What does a former set theorist do in mathematical biology?

Winfried Just
Department of Mathematics, Ohio University

Based on joint work with Ying Xin
Johns Hopkins University Medical School

Presented jointly with Sergio García
Department of Mathematics, York University

Toronto Set Theory Seminar
Fields Institute, October 4, 2019

A famous article

There is a famous article titled:

Mathematics Is Biology's Next Microscope, Only Better; Biology Is Mathematics' Next Physics, Only Better

by Joel E. Cohen in PLoS Biology, 2004.

It sums up the sentiment around the time when I made the switch to mathematical biology:

Biology was believed to supplant physics as the primary source of new mathematical problems.

A Set Theorist's Question: Could biology also become a new source for problems in the foundations of mathematics?

Why is mathematical biology even possible?

Mathematical biology is implicitly based on the assumption that low-dimensional approximate models for explaining relevant behavior of complex, high-dimensional biological systems always exist.

Empirical evidence indicates that this may actually be true, at least to a large extent.

Question: But why??

Conjecture: This question can be formulated in a mathematically rigorous way.

The answer might shed light on *design principles* of biological systems.

Measuring complexity of a dynamical system: The idea

For the remainder of this talk, a dynamical system is a pair (X, F), where X is a compact metric space with distance function D and $F : X \to X$ is a homeomorphism.

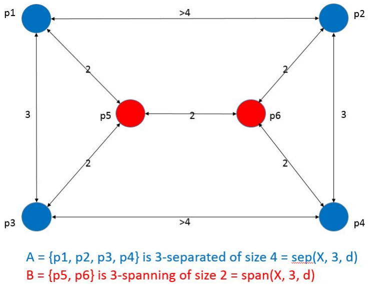
- A (forward) trajectory is a sequence $(F^t(x))_{t=0}^{\infty}$ for some $x \in X$.
- For $\varepsilon > 0$, let $N_T(\varepsilon, D)$ be the number of trajectories that are distinguishable at resolution ε within T time steps.
- In complex systems, for sufficiently small $\varepsilon > 0$, this number $N_T(\varepsilon, D)$ scales approximately like $B(\varepsilon)^T$ for some $B(\varepsilon) > 1$.
- Thus we can use the growth rate of $N_T(\varepsilon, D)$ to define a measure of complexity of the system: Roughly speaking, the larger the base $B(\varepsilon)$, the more complex is the system.

Separation numbers and spanning numbers

Let (X, d) be metric space, and let $\varepsilon > 0$.

We define the separation number $sep(X, \varepsilon, d)$ as the **largest** size of a subset $Y \subset X$ such that $d(x, x') \geq \varepsilon$ for all $x, x' \in Y$,

and the spanning number $span(X, \varepsilon, d)$ as the **smallest** size of a subset $Y \subset X$ such that for all $x \in X$ there exists $y \in Y$ with $d(x, y) < \varepsilon$.



Two definitions of $N_T(\varepsilon, D)$

We could define $N_T(\varepsilon, D)$ as the largest size $sep(X, \varepsilon, D_T)$ of a (T, ε)-separated subset of X, that is, of a set $Y \subseteq X$ such that for all $x, x' \in Y$ there exists a $0 \leq t < T$ with $D(F^t(x), F^t(x')) \geq \varepsilon$.

Or we could define $N_T(\varepsilon, D)$ as the smallest size $span(X, \varepsilon, D_T)$ of a (T, ε)-spanning subset of X, that is, of a set $Y \subseteq X$ such that for all $x \in X$ there exists $y \in Y$ such that for all $0 \leq t < T$ we have $D(F^t(x), F^t(y)) < \varepsilon$.

The separation numbers $sep(X, \varepsilon, D_T)$ and spanning numbers $span(X, \varepsilon, D_T)$ are always finite and satisfy the inequality $sep(X, \varepsilon, D_T) \geq span(X, \varepsilon, D_T) \geq sep(X, 2\varepsilon, D_T)$.

Measuring complexity: Topological entropy

- For $\varepsilon > 0$, let $N_T(\varepsilon, D)$ be the number of trajectories that are distinguishable at resolution ε within T time steps.
- In complex systems, for sufficiently small $\varepsilon > 0$, this number $N_T(\varepsilon, D)$ scales approximately like $B(\varepsilon)^T$ for some $B(\varepsilon) > 1$.
- Thus we can use the growth rate of $N_T(\varepsilon, D)$ to define a measure of complexity of the system: Roughly speaking, the larger the base $B(\varepsilon)$, the more complex is the system.

