Factor complexity and Symbolic dynamics

Paulina Cecchi Bernales.

Departamento de Matemáticas–Facultad de Ciencias–Universidad de Chile.

EPAA Kickoff meeting, UBA

Buenos Aires, October 22, 2024.

Words, languages

- A finite set of symbols, $\mathcal{A} = \{0, 1, \cdots, d-1\}$, $d \geq 2$, is called an alphabet.
- Given an alphabet $\mathcal{A}$, an infinite word with symbols in $\mathcal{A}$ is an element of $\mathcal{A}^{\mathbb{N}}$ or $\mathcal{A}^{\mathbb{Z}}$,

$$x = x_0 x_1 x_2 x_3 \cdots, \quad x_i \in \mathcal{A}.$$

- A finite word with symbols in $\mathcal{A}$ is an element of $\mathcal{A}^\star := \bigcup_{n\in\mathbb{N}} \mathcal{A}^n$.
- For $x = (x_n)_{n\in\mathbb{N}} \in \mathcal{A}^{\mathbb{N}}$, a factor of x is a finite word appearing in x,

$$w = x_j \cdots x_k, \quad k \geq j.$$

- The length of w is $k - j + 1$ and is denoted $|w|$.

Factor complexity

- The language $\mathcal{L}_x$ of x is the set of all factors of x.
- The factor complexity of $x \in \mathcal{A}^{\mathbb{N}}$, is the map $p_x : \mathbb{N} \to \mathbb{N}$ given by

$$p_x(n) = |\mathcal{L}_x \cap \mathcal{A}^n|$$

- For instance, the complexity of

$$x = 010101010101010101 \cdots$$

satisfies $p_x(0) = 1$ (empty word) and $p_x(n) = 2$ for all $n \geq 1$.

Minimal complexity

- Periodic sequences: if $w \in \mathcal{A}^*$ is a non-empty and primitive word, and

$$x = wwww \cdots ,$$

then $p_x(0) = 1$, $p_x(1) = |\mathcal{A}|$ and $p_x(n) = |w|$ for all $n \geq |w|$.

- *Eventually* periodic sequences: if $w \in \mathcal{A}^*$ is a non-empty and primitive finite word, and

$$x = twwww \cdots ,$$

where t does not end with the same letter than w, then $p_x(n) = |tw|$ for each $n \geq |tw|$.

Maximal complexity

- If $|\mathcal{A}| = d$, then $p_x(n) \leq d^n$ for any $x \in \mathcal{A}^{\mathbb{N}}$, for any $n \in \mathbb{N}$.
- $\mathcal{A}^\star$ is countable. Let $\{w_0, w_1, w_2, \cdots\}$ be an enumeration.
- For instance, if $\mathcal{A} = \{a, b\}$,

  $\mathcal{A}^\star = \{\varepsilon, a, b, aa, ab, ba, bb, aaa, aab, aba, abb, baa, bab, bba, \cdots\}$
- Let x be the word obtained from the concatennation of the $w_i's$,

$$x = w_0 w_1 w_2 w_3 \cdots$$

- By construction, $p_x(n) = d^n$.

The Morse–Hedlund Theorem

- The map p_x is non-decreasing: $p_x(n) \leq p_x(n+1)$.
- **Theorem [Hedlund–Morse '40]**: Let x be an infinite word on a given alphabet, let p_x its factor complexity. Then, either x is eventually periodic or p_x is strictly increasing.
- **Corollary**: If there is an $n \in \mathbb{N}$ such that $p_x(n) \leq n$, then x is eventually periodic.
- **Question**: ¿Are there infinite words which are not eventually periodic and $p_x(n) = n + 1$ for all $n \in \mathbb{N}$?

Sturmian words

- Consider the following sequence of finite words $(f_i)_{i\in\mathbb{N}}$ in $\mathcal{A} = \{a, b\}$,

$$f_i = \begin{cases} a & \text{if } i = 0 \\ ab & \text{if } i = 1 \\ f_{i-1}f_{i-2} & \text{if } i > 1. \end{cases}$$

- Since each f_i is a *prefix* of f_{i+1}, there exists a unique infinite word x_F such that each f_i is a prefix of x_F.
- x_F is called the Fibonacci infinite word,

$$x_F = abaababaabaababaab\cdots$$

- x_F has complexity $p_{x_F}(n) = n + 1$ for all $n \in \mathbb{N}$.
- Infinite words with complexity $n + 1$ are known as *Sturmian words*.

Integer base expansion of a real number

- Let $0 < \alpha < 1$ be a real number, let $d \geq 2$ a positive integer. Consider the base-d expansion of α,

$$\sum_{n \geq 0} \frac{a_n}{d^{n+1}} = 0.a_0 a_1 a_2 \cdots , \quad a_i \in \{0, 1, \cdots , d-1\}.$$

- To α we associate the infinite word

$$x_{\alpha,d} = a_0 a_1 a_2 a_3 \cdots \in \{0, 1, \cdots , d-1\}^{\mathbb{N}}.$$

- **Theorem**: A real number $0 < \alpha < 1$ es rational if and only if $x_{\alpha,d}$ is eventually periodic.

