

Outline

1 Notation and preliminaries

Notation

Algorithmic randomness

2 Randomness by martingales

Gambling strategies for automata

Finite-state dimension

3 Incompressibility

Sequential decompression

4 Automata and measure

Automata on ω -words

Subword complexity

5 Other concepts

Finite-state genericity

Unpredictability

Notation: Strings and Languages

Finite Alphabet $X = \{0, \dots, r-1\}$, cardinality $|X| = r$

Finite strings (words) $w = x_1 \cdots x_n \in X^*$, $x_i \in X$

Length $|w| = n$

Languages $W \subseteq X^*$

Infinite strings (ω -words) $\xi = x_1 \cdots x_n \cdots \in X^\omega$

Prefixes of infinite strings $\xi[0..n] \in X^*$, $|\xi[0..n]| = n$

ω -Languages $F \subseteq X^\omega$

X^ω as CANTOR space

Metric: $\rho(\eta, \xi) := \inf\{r^{-|w|} : w \in \mathbf{pref}(\eta) \cap \mathbf{pref}(\xi)\}$

Balls: $w \cdot X^\omega = \{\eta : w \in \mathbf{pref}(\eta)\} = \{\eta : w \sqsubset \eta\}$

Diameter: $\text{diam } w \cdot X^\omega = r^{-|w|}$

$\text{diam } F = \inf\{r^{-|w|} : F \subseteq w \cdot X^\omega\}$

Open sets: $W \cdot X^\omega = \bigcup_{w \in W} w \cdot X^\omega$

Closure: (Smallest closed set containing F)

$cl_\rho(F) = \{\xi : \mathbf{pref}(\xi) \subseteq \mathbf{pref}(F)\}$

Fact

$F \subseteq X^\omega$ is **closed** if and only if $\mathbf{pref}(\xi) \subseteq \mathbf{pref}(F)$ implies $\xi \in F$.

Algorithmic randomness

measure-theoretic paradigm

An ω -word is random if and only if it is not contained in a constructive null-set.

unpredictability paradigm

An ω -word is random if and only if no constructive predicting strategy can win against it.

incompressibility (complexity-theoretic) paradigm

An ω -word is random if and only if one cannot constructively compress infinitely many of its prefixes.

Measure

Measure on base sets: $\mu(w \cdot X^\omega) := r^{-|w|}$

Constructive null-sets: Unions of ω -languages of the form $\bigcap_{n \in \mathbb{N}} V_n \cdot X^\omega$,

where $V \subseteq \{(v, n) : v \in X^* \wedge n \in \mathbb{N}\}$ is constructive,

$V_n := \{v : (v, n) \in V\}$, and

$\mu(V_n \cdot X^\omega) \leq r^{-n}$.

Definition (Randomness)

$\xi \in X^\omega$ is *random* if and only if no constructive null-set contains ξ .

Predicting strategy: Gambling

Our model:

- Playing against an ω -word $\xi \in X^\omega$.
- Gambling strategy $\Gamma : X^* \times X \rightarrow [0, 1]$ (bet on outcome $x \in X$)
 $\sum_{x \in X} \Gamma(w, x) \leq 1$ for $w \in X^*$
- yields a (super-)martingale $\mathcal{V}_\Gamma : X^* \rightarrow \mathbb{R}_+$
- $\mathcal{V}_\Gamma(\xi[0..n])$ is the capital after the n th round, that is,

$$\mathcal{V}_\Gamma(\xi[0..n]) = r \cdot \Gamma(\xi[0..n], x) \cdot \mathcal{V}_\Gamma(\xi[0..n-1]), \text{ for } \xi(n) = x$$

