

Theorem

Let Z be a computable, infinite and co-infinite set of natural numbers. If the sequence $A = A_0 \oplus_Z A_1$ is KL-random, then at least one of A_0, A_1 is Martin-Löf random.

- Suppose neither A_0 nor A_1 is ML-random.
- Then there are Martin-Löf tests $(U_n^0 \colon n \in \mathbb{N})$ and $(U_n^1 \colon n \in \mathbb{N})$ with $U_n^i = \{u_{n,0}^i, u_{n,1}^i, \dots\}$, such that (U_n^i) covers A_i.
- Define functions f_0, f_1 by $f_i(n) = \min\{k \in \mathbb{N} \colon u_{n,k}^i \sqsubset A_i\}$.
- For some i, $(\overset{\infty}{\exists} m)\ f_i(m) \geq f_{1-i}(m)$.

Splitting Properties

Theorem

Let Z be a computable, infinite and co-infinite set of natural numbers. If the sequence $A = A_0 \oplus_Z A_1$ is KL-random, then at least one of A_0, A_1 is Martin-Löf random.

- Define a new test (V_n) by

$$V_n = \bigcup_{m>n} \{u_0^{1-i}, \ldots, u_{f_i(m)}^{1-i}\}.$$

- (V_n) is a Schnorr test relative to the oracle A_i and covers A_{1-i}, so A_{1-i} is not Schnorr^{A_i}-random.
- KL-randomness implies Schnorr-randomness, so A_{1-i} is not KL^{A_i}-random, and hence A is not KL-random.

Splitting Properties

This result can be improved.

Z has density δ if $\lim_{m\to\infty} |\{Z \cap \{0,\ldots,m-1\}|/m = \delta$.

Theorem

Let A be a KL-random sequence and let $\delta < 1$ be rational. Then there is a computable set Z of density at least δ such that $A\restriction_Z$ is ML-random.

Proof uses a result by Van Lambalgen (1987), who showed that $A = A_0 \oplus_Z A_1$ is ML-random if and only if A_0 is ML-random and A_1 is ML^{A_0}-random.

Splitting Properties

This result can be improved.

Z has density δ if $\lim_{m\to\infty} |\{Z \cap \{0,\dots,m-1\}|/m = \delta$.

Theorem

Let A be a KL-random sequence and let $\delta < 1$ be rational. Then there is a computable set Z of density at least δ such that $A\restriction_Z$ is ML-random.

Proof uses a result by Van Lambalgen (1987), who showed that $A = A_0 \oplus_Z A_1$ is ML-random if and only if A_0 is ML-random and A_1 is ML^{A_0}-random.

Splitting Properties

This result can be improved.

Z has density δ if $\lim_{m\to\infty} |\{Z \cap \{0,\ldots,m-1\}|/m = \delta$.

Theorem

Let A be a KL-random sequence and let $\delta < 1$ be rational. Then there is a computable set Z of density at least δ such that $A\restriction_Z$ is ML-random.

Proof uses a result by Van Lambalgen (1987), who showed that $A = A_0 \oplus_Z A_1$ is ML-random if and only if A_0 is ML-random and A_1 is ML^{A_0}-random.

A Counterexample

The splitting property of KL-random sequences is not a sufficient criterion for ML-randomness.

Theorem

There is a sequence A which is not computably random such that for each computable infinite and co-infinite set Z, $A \upharpoonright_Z$ is ML-random.

A Counterexample

The splitting property of KL-random sequences is not a sufficient criterion for ML-randomness.

Theorem

There is a sequence A which is not computably random such that for each computable infinite and co-infinite set Z, $A\restriction_Z$ is ML-random.

Kolmogorov-Loveland Stochasticity

One can modify betting strategies to obtain the concept of selection rules

Selection rules

- Select a position $k \in \mathbb{N}$.
- Specify whether to include the bit $A(k)$ in the selected subsequence.
- After the bit is revealed pick a new position not selected before.

Kolmogorov-Loveland Stochasticity

One can modify betting strategies to obtain the concept of selection rules

Selection rules

- Select a position $k \in \mathbb{N}$.
- Specify whether to include the bit $A(k)$ in the selected subsequence.
- After the bit is revealed pick a new position not selected before.

Kolmogorov-Loveland Stochasticity

Selection rules

- Select a position $k \in \mathbb{N}$.
- Specify whether to include the bit $A(k)$ in the selected subsequence.
- After the bit is revealed pick a new position not selected before.

A sequence A is Kolmogorov-Loveland stochastic if every infinite subsequence of A selected by a computable selection rule has limit frequency $1/2$.

Every ML-random sequence is KL-stochastic.
Shen (1988) showed that there are KL-stochastic sequences not ML-random.

Constructive Dimension

- There exists an interesting connection between the asymptotic complexity of sequences and Hausdorff dimension.
- Hausdorff dimension is defined via Hausdorff measures. Similar to Lebesgue measure, one can define effective versions [Lutz 2000].
- Constructive dimension, $\dim_{\mathrm{H}}^1$, can be characterized in terms of Kolmogorov complexity.

Theorem (Ryabko; Mayordomo)

The constructive dimension of a sequence A is given by

$$\dim_{\mathrm{H}}^1 A = \liminf_{n \to \infty} \frac{K(A\restriction_n)}{n}.$$

The Dimension of KL-Stochastic Sequences

It turns out that even the KL-stochastic sequences are already very close to incompressible.

Theorem

If R is KL-stochastic, then $\dim^1_{\mathrm{H}} R = 1$*.*

This implies, in particular, that all KL-random sequences have dimension 1, too.

Conclusion

- Non-monotonicity makes betting strategies much more powerful.
- In many ways, KL-randomness behaves like Martin-Löf randomness.
- However, none of the properties studied is a sufficient condition for ML-randomness.
- A proof that KL-randomness is equivalent to ML-randomness would would give a striking argument against Schnorr's criticism of Martin-Löf randomness.