Define the topological entropy of the dynamical system (X, F) as:

$h(X, F) = \lim_{\varepsilon \to 0^+} \limsup_{T \to \infty} \frac{\ln N_T(\varepsilon, D)}{T}$.

In particular, can define the topological entropy $h(X, F)$ as:

$\lim_{\varepsilon \to 0^+} \limsup_{T \to \infty} \frac{\ln sep(X, \varepsilon, D_T)}{T} = \lim_{\varepsilon \to 0^+} \limsup_{T \to \infty} \frac{\ln span(X, \varepsilon, D_T)}{T}$.

Question: But why lim sup? Could we use lim instead?

It had been widely believed for the last 50 years or so that the answer is negative for the definitions based on separation or spanning numbers. But no actual counterexamples were known.

Our main theorem

Theorem

There exists a system (X, F) *with a metric* D *on* X *such that for some* $\varepsilon > 0$ *we have:*

$$\liminf_{T \to \infty} \frac{\ln sep(X, \varepsilon, D_T)}{T} < \limsup_{T \to \infty} \frac{\ln sep(X, \varepsilon, D_T)}{T},$$

$$\liminf_{T \to \infty} \frac{\ln span(X, \varepsilon, D_T)}{T} < \limsup_{T \to \infty} \frac{\ln span(X, \varepsilon, D_T)}{T}.$$

The system (X, F) *is minimal; i.e., every forward trajectory is dense in* X*.*

Remark: We have more results along these lines. They are included in Ying's dissertation (2018).

But the blue part was proved more recently.

What would a shockingly simple example look like?

Let $X^+ = {}^{\mathbb{Z}}A$ be the space of two-sided sequences x of symbols from some given finite alphabet A with $|A| > 1$.

Define a metric ρ on ${}^{\mathbb{Z}}A$ as follows: $\rho(x, y) = \kappa^{-\Delta(x,y)}$, where $\kappa > 1$ and $\Delta(x, y)$ is the first place (in a suitable enumeration of $\mathbb{Z}$) where x and y differ.

Let $\sigma : X^+ \to X^+$ be the shift operator defined by
$\sigma(x)(i) = x(i + 1)$ for all $i \in \mathbb{Z}$.

Then (X^+, ρ) is called a full shift. It is a compact metric space, homeomorphic to the Cantor set, and $\sigma : X^+ \to X^+$ is a homeomorphism.

Subshift systems

Now let $X \subset {}^{\mathbb{Z}}A$ be closed both topologically wrt to ρ and wrt to σ. Then X is a subshift and (X, σ) is a standard subshift system.

Could we use a standard subshift system in our construction?

No. This is a well-known result.

But let d be any metric on A and consider the following metric D on ${}^{\mathbb{Z}}A$:

$$D(x, y) = d(x(\Delta(x, y)), y(\Delta(x, y)))\kappa^{-\Delta(x,y)},$$

where κ is sufficiently large so that this is actually a metric.

Then ρ and D generate the same topology on X^+, and when X is a subshift, we obtain a near-subshift system (X, σ) with near-subshift metric D.

Could we use a near-subshift system in our construction?

Some remarks on the terminology

While standard subshift systems have been extensively studied in the literature, the phrase "near-subshift" was coined by us. As far as we know, these systems had not been studied before. This is not surprising, as topologically they are the same as the standard ones.

The notion of a near-subshift metric is defined relative to an underlying metric d on A. The standard metric ρ is a special case for $d(a, b) = 1$ whenever $a \neq b$. In our construction of a near-subshift example we can use $A = \{0, 1, 2, 4, 5\}$ with $d(a, b) = |a - b|$ and $\kappa = 6$.

So, could we use a near-subshift system in our construction?

We actually started out trying to construct a near-subshift system (X, σ, D) with

$$\liminf_{T \to \infty} \frac{\ln sep(X, \varepsilon, D_T)}{T} < \limsup_{T \to \infty} \frac{\ln sep(X, \varepsilon, D_T)}{T}.$$

This didn't work. We always seemed to run into the same technical issue that we still don't fully understand.

Conjecture: There are no such examples.

Only quite recently did we seriously try to construct a counterexample for the spanning numbers.