Complexity and transcendence

- **Theorem [Ferenczi–Maduit '97]**: If the binary expansion of α satisfies $p_{x_{\alpha,2}}(n) = n + 1$ for all $n \in \mathbb{N}$, then α is transcendent.
- Generalizations: for a d-letter alphabet, there are two ways of generalize the notion of *Sturmian word* based on their complexity,
  - $p_x(n) = n + d - 1$,
  - $p_x(n) = (d - 1)n + 1$.
- **Theorem [Ferenczi–Maduit '97], [Risley–Zamboni '00]**: If for some $d \geq 2$, the base-d expansion of α satisfies $p_{x_{\alpha,d}}(n) = n + d - 1$ for all $n \in \mathbb{N}$, or $p_{x_{\alpha,d}}(n) = (d - 1)n + 1$ for all $n \in \mathbb{N}$ then α is transcendent.

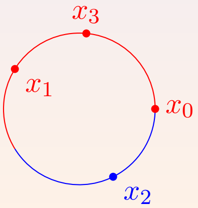
Dynamical systems

- **Dynamical system**: Set with some structure and structure-preserving group action on it.
- Throughout this talk, $G = \mathbb{Z}$.
- Topological dynamical system: (X, T), where X is a metric space and $T : X \to X$ is a homeomorphism (defines a continuous action of $\mathbb{Z}$ on X). $\leftarrow$ Topological dynamics.
- Measure-theoretic dynamical system: (X, T, μ), where X is a measure space, T is a measure preserving bijection $\mu(T^{-1}A) = \mu(A) \quad \forall A \in \mathcal{B}(X)$. $\leftarrow$Ergodic theory
- Both: X is a metric space and μ is a measure defined on the Borel $\sigma-$algebra $\mathcal{B}_X$, T is a measure preserving homeomorphism.

Orbit coding

- Example: rotation on $\mathbb{S}^1$. Consider the rotation of angle $\alpha \in \mathbb{R} \setminus \mathbb{Q}$ on the unit circle, $R_\alpha : [0,1) \to [0,1)$, $R_\alpha(x) = x + \alpha \mod 1$.
- Consider the *orbit coding sequence* $c(x) = (c(x)_i)_{i \in \mathbb{Z}}$ associated to $x \in \mathbb{S}^1$,

$$c(x)_i = \begin{cases} 0 & \text{if } T^i(x) \in [0, 1-\alpha) \\ 1 & \text{if } T^i(x) \in [1-\alpha, 1) \end{cases}$$



$w = 0010$

Subshifts

- Let $\mathcal{A} = \{0, 1, \cdots, d-1\}$ be an alphabet, $d \geq 2$.
- Consider $\mathcal{A}^{\mathbb{Z}} = \{(x_n)_{n\in\mathbb{Z}} \mid x_n \in \mathcal{A} \quad \forall n \in \mathbb{Z}\}$ ((bi-)infinite words with symbols in $\mathcal{A}$), equiped with the product topology of the discrete topology on each copy of $\mathcal{A}$.
- Let $S : \mathcal{A}^{\mathbb{Z}} \to \mathcal{A}^{\mathbb{Z}}$ be the shift map:
  $S((x_n)_{n\in\mathbb{Z}}) = (x_{n+1})_{n\in\mathbb{Z}}$.
- If $X \subseteq \mathcal{A}^{\mathbb{Z}}$ is closed and shift-invariant $(S(X) = X)$, $(X, S\mid_X)$ is a topological dynamical system called shift or subshift on $\mathcal{A}$.
- The whole system $(\mathcal{A}^{\mathbb{Z}}, S)$ is called fullshift on $\mathcal{A}$.

Complexity and Entropy

- Let (X, S) be a subshift on the alphabet $\mathcal{A}$. Let $\mathcal{L}_X$ be the set of all factor appearing on all elements of X.
- The factor complexity of X is the map $p_X : \mathbb{N} \to \mathbb{N}$ given by

$$p_X(n) = |\mathcal{L}_X \cap \mathcal{A}^n|.$$

- We consider minimal symbolic systems: every orbit is dense.
- Equivalently, the language $\mathcal{L}_X$ is uniformly recurrent: every factor appearing in any element $x \in X$, appears infinitely often and with bounded gaps.
- In minimal systems, $\mathcal{L}_x = \mathcal{L}_X$ for all $x \in X$.

Complexity and Entropy

- The topological entropy of a symbolic system corresponds to the limit $\lim_{n\to\infty} \frac{\log(p_X(n))}{n}$.
- Factor complexity is a finer notion of randomness.
- Among zero topological entropy there is a wide variety of different complexity behaviors.