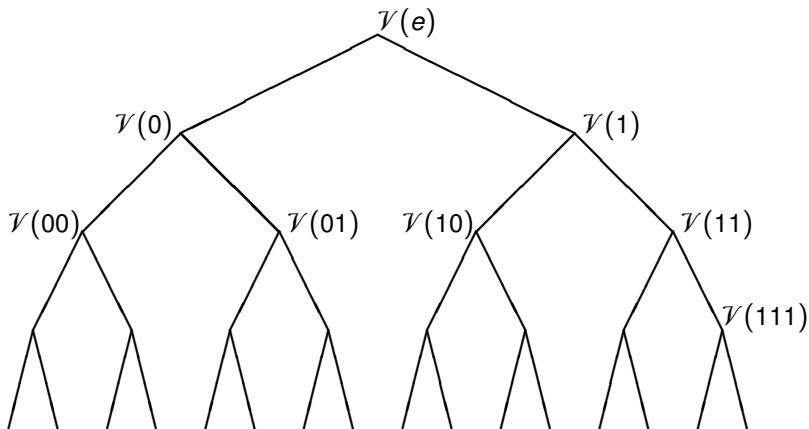
Fact (super-martingale property)

$$\mathcal{V}_\Gamma(w) \geq \frac{1}{r} \cdot \sum_{x \in X} \mathcal{V}_\Gamma(wx)$$

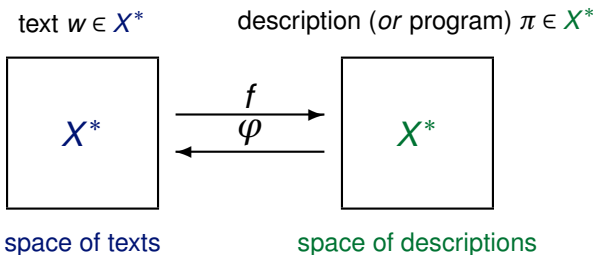
Definition (Randomness)

$\xi \in X^\omega$ is *random* if and only if no constructive gambling strategy Γ can win against ξ , that is, $\limsup_{n \rightarrow \infty} \mathcal{V}_\Gamma(\xi[0..n]) < \infty$.

Gambling strategies: Martingale \mathcal{V}



Compression: The principle of loss-less compression



f is injective and $\varphi(f(w)) = w$ for all $w \in X^*$

Complexity of w w.r.t. φ : $C_\varphi(w) := \inf\{|\pi| : \varphi(\pi) = w\}$

Definition (Randomness = Incompressibility)

$\xi \in X^\omega$ is *random* if and only if all constructive decompression functions φ satisfy $\exists c \forall n (C_\varphi(\xi[0..n])) \geq n - c$, that is, prefixes of ξ cannot be compressed.

Gambling finite automaton

Definition (Betting automaton)

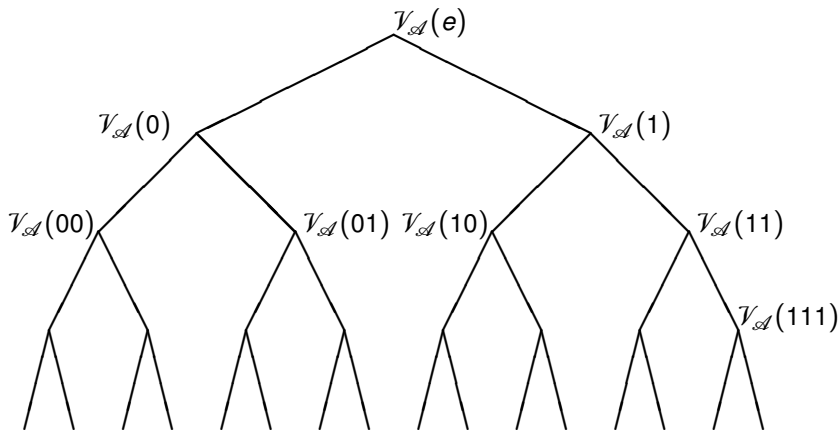
$\mathcal{A} = [X, Q, \mathbb{R}_{\geq 0}, q_0, \delta, \nu]$ is a finite-state betting automaton : \iff

- 1 Q is a finite set (of states), $q_0 \in Q$,
- 2 $\delta : Q \times X \rightarrow Q$,
- 3 $\nu : Q \times X \rightarrow \mathbb{R}_{\geq 0}$ and $\sum_{x \in X} \nu(q, x) \leq 1$, for all $q \in Q$.