Near-subshift systems can be counterexamples for the spanning numbers

Theorem

There exists a near-subshift system (Z, σ, D) *such that*

$$\liminf_{T \to \infty} \frac{\ln span(Z, 2, D_T)}{T} < \limsup_{T \to \infty} \frac{\ln span(Z, 2, D_T)}{T}.$$

Proof: We start by picking sequences $(T(n))_{n \in \mathbb{N}}$ and $(T^+(n))_{n \in \mathbb{N}}$ of positive integers such that the following inequalities hold:

$T(0) < T^+(0) < T(1) < T^+(1) < \dots < T(n) < T^+(n) < \dots$

The idea is to construct the system in such a way that

$$\limsup_{n \to \infty} \frac{\ln span(Z, 2, D_{T(n)})}{T(n)} \leq 0.6 \ln 2,$$

$$\liminf_{n \to \infty} \frac{\ln span(Z, 2, D_{T^+(n)})}{T^+(n)} \geq 0.9 \ln 2.$$

Outline of our construction

Our alphabet A and the metric d on A here will be given by:

$$A = \{0, 1, 2, 4, 5\},$$
$$d(a, b) = |a - b|.$$

We can define D based on d with $\kappa = 6$.

We will construct Z as the union of two disjoint subshifts $Z = X \dot{\cup} Y$ such that:

- For every $n \in \mathbb{N}$, we can find $(T(n), 2)$-spanning sets for Z of size at most $2^{0.6T(n)}$ that heavily relies on elements of X,
- For every $n > 0$, every $(T^+(n), 2)$-spanning sets for Z must contain a subset of Y of size at least $2^{0.9T^+(n)}$.

Terminology: A few words about words

- A word of length T in A is a finite sequence $\psi \in {}^{T}A$.
- For a subset $I \subset \mathbb{Z}$ and an integer t we let $I + t = \{i + t : \ i \in I\}$.
- We extend the shift operator σ to the family of words in A, so that when ψ is a word of length T, then $\sigma^t(\psi)$ is a function defined on the set $T - t$ (not necessarily a word) that takes the values $\sigma^t(\psi)(j) = \psi(j + t)$ for all $j \in T - t$.
- Let φ, ψ be words of lengths $T^- \leq T$, respectively. Then φ is a subword of ψ if there exists t with $0 \leq t \leq T - T^-$ such that $\varphi = \sigma^t(\psi) \restriction T^-$. Similarly, φ is a subword of $z \in {}^{\mathbb{Z}}A$ if there exists some $t \in \mathbb{Z}$ such that $\varphi = \sigma^t(z) \restriction T^-$.
- Let φ, ψ be words of lengths T, T', respectively. Then the concatenation $\varphi^\frown\psi$ is the word of length $T + T'$ such that $\varphi^\frown\psi \restriction T = \varphi$ and $\sigma^T(\varphi^\frown\psi) \restriction T' = \psi$.

Background: A few more words about words

A standard argument shows that $X \subseteq {}^{\mathbb{Z}}A$ is a subshift if, and only if, $X \neq \emptyset$ and there exists a (possibly empty) set W^- of forbidden words such that X consists of all elements of ${}^{\mathbb{Z}}A$ that do not contain any element of W^- as a subword.

Conversely, if X is any subshift, then we can define, for each $T \geq 0$, the set $W_T^+ = W_T^+(X)$ of all permitted words of length T as the set of all words that are subwords of some $x \in X$.

The size of this set is closely related to the spanning and separation numbers: Let (X, σ, D) be any near-subshift system with $X \subseteq {}^{\mathbb{Z}}A$. Let D be defined based on any metric d on A with $\min\{d(a,b) : a \neq b \in A\} = 1$. Then

$$\forall T > 0 \quad sep(X, 1, D_T) = span(X, 1, D_T) = |W_T^+(X)|.$$

Permitted words in our construction

In order to define the subshift X, we will construct the families W_T^X and W_T^Y of permitted words of length T for the subshifts X and Y, respectively.

We will do this in stages, by first constructing families $\mathcal{X}_n^-, \mathcal{Y}_n^-$ of words of length $T(n)$ and $\mathcal{X}_n, \mathcal{Y}_n$ of words of length $T^+(n)$. Then ψ will be a permitted word of length $T(n)$ for X iff ψ is a subword of a concatenation of two words in $\mathcal{X}_n^-$, and ψ will be a permitted word of length $T^+(n)$ for X iff ψ is a subword of a concatenation of two words in $\mathcal{X}_n$. Analogously for Y.