Complexity and Conjugacy

- Two topological dynamical systems (X_1, T_1) and (X_2, T_2) are conjugate if there exists a homeomorphism $h : X_1 \to X_2$ such that
$$h \circ T_1 = T_2 \circ h$$
- The factor complexity p_X is not preserved under conjugacy.
- However, if (X, S) and (Y, S) are conjugate subshifts, then $\exists C$ such that, $\forall n > C$
$$p_X(n - C) \le p_Y(n) \le p_X(n + C)$$ **[Ferenczi '96].**
- In particular, the asymptotic behavior is the same.
- The behavior of the factor complexity function imposes a restriction on the conjugacy class of a given symbolic system.

Invariant measures

- Let (X, T) be a topological dynamical system. A Borel probability measure μ on X is T-invariant if

$$\forall A \in \mathcal{B}_X, \mu(T^{-1}A) = \mu(A).$$

- An invariant measure is called ergodic if $\forall A \in \mathcal{B}_X$, $T^{-1}A = A \implies \mu(A) = 0 \vee \mu(A) = 1$.

- Let $\mathcal{M}(X, T)$ denote the set of all invariant probability measures on X. It has a simplex structure, whose extreme points are the ergodic measures of the system.

- If $|\mathcal{M}(X, T)| = 1$, the system is said to be uniquely ergodic.

Invariant measures

- Let (X, S) be a minimal subshift, let p_X be its complexity function. If $\liminf_{n\to\infty} \frac{p_X(n)}{n} < \alpha < +\infty$, then the number of extreme points of $\mathcal{M}(X, S)$ is at most $\max(\lfloor\alpha\rfloor, 1)$ **[Boshernitzan '84]**.

- If $\limsup_{n\to\infty} \frac{p_X(n)}{n} < 3$, then (X, S) is uniquely ergodic **[Boshernitzan '84]**

- The behaviour of the factor complexity function imposes a restriction on the number of ergodic measures of a minimal symbolic system.

- The Boshernitzan condition is optimal.

Invariant measures

- **Theorem [Cyr–Kra '20]**: If p_n is any sequence of natural numbers such that $\liminf_{n\to\infty} \frac{p_n}{n} = \infty$, then there exists a minimal subshift (X, S) such that $\mathcal{M}(X, S)$ has uncountably many extreme points, and which satisfies

$$\liminf_{n\to\infty} \frac{p_X(n)}{p_n} = 0.$$

Invariant measures

- **Theorem [CB-Donoso '22]**: If p_n is any sequence of natural numbers such that $\liminf_{n\to\infty} \frac{p_n}{n} = \infty$ and K is any Choquet simplex, then there exists a minimal subshift (Y, S) such that
  - $\mathcal{M}(Y, S) \cong K$,
  - The factor complexity of (Y, S) satisfies $\lim p_Y(n)/p_n = 0$.
- **Theorem [CB-Donoso '24]**: Let K be a Choquet simplex. Let $(g_n)_{n\in\mathbb{N}}$ be a sequence of positive real numbers which is subexponential. Then, there exists a zero-entropy minimal subshift (Y, S) such that
  - $\mathcal{M}(Y, S) \cong K$,
  - The factor complexity of (Y, S) satisfies $\liminf g_n/p_X(n) = 0$.

Invariant measures and Entropy

- **Theorem** **[CB-Donoso '24]**: Let K be a Choquet simplex and $\alpha \geq 1$ be given. Let $(g_n)_{n\in\mathbb{N}}$ be a sequence of positive real numbers satisfying $\lim \frac{\log(g_n)}{n} = \log(\alpha)$. Then, there exists a minimal subshift (X, S) such that
  - $\mathcal{M}(X, S) \cong K$.
  - $h(X, S) = \log(\alpha)$.
  - The complexity of (X, S) satisfies $\liminf g_n/p_X(n) = 0$.

Questions

- Let K be a Choquet simplex and let $(p_n)_{n\in\mathbb{N}}$ $(g_n)_{n\in\mathbb{N}}$ be sequences of positive real numbers satisfying $\lim \frac{\log(g_n)}{n} = \log(\alpha)$.
  Is there a minimal subshift (X, S) such that
  - $\mathcal{M}(X, S) \cong K$.
  - $h(X, S) = \log(\alpha)$.
  - The complexity of (X, S) satisfies $\liminf p_X(n)/p_n = 0$???

Questions

- Let K be a Choquet simplex, let $(p_n)_{n\in\mathbb{N}}$ and $(g_n)_{n\in\mathbb{N}}$ two sequences of real numbers such that p_n is superlinear, g_n is at most exponential, and $\lim_{n\to\infty} \frac{g_n}{p_n} = 0$.
  Is there a minimal subshift (X, S) such that
  - $\mathcal{M}(X, S) \cong K$.
  - The complexity of (X, S) satisfies $\lim_{n\to\infty} \frac{g_n}{p_X(n)} = \lim_{n\to\infty} \frac{p_X(n)}{p_n} = 0$???

GRACIAS!