Definition (Capital function of \mathcal{A})

$$\begin{aligned} \mathcal{V}_{\mathcal{A}}(e) &:= 1, \text{ and} \\ \mathcal{V}_{\mathcal{A}}(wx) &:= r \cdot \nu(\delta(q_0, w), x) \cdot \mathcal{V}_{\mathcal{A}}(w) \end{aligned}$$

Again: Gambling strategies: Martingale $\mathcal{V} = \mathcal{V}_{\mathcal{A}}$ ($X = \{0, 1\}$)



BOREL normality

Definition

An ω -word $\xi \in X^\omega$ is *BOREL normal* iff every subword (infix) $w \in X^*$ appears with the same frequency.

$$\forall w \left(\lim_{n \rightarrow \infty} \frac{|\{i : i \leq n \wedge \xi[0..i] \in X^* \cdot w\}|}{n} \right) = r^{-|w|}$$

The theorem of SCHNORR and STIMM

Theorem (SCHNORR/STIMM '72)

If $\xi \in X^\omega$ is BOREL normal then for every finite automaton \mathcal{A} it holds

- ① $\forall^\infty n (n \in \mathbb{N} \rightarrow \mathcal{V}_{\mathcal{A}}(\xi[0..n]) = \mathcal{V}_{\mathcal{A}}(\xi[0..n+1]))$, or
- ② $\exists \rho (0 \leq \rho < 1 \wedge \forall^\infty n (n \in \mathbb{N} \rightarrow \mathcal{V}_{\mathcal{A}}(\xi[0..n]) \leq \rho^n)$.

If $\xi \in X^\omega$ is **not** BOREL normal then there are a finite automaton \mathcal{A} and $\gamma > 0$ such that

- ③ $\exists^\infty n (n \in \mathbb{N} \wedge \mathcal{V}_{\mathcal{A}}(\xi[0..n]) \geq r^{\gamma \cdot n})$.

Partial Randomness: Finite-state dimension [DAI ET AL.'04]

Finite-state dimension tries to measure, for $\xi \in X^\omega$, the largest exponent γ_0 with

$$\mathcal{V}_{\mathcal{A}}(\xi[0..n]) \approx r^{\gamma_0 \cdot n + o(n)}.$$

for some finite automaton \mathcal{A} 'best fitted' to ξ .

More precisely, $\dim_{FS}(\xi) = 1 - \gamma_0 : \iff$

$$\exists \mathcal{A} \left(\limsup_{n \rightarrow \infty} \frac{\mathcal{V}_{\mathcal{A}}(\xi[0..n])}{r^{\gamma \cdot n}} > 0 \right) \text{ for } \gamma < \gamma_0, \text{ and}$$

$$\forall \mathcal{A} \left(\lim_{n \rightarrow \infty} \frac{\mathcal{V}_{\mathcal{A}}(\xi[0..n])}{r^{\gamma \cdot n}} = 0 \right) \text{ for } \gamma > \gamma_0.$$

Observe

The higher the dimension $\dim_{FS}(\xi)$ the 'more random' the ω -word.

Corollary

$\dim_{FS}(\xi) = 1$ if and only if ξ is BOREL normal.

Uniform extension to $F \subseteq X^\omega$ [DAI ET AL.'04]

Definition (Finite-state Dimension (DAI ET AL. '04))

$$\alpha_{\mathcal{A}}(F) := \inf\left\{\alpha : \forall \xi (\xi \in F \rightarrow \limsup_{n \rightarrow \infty} \frac{\mathcal{V}_{\mathcal{A}}(\xi[0..n])}{r^{(1-\alpha) \cdot n}} > 0)\right\}$$

$$\dim_{FS}(F) := \sup\{\alpha_{\mathcal{A}}(F) : \mathcal{A} \text{ is a finite automaton}\}$$

Observe

$1 - \alpha$ corresponds to the exponent γ .

Proposition

\dim_{FS} is monotone and stable: $\dim_{FS}(F \cup F') = \max\{\dim_{FS} F, \dim_{FS} F'\}$

Example

\dim_{FS} is not countably stable:

$$\dim_{FS}\{w \cdot v^\omega\} = 0 \text{ and } \dim_{FS}\{w \cdot v^\omega : w, v \in X^*\} = 1.$$

Finite-state dimension: Frequency

Let $h(\alpha) := -\alpha \cdot \log_2 \alpha - (1 - \alpha) \cdot \log_2(1 - \alpha)$ be the binary SHANNON entropy and let

$$\text{FREQ}(\alpha) := \left\{ \xi : \xi \in \{0, 1\}^\omega \wedge \lim_{n \rightarrow \infty} \frac{|\xi[0..n]|_1}{n} = \alpha \right\}$$

Theorem (DAI ET AL.'04)

Let $\alpha \in [0, 1]$ be rational. Then the following hold.