We will need another parameter, a fixed sequence $(\alpha_n)_{n\in\mathbb{N}}$ of reals such that

- $0.1 = \alpha_0 \geq \alpha_n > 2\alpha_{n+1} > 0$ for all $n \in \mathbb{N}$.
- $\lim_{n\to\infty} \alpha_n = 0$.

Permitted words for X

- The family $\mathcal{X}_0^- := \{\psi_0, \varphi_0\} \subset {}^{T(0)}\{1, 4\}$ consists of the function ψ_0 that is constantly equal to 1, and the function φ_0 that takes the value $\varphi_0(0) = 4$ and the value $\varphi_0(i) = 1$ for $0 < i < 10$.
- For any $n \in \mathbb{N}$, the set $\mathcal{X}_n$ consists of all words of length $T^+(n)$ that contain at most $\alpha_{n+1} T^+(n)$ occurrences of 4 and that are concatenations of successive words in $\mathcal{X}_n^-$.
- For any $n \in \mathbb{N}$, the set $\mathcal{X}_{n+1}^-$ consists of all words of length $T(n+1)$ that are concatenations of successive words in $\mathcal{X}_n$.

We will choose all numbers $T(n)$ so that they are multiples of 10. Then we get for every n:

$$span(X, 2, D_{T(n)}) \leq |W^X_{T(n)}| \leq 2^{0.1\,T(n)}.$$

Regular permitted words for Y

- The family $\mathcal{Y}_0^-$ consists of all functions $\psi \in {}^{T(0)}\{0,1,2,5\}$ such that $\psi(0) \in \{1,5\}$ and $\psi(i) \in \{0,2\}$ for $0 < i < 10$.
- All words in $\mathcal{Y}_0^-$ are considered regular words.
- For any $n \in \mathbb{N}$, the set $\mathcal{Y}_n$ consists of all words of length $T^+(n)$ that are of two kinds, regular or irregular:
  - Regular words $\varphi \in \mathcal{Y}_n$ that are concatenations of regular words in $\mathcal{Y}_n^-$.
- For any $n \in \mathbb{N}$, the set $\mathcal{Y}_{n+1}^-$ consists of all words of length $T(n+1)$ that are of two kinds, regular or irregular:
  - Regular words $\varphi \in \mathcal{Y}_{n+1}^-$ that are concatenations of regular words in $\mathcal{Y}_n$ and satisfy

$$\exists \psi \in \mathcal{X}_{n+1}^- \quad \{i \in T(n+1) : \varphi(i) = 5\} = \{i \in T(n+1) : \psi(i) = 4\}.$$

"Almost" $(T(n), 2)$-spanning sets of Z

Let us consider a regular word $\varphi \in \mathcal{Y}_n^-$. Then there exists a word $\psi \in \mathcal{X}_n^-$ such that $\psi(i) = 1$ whenever $\varphi(i) \in \{0, 2\}$, and $\psi(i) = 4$ whenever $\varphi(i) = 5$. Thus $D_{T(n)}(\varphi, \psi) = 1$.

In an ideal world where $W_{T(n)}^Y$ would consist only of regular words in $\mathcal{Y}_n^-$, this would imply that we could choose a $(T(n), 2)$-spanning set for X that consists entirely of elements of X, so that $span(Z, 2, D_{T(n)}) \leq 2^{0.1T(n)}$.

In the real world this doesn't quite work, because:

- We need to also deal with words φ that are concatenations of two words in $\mathcal{Y}_n^-$. This is straightforward but a bit technical and requires some additional conditions on the numbers $T(n)$ and $T^+(n)$. We will omit details here.
- Our definition of the subshift Y in terms of permitted rather than forbidden words requires us to also include some irregular words so that we do get a subshift.

Irregular words for Y

We omit the precise definition of irregular words here. They are a bit like irregular verbs in learning a natural language: A nuisance to deal with, but fortunately there aren't too many of them.

The size of the set $V^Y_{T(n)}$ of irregular words of length $T(n)$ satisfies:

$$|V^Y_{T(n)}| \leq 10T(n)2^{(T(n)+T^+(n-1))/2}.$$

Under suitable technical assumptions on $T(n)$ and $T^+(n-1)$, this allows us to include all irregular words, together with all words in $W^X_{T(n)}$ in any spanning set without exceeding the upper bound $2^{0.6T(n)}$ on its size.

How about $span(Z, 2, D_{T^+(n)})$?

Recall some of our definitions:

- For any $n \in \mathbb{N}$, the set $\mathcal{X}_n$ consists of all words of length $T^+(n)$ that contain at most $\alpha_{n+1} T^+(n)$ occurrences of 4 and that are concatenations of successive words in $\mathcal{X}_n^-$.
  - This implies a similar restriction on the number of occurrences of 4 in any word $\psi \in W^X_{T^+(n)}$:
    There can be at most $2\alpha_{n+1} T^+(n)$ of them.
- Regular words $\varphi \in \mathcal{Y}_n$ are concatenations of regular words in $\mathcal{Y}_n^-$.
  - This does imply a restriction on the number of occurrences of 5 in regular words in $\mathcal{Y}_n$:
    Only roughly $\alpha_n T^+(n)$ of them can occur.
    But this restriction is more lenient, since $\alpha_n > 2\alpha_{n+1}$.

We can show that there exists $y^* \in Y$ such that $D_{T^+(n)}(x, y^*) \geq 2$ whenever $x \in X$.

How about $span(Z, 2, D_{T^+(n)})$? Completed.

Consider y^* as on the previous slide and let $Y^* \subset Y$ be the set of all $y \in Y$ such that y and y^* take the value 5 on the same subset of $T^+(n)$.

Then also $D_{T^+(n)}(x, y) \geq 2$ whenever $x \notin Y^*$ and $y \in Y^*$.

There exists a $(T^+(n), 2)$-separated subset $Y^{**} \subseteq Y^*$ of size $|Y^{**}| \geq 2^{0.9T^+(n)}$.

Every $(T^+(n), 2)$-spanning subset of Z must contain such a set.

It follows that $span(Z, 2, D_{T^+(n)}) \geq 2^{0.9T^+(n)}$.

This completes the proof of our theorem. $\square$

Recall our main theorem

Theorem

There exists a minimal system (X, F) with a metric D on X such that for some $\varepsilon > 0$ we have:

$$\liminf_{T\to\infty} \frac{\ln sep(X,\varepsilon,D_T)}{T} < \limsup_{T\to\infty} \frac{\ln sep(X,\varepsilon,D_T)}{T},$$

$$\liminf_{T\to\infty} \frac{\ln span(X,\varepsilon,D_T)}{T} < \limsup_{T\to\infty} \frac{\ln span(X,\varepsilon,D_T)}{T}.$$

To make the exposition more transparent, we will defer metaphysics, the definitions of X, D, F, until much later and focus on how certain features of these objects make our system behave in the way asserted by the theorem.

The first ingredient of our construction is already familiar

We will need positive integers $T(n)$ and $T^+(n)$ with

$$1 < T(0) < T^+(0) < \cdots < T(n) < T^+(n) < T(n+1) < T^+(n+1) < \ldots$$

These sequences will be parameters in our construction and will satisfy a number of ugly but otherwise unproblematic technical conditions.

We will then prove, for a suitably chosen parameter $\varepsilon > 0$, the following inequalities:

$$\liminf_{n\to\infty} \frac{\ln \mathit{sep}(X, \varepsilon, D_{T^+(n)})}{T^+(n)} > \limsup_{n\to\infty} \frac{\ln \mathit{sep}(X, \varepsilon, D_{2T(n)})}{2T(n)}.$$

$$\liminf_{n\to\infty} \frac{\ln \mathit{span}(X, \varepsilon, D_{T^+(n)})}{T^+(n)} > \limsup_{n\to\infty} \frac{\ln \mathit{span}(X, \varepsilon, D_{2T(n)})}{2T(n)}.$$

A little more precisely ...

More precisely, we will construct things so that for some $\lambda < 0.9$:

$$\liminf_{n\to\infty} \frac{\ln sep(X, \varepsilon, D_{T^+(n)})}{T^+(n)} \geq 0.9 \ln 2.$$

$$\limsup_{n\to\infty} \frac{\ln sep(X, \varepsilon, D_{2T(n)})}{2T(n)} \leq \lambda \ln 2.$$

$$\liminf_{n\to\infty} \frac{\ln span(X, \varepsilon, D_{T^+(n)})}{T^+(n)} \geq 0.9 \ln 2$$

$$\limsup_{n\to\infty} \frac{\ln span(X, \varepsilon, D_{2T(n)})}{2T(n)} \leq \lambda \ln 2.$$

What do we need to know right now about X and F?