- 1 There is an ω -word $\xi \in X^\omega$ having $\dim_{FS}(\xi) = \alpha$, and
- 2 $\dim_{FS}(\text{FREQ}(\alpha)) = h(\alpha)$.

The entropy of languages (CHOMSKY/MILLER '58)

Definition (*Entropy* of a language $W \subseteq X^*$)

$$\mathbf{H}_W := \limsup_{n \rightarrow \infty} \frac{\log_r(1 + |W \cap X^n|)}{n}$$

Proposition

Let $W \subseteq X^*$. Then

- 1 $\limsup_{n \rightarrow \infty} \frac{|W \cap X^n|}{r^{\alpha \cdot n}} = \begin{cases} 0, & \text{if } \alpha > \mathbf{H}_W \text{ and} \\ \infty, & \text{if } \alpha < \mathbf{H}_W, \text{ or} \end{cases}$
- 2 $\sum_{w \in W} |X|^{-\alpha \cdot |w|} = \begin{cases} \infty, & \text{if } \alpha < \mathbf{H}_W \text{ and} \\ < \infty, & \text{if } \alpha > \mathbf{H}_W. \end{cases}$

Connection to regular languages

$$\mathbf{infix}(P) := \{w : \exists v(w \cdot v \in \mathbf{pref}(P))\}$$

Lemma (Bourke et al. '05)

$$\dim_{FS} F \leq \inf\{\mathbf{H}_W : W \subseteq X^* \wedge \mathbf{infix}(F) \subseteq W \wedge W \text{ is regular}\}$$

Fact

If $W \subseteq X^*$ is regular then

$$\mathbf{H}_W = \mathbf{H}_{\mathbf{pref}(W)} = \mathbf{H}_{\mathbf{infix}(W)}.$$

Decompression by transducers

Definition

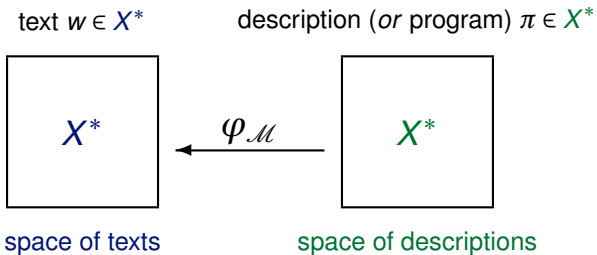
$\mathcal{M} = [X, Y, Q, q_0, \delta, \lambda]$ is a *generalised sequential machine* (or *finite transducer*) : \iff

- ① Q is a finite set (of states), $q_0 \in Q$,
- ② $\delta : Q \times X \rightarrow Q$,
- ③ $\lambda : Q \times X \rightarrow Y^*$.

$\varphi_{\mathcal{M}}$ is the mapping related to \mathcal{M} if $\varphi_{\mathcal{M}}(w) = \lambda(q_0, w)$.

In the sequel we will only consider transducers with $Y = X$.

Incompressibility: Complexity



Complexity of w w.r.t. to the transducer \mathcal{M} :

$$C_{\mathcal{M}}(w) := \inf\{|\pi| : \varphi_{\mathcal{M}}(\pi) = w\}$$

Fact (Combinatorial lower bound)

If $W \subseteq X^$ has at least m elements then there is a $w \in W$ such that $C_{\varphi}(w) \geq \log_r m$.*

Decompression and finite-state dimension

Definition (Decompression along an input)

$$\vartheta_{\mathcal{M}}(\eta) := \liminf_{n \rightarrow \infty} \frac{n}{|\varphi_{\mathcal{M}}(\eta[0..n])|},$$

where \mathcal{M} is a finite transducer and $\varphi_{\mathcal{M}}$ its related mapping.