Our construction uses a lot of words in the alphabet $\{0, 1\}$ together with a counter that keeps track of their origins.

We construct (X, F) so that with every $x \in X$ and $n \in \mathbb{N}$ we can associate an integer

$k_n(x) \in T^+(n) := \{0, \ldots, T^+(n) - 1\}$ and a word

$\psi_n \in {}^{T^+(n)}\{0, 1\}$ of length $T^+(n)$ so that for all $x \in X$ and $n \in \mathbb{N}$:

(Fk) $k_n(F(x)) = k_n(x) + 1 \mod T^+(n)$

(Fx) If $k_n(x) < T^+(n) - 1$, then $\psi_n(x) = \psi_n(F(x))$.

We will require that $\psi_n \in \mathcal{X}_n$, where $\mathcal{X}_n$ is a subset of ${}^{T^+(n)}\{0, 1\}$ that satisfies certain conditions, in particular, is of size at least $|\mathcal{X}_n| \geq 2^{0.9T^+(n)}$.

The key property of the metric D

The metric D on X is then defined in terms of a sequence of conditions $Cond_n$ on triplets (φ, ψ, k), where φ, ψ are words of length $T^+(n)$ and $k \in T^+(n)$, such that

(Dε) $D(x, x') \geq \varepsilon$

if, and only if,

$$\forall n \in \mathbb{N} \;\; (k_n(x) \neq k_n(x') \;\vee\; Cond_n(\psi_n(x), \psi_n(x'), k_n(x))).$$

Keeping some separation numbers small

We want to assure that for some fixed $\lambda < 0.9$ and all $n \in \mathbb{N}$

$$span(X, \varepsilon, D_{2T(n)}) \leq sep(X, \varepsilon, D_{2T(n)}) < 2^{\lambda 2T(n)}, \text{ that is,}$$

$$\frac{\ln span(X, \varepsilon, D_{2T(n)})}{2T(n)} \leq \frac{\ln sep(X, \varepsilon, D_{2T(n)})}{2T(n)} < \lambda \ln 2.$$

As long as our parameter satisfy the condition $T^+(n) \leq T(n)2^{0.01T(n)}$, it suffices to show that

$$sep(X, \varepsilon, D_{2T(n)}) \leq T^+(n)2^{1.75T(n)},$$

since for all $\lambda > 0.88$ and all sufficiently large n we will then have

$$sep(X, \varepsilon, D_{2T(n)}) \leq T^+(n)2^{1.75T(n)} \leq T(n)2^{1.76T(n)} < 2^{\lambda 2T(n)}.$$

We need to design the conditions $Cond_n$ in the definition of the metric D so that this works.

Question: What could we use for achieving this goal?

Colorings

Let $C(n) = T^+(n)/T(n)$. We partition the interval $T^+(n) := \{0, \dots, T^+(n) - 1\}$ into consecutive subintervals I_j^n of length $T(n)$ each, where j ranges from 1 to $C(n)$.

For a subset $S \subseteq {}^{T^+(n)}\{0, 1\}$, let $[S]^2$ denote the set of all unordered pairs $\{\varphi, \psi\}$ of different words from S.

Moreover, let $[C(n)] = \{1, 2, \dots, C(n)\}$.

We will consider colorings $c : [S]^2 \to [C(n)]$,

where $S \subseteq {}^{T^+(n)}\{0, 1\}$ for some $n \in \mathbb{N}$.

A subset $S^- \subseteq S$ will be called ≤ 2-*chromatic* for c if the restriction of c to $[S^-]^2$ takes at most 2 values from the set $[C(n)]$.

The conditions $Cond_n$

Our construction uses a sequence $(c_n)_{n \in \mathbb{N}}$ of suitable colorings as parameters.

In particular, the domain of c_n will be $[\mathcal{X}_n]^2$, and the colorings c_n will not admit ≤ 2-chromatic subsets of size $\geq 2^{0.75\,T(n)}$.

We prove existence of suitable colorings using the probabilistic method.

The conditions $Cond_n$ for the definition of the metric D will then take the form:

$$Cond_n(\varphi, \psi, k_n) \Leftrightarrow (\varphi(k_n) \neq \psi(k_n) \ \&\ k_n \in I^n_{c_n(\varphi,\psi)}).$$

Let us illustrate how this works for keeping some separation numbers small.

Consider a $(2T, \varepsilon)$-separated subset $A \subset X$ such that $k_n(x) = \tau$ for all $x \in A$. Since there are only $T^+(n)$ possible choices for τ, for the inequality $sep(X, \varepsilon, D_{2T(n)}) \leq T^+(n)2^{1.75T(n)}$ it will suffice to show that such A has size at most $2^{1.75T(n)}$.

Let us focus in this illustration on the simplest case where τ with $0 \leq \tau < \tau + 2T(n) < T^+(n)$.

Then by (Fk) and (Fx), for all $t \in 2T$ and $x \in A$ we have $k_n(F^t(x)) = \tau + t$ and $\psi_n(F^t(x)) = \psi_n(x)$.

Since A was assumed to be $(2T, \varepsilon)$-separated, for every $x \neq x' \in A$ there exists $t < 2T$ such that $D(F^t(x), F^t(x')) \geq \varepsilon$.

By (Dε), this implies that for every $x \neq x' \in A$ there exists $t < 2T$ such that $Cond_n(\psi_n(F^t(x)), \psi_n(F^t(x')), \tau + t)$ holds. In view of (31) this now implies that for every $x \neq x' \in A$ there exists $t < 2T$ such that $\psi_n(x)(t + \tau) \neq \psi_n(x')(t + \tau)$ and $\tau + t \in I^n_{c_n(\psi_n(x), \psi_n(x'))}$.

Small separation numbers, completed

Now assume towards a contradiction that A has more than $2^{1.75T(n)}$ elements and for every $x \neq x' \in A$ there exists $t < 2T$ such that $\psi_n(x)(t+\tau) \neq \psi_n(x')(t+\tau)$ and $\tau + t \in I^n_{c_n(\psi_n(x),\psi_n(x'))}$.

Then there exists a subset $A^- \subset A$ of size $> 2^{0.75T(n)}$ such that $\psi_n(x), \psi_n(x')$ take the same values on $\tau + t$ for all $\tau + T(n) \leq t < \tau + 2T(n)$, so that for $x \neq x' \in A^-$ there exists $t < T$ such that $\psi_n(x)(t+\tau) \neq \psi_n(x')(t+\tau)$ and $\tau + t \in I^n_{c_n(\psi_n(x),\psi_n(x'))}$.

But note that since each of the intervals I^n_j has length T, the values $\tau, \tau+1, \dots, \tau+T-1$ can belong to at most two of the intervals I^n_j.

It follows that the set $S := \{\psi_n(x) : \ x \in A^-\}$ has the same size as the set A^- and that the restriction of c_n to the set $[S]^2$ takes at most 2 values from the set $C(n)$. Thus the size of A^- cannot exceed the maximal size of a $\leq$ 2-chromatic subset of c_n. This contradicts our choices of A^- and c_n.

Large $(T^+(n), \varepsilon)$-separated subsets of X

For each n, we construct a set W_n such that:

$\forall \psi \in \mathcal{X}_n \, \exists x \in W_n \;\; \psi_n(x) = \psi$. Thus $|W_n| \geq 2^{0.9T^+(n)}$.

$\forall m \in \mathbb{N} \, \forall x \in W_n \;\; k_m(x) = 0$.

$\forall x \neq x' \in W_n \, \exists t \in T^+(n) \, \forall m \in \mathbb{N} \; Cond_m(\psi_m(F^t(x)), \psi_m(F^t(x)), k_m))$, where $k_m = t \mod T^+(m)$.

By Property (Fk), for $x, x' \in W_n$ we then have
$k_m(F^t(x)) = k_m(F^t(x')) = t \mod T^+(m)$ for all m,
and Property (Dε) will imply that

$$sep(X, \varepsilon, D_{T^+(n)}) \geq 2^{0.9T^+(n)}, \text{ so that } \frac{\ln sep(X, \varepsilon, D_{T^+(n)})}{T^+(n)} \geq 0.9 \ln 2.$$

We get the analogue of the last inequality for spanning numbers by showing that for any $x' \in X$ the inequality $D_{T^+(n)}(x', x) < \varepsilon$ can hold for at most one $x \in W_n$.

How do we get large spanning numbers?

We will not define the sets W_n here. Instead, as we go, we will list additional properties that we need to get large spanning numbers.

Fix $n \in \mathbb{N}$ and let $x' \in X$. We will need the following property:

(X1) $x \neq x' \in W_n \quad \Longrightarrow \quad \psi_n(x) \neq \psi_n(x')$.