Let $\bar{\varphi}(\eta) := \lim_{v \rightarrow \eta} \varphi(v)$ or $\mathbf{pref}(\bar{\varphi}(\eta)) = \mathbf{pref}(\varphi(\mathbf{pref}(\eta)))$

The value $\vartheta_{\mathcal{M}}(\eta) \cdot n$ is the asymptotic necessary amount of information to produce the first n symbols of $\bar{\varphi}_{\mathcal{M}}(\eta)$ (via \mathcal{M} and η).

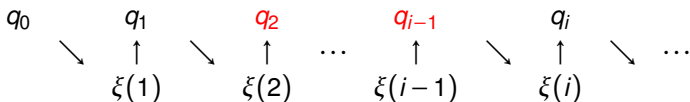
Theorem (SHEINWALD ET AL. '91 AND '95)

$$\dim_{FS}(\xi) = \inf \{ \vartheta_{\mathcal{M}}(\eta) : \mathcal{M} \text{ finite transducer} \wedge \xi = \bar{\varphi}_{\mathcal{M}}(\eta) \}$$

Automata on ω -words: Büchi-automata

Automaton: $\mathcal{A} = (X, Q, \Delta, q_0, Q_{\text{fin}})$ with
 $\Delta \subseteq Q \times X \times Q$, $q_0 \in Q$, $Q_{\text{fin}} \subseteq Q$

Run on ξ : $(q_i)_{i \in \mathbb{N}}$ with $\forall i \geq 0 : (q_i, \xi(i+1), q_{i+1}) \in \Delta$



\mathcal{A} accepts ξ : $\exists (q_i)_{i \in \mathbb{N}} \quad \forall i \geq 0 : (q_i, \xi(i+1), q_{i+1}) \in \Delta \quad \wedge$
 $\exists^\infty k : q_k \in Q_{\text{fin}}$

\mathcal{A} accepts F : $F = \{ \xi : \mathcal{A} \text{ accepts } \xi \}$

Regular ω -languages

Definition (Regular ω -language)

An ω -language $F \subseteq X^\omega$ is called *regular* if and only if F is accepted by a finite automaton

Theorem (BÜCHI '62)

- 1 An ω -language $F \subseteq X^\omega$ is regular if and only if $F = \bigcup_{i=1}^n W_i \cdot V_i^\omega$ for some $n \in \mathbb{N}$ and regular languages $W_i, V_i \subseteq X^*$.
- 2 The set of regular ω -languages over X is closed under Boolean operations.

Deterministic BÜCHI automata

$F \in \mathcal{DB}$ if and only if F is accepted by a deterministic automaton.

Theorem (LANDWEBER '69)

- 1 $F \in \mathcal{DB}$ if and only if F is regular and of the form $F = \bigcap_{i \in \mathbb{N}} V_i \cdot X^\omega$,
that is, a countable intersection of ω -languages open in CANTOR space.
- 2 If $F \subseteq X^\omega$ is regular and closed in CANTOR space then $F \in \mathcal{DB}$.
- 3 \mathcal{DB} is closed under union and intersection but not under complement.
- 4 If $F \subseteq X^\omega$ is regular then F is a countable union of $F_i \in \mathcal{DB}$.

Density in X^ω

Definition (Nowhere dense in X^ω)

An ω -language F is *nowhere dense* in X^ω if and only if its closure does not contain a non-empty open subset, that is, for all $w \in \mathbf{pref}(F)$ there is a $v \in X^*$ such that $w \cdot v \notin \mathbf{pref}(F)$.

Definition (BAIRE category)

An ω -language F is *meagre* or of *first BAIRE category* if and only if it is a countable union of nowhere dense sets.

It is of *second BAIRE category* if and only if it is not of first BAIRE category.

Complements of meagre sets are referred to as *co-meagre* or *residual*.

Density of regular ω -languages: Forbidden subwords

Theorem (St'76)

Let F be a regular ω -language.

- ① F is nowhere dense if and only if there is word $w \in X^*$ such that

$$F \subseteq X^\omega \setminus X^* \cdot w \cdot X^\omega.$$

- ② F is of first BAIRE category if and only if

$$F \subseteq \bigcup_{w \in X^*} X^\omega \setminus X^* \cdot w \cdot X^\omega.$$

Balanced measures

Definition (Balance condition for measures)

A finite positive measure μ on X^ω is referred to as *balanced* if and only if

- ① $\mu(w \cdot X^\omega)$ exists for all $w \in X^*$, and
- ② there is a constant $c_\mu > 0$ such that

$$c_\mu \cdot \mu(w \cdot X^\omega) \leq \mu(w \cdot x \cdot X^\omega)$$

for all $w \in X^*$ and $x \in X$.

Observe

The mass $\mu(w \cdot X^\omega)$ of $w \cdot X^\omega$ distributes as

$$\mu(w \cdot X^\omega) = \sum_{x \in X} \mu(w \cdot x \cdot X^\omega).$$

Regular null-sets

Theorem (St'76, St'98)

Let F be a regular ω -language and let μ be a balanced measure on X^ω .

- 1 If $F \in \mathcal{DB}$ then $\mu(F) = 0$ if and only if there is word $w \in X^*$ such that

$$F \subseteq X^\omega \setminus X^* \cdot w \cdot X^\omega.$$

- 2 $\mu(F) = 0$ if and only if

$$F \subseteq \bigcup_{w \in X^*} X^\omega \setminus X^* \cdot w \cdot X^\omega.$$

Definition (Disjunctivity)

An ω -word $\xi \in X^\omega$ is called *disjunctive* (or *rich* or *saturated*) if and only if it contains every word $w \in X^*$ as subword (infix) [$\mathbf{infix}(\xi) = X^*$].

Conclusion: Randomness in the measure-theoretical sense

Automatic randomness = Disjunctivity

Measure and category for regular ω -languages [St'76, St'98, Varacca/Völzer '06 and Völzer/Varacca '12]

Let μ be a finite, positive and balanced measure on X^ω .

Then every regular ω -language is measurable, and, moreover, for regular ω -languages $F \subseteq X^\omega$ the following equivalences between *measure* and '*category*' hold.

	Measure	Category (Density)
very large	$\mu(F) = \mu(X^\omega)$	F is residual (co-meagre)
large	$\mu(F) \neq 0$	F is of 2 nd BAIRE category
small	$\mu(F) = 0$	F is of 1 st BAIRE category (meagre)
very small	$\mu(cI_\rho(F)) = 0$	F is nowhere dense

Partial randomness: Subword complexity

Definition (Asymptotic subword complexity of $P \subseteq X^* \cup X^\omega$)

$$\tau(P) := \limsup_{n \rightarrow \infty} \frac{\log_r |\mathbf{infix}(P) \cap X^n|}{n}$$

$$\mathbf{infix}(P) \cap X^{n+m} \subseteq (\mathbf{infix}(P) \cap X^n) \cdot (\mathbf{infix}(P) \cap X^m)$$

Fact

The limit exists and equals $\tau(P) = \inf \left\{ \frac{\log_r |\mathbf{infix}(P) \cap X^n|}{n} : n \in \mathbb{N} \right\}$.

Proposition

τ is monotone and stable: $\tau(P \cup P') = \max\{\tau(P), \tau(P')\}$

Example

τ is not countably stable:

$$\tau(w \cdot v^\omega) = 0 \text{ and } \tau(X^* \cdot v^\omega) = 1.$$

Subword complexity and the entropy of regular languages

Proposition

$0 \leq \tau(P) \leq 1$ and an ω -word $\xi \in X^\omega$ is disjunctive if and only if $\tau(\xi) = 1$.

Fact

$$\tau(P) = \inf\{\mathbf{H}_W : W \subseteq X^* \wedge \mathbf{infix}(P) \subseteq W \wedge W \text{ is regular}\}$$

Corollary

If $F \subseteq X^*$ is regular and closed, that is, $F = \{\xi : \mathbf{pref}(\xi) \subseteq \mathbf{pref}(F)\}$, then

$$\tau(F) = \mathbf{H}_{\mathbf{pref}(F)} = \mathbf{H}_{\mathbf{infix}(F)}.$$

Relations between \dim_{FS} and τ

Corollary

$$\dim_{FS} F \leq \tau(F)$$

Example

Every BOREL normal ω -word is disjunctive.

The ω -word $\eta = \prod_{w \in X^*} 0^{|w|} \cdot w$ is disjunctive but not BOREL normal.

Proposition

Let $F \subseteq X^\omega$ be non-empty and regular.

- ① Then $\max_{\xi \in F} \tau(\xi)$ exists and $\max_{\xi \in F} \tau(\xi) = \max_{\xi \in F} \dim_{FS} \{\xi\}$.
- ② If, moreover, $F = \{\xi : \mathbf{pref}(\xi) \subseteq \mathbf{pref}(F)\}$ then $\tau(F) = \dim_{FS} F$.

The range of \dim_{FS} and τ

Theorem (MOLDAGALIYEV ET AL. '18)

For every $t, 0 \leq t \leq 1$, there is a $\xi \in X^\omega$ such that $\tau(\xi) = t$.

Corollary

For every $t, 0 \leq t \leq 1$, there is a $\xi \in X^\omega$ such that $\dim_{FS}(\xi) = t$.

Finite-state genericity [AMBOS-SPIES/BUSSE'03]

Definition (k -labelled automaton)

Let $k \geq 1$. A finite automaton $\mathcal{K} = [X, Q, q_0, \delta, \lambda_k]$ is called a k -labelled automaton if and only if $\lambda_k : Q \rightarrow X^k$.

Definition

- 1 An ω -word $\xi \in X^\omega$ meets \mathcal{K} if and only if $\xi[0..n]\lambda_k(\delta(q_0, \xi[0..n])) \in \mathbf{pref}(\xi)$ for some $n \in \mathbb{N}$.
- 2 An ω -word $\xi \in X^\omega$ is *finite-state generic* if and only if, for every k , the ω -word ξ is met by every k -labelled automaton \mathcal{K} .

Theorem (AMBOS-SPIES/BUSSE'03)

An ω -word ξ is *disjunctive* if and only if, for every k , it is met by every k -labelled automaton $\mathcal{K} = [X, Q, q_0, \delta, \lambda_k]$.

Predicting automaton

- Playing against an ω -word $\xi \in X^\omega$.
- Knowing $\xi[0..n-1]$ predict the next symbol $\xi(n)$ or Skip.
- Predict infinitely often.
- All predictions have to be correct!

Definition (Predicting automaton)

$\mathcal{A} = [X, Q, q_0, \delta, \lambda]$ is a finite-state predicting automaton : \iff

- 1 Q is a finite set (of states), $q_0 \in Q$,
- 2 $\delta : Q \times X \rightarrow Q$,
- 3 $\lambda : Q \rightarrow X \cup \{e\}$. [e – empty word, that is, Skip]

Prediction

Definition (TADAKI '14)

A predicting automaton $\mathcal{A} = [X, Q, q_0, \delta, \lambda]$ predicts $\xi \in X^\omega$ if and only if

- ① $\lambda(\delta(q_0, \xi[0..n-1])) = \xi(n)$ for infinitely many n , and
- ② if $\lambda(\delta(q_0, \xi[0..n-1])) \neq \xi(n)$ then $\lambda(\delta(q_0, \xi[0..n-1])) = e$.

Theorem

Let $\mathcal{A} = [X, Q, q_0, \delta, \lambda]$ be a predicting automaton.

- ① If \mathcal{A} predicts ξ then ξ is not disjunctive.
- ② If, moreover, $X = \{0, 1\}$ then every non-disjunctive ξ is predicted by some automaton \mathcal{A}_ξ .

Weak Prediction

Definition

A predicting automaton $\mathcal{A} = [X, Q, q_0, \delta, \lambda]$ *weakly predicts* $\xi \in X^\omega$ if and only if

- 1 $\lambda(\delta(q_0, \xi[0..n-1])) \in X$ for infinitely many n , and
- 2 if $\lambda(\delta(q_0, \xi[0..n-1])) \in X$ then $\lambda(\delta(q_0, \xi[0..n-1])) \neq \xi(n)$.

Theorem

An ω -word ξ is weakly predictable by some automaton $\mathcal{A} = [X, Q, q_0, \delta, \lambda]$ if and only if it is non-disjunctive.