Then there exists at most one $x \in W_n$ with $\psi_n(x) = \psi_n(x')$.

So assume $x \in W_n$ is such that $\psi_n(x) \neq \psi_n(x')$. It suffices to show that $D_{T^+(n)}(x, x') \geq \varepsilon$.

For that we need some $t \in T^+(n)$ with $D(F^t(x), F^t(x')) \geq \varepsilon$.

Note that we need to find such a t that works at all levels m simultaneously.

Note also that the first clause of (Dε) essentially tells us that levels m with $k_m(x) \neq k_m(x')$ are unproblematic, so we will restrict our illustration here to the most interesting situation where $k_m(x') = k_m(x) = 0$ for all m. Then we also have $k_m(F^t(x')) = k_m(F^t(x))$ for all m and all t by Property (Fk).

How do we get large spanning numbers? Continued.

Now we need some t with $t = k_n(F^t(x)) = k_n(F^t(x'))$ such that

- $\psi_n(x)(t) \neq \psi_n(x')(t)$.
- $t \in I^n_{c_n(\psi_n(x),\psi_n(x'))}$.

We get such t from the following property of our colorings:

(CD) $\psi \neq \varphi \in \mathcal{X}_n$, then there exists at least one $k \in I^n_{c_n(\psi,\varphi)}$ such that $\psi(k) \neq \varphi(k)$.

Note then any $t \in I^n_{c_n(\psi_n(x),\psi_n(x'))}$ with $\psi_n(x)(t) \neq \psi_n(x')(t)$ will work for satisfying the clause in (Dε) that deals with level n.

We still need to find a t, in this same interval, that covers the clauses of (Dε) that deal with levels $m < n$ and with levels $m > n$.

Dealing with levels $m < n$

We need the following properties:

(TC) $T^+(n)$ is an integer multiple of $T^+(m)$ for $m < n$.

(XC) For $x \in X$, $m < n$, and $\ell T^+(m) < T^+(n)$, the restriction of $\psi_n(x)$ to $\{\ell T^+(m), \dots, (\ell+1)T^+(m) - 1\}$ is equal to $\psi_m(F^{\ell T^+(m)}(x))$.

Assume by induction that

- $\psi_n(x) \restriction I^n_{c_n(\psi_n(x),\psi_n(x'))} \neq \psi_n(x') \restriction I^n_{c_n(\psi_n(x),\psi_n(x'))}$.

Let $m = n - 1$. Then we find $\tau = \ell T^+(m)$ with

- $\psi_m(F^\tau(x)) \neq \psi_m(F^\tau(x'))$.
- $k_m(x) = k_m(x') = 0$.

Now we can find $t' \in I^m_{c_m(\psi_m(F^\tau(x)),\psi_m(F^\tau(x')))}$ as on the previous slide, and $t = (c_n(\psi_n(x), \psi_n(x')) - 1)C(n) + \ell T^+(m) + t'$ will work at both levels m and n.

By iterating this argument, we find t that works at al levels $m \leq n$.

Dealing with levels $m > n$

The t we have found so far is an element of I_1^m for $m > n$.

This t will automatically work at levels $m > n$ if we have the following properties:

(WS) For $x \in W_m$ and $m < n$, the function $\psi_n(x)$ is "special."

(C1) If $\psi \neq \varphi \in \mathcal{X}_n$ and at least one of ψ, φ is special, then $c_n(\psi, \varphi) = 1$.

So what about that definition of (X, F)?

We let X consist of pairs $x = (y, \kappa)$, where $y \in {}^{\mathbb{Z}}\{0, 1\}$ and $\kappa : \mathbb{N} \to \mathbb{N}$ is a function such that $\kappa(n) \in T^+(n)$ for all n.

X will not consist of all such pairs, only of the pairs that are allowed by our conditions.

$F(y, \kappa) = (\sigma(y), \kappa \oplus 1)$, where:

- σ is the usual shift operator,
- $(\kappa \oplus 1)(n) = \kappa(n) + 1 \mod T^+(n)$ for all $n \in \mathbb{N}$.

We let $k_n(x) = \kappa(n)$ and $\psi_n(x) = \sigma^{-\kappa(n)}(y) \restriction T^+(n)$.

We then define the sets W_n so that they satisfy all relevant conditions.

And what about that definition of D?

We define D so that:

- (Dε) holds, and
- For each $x = (y, \kappa) \in X$ and $N \in \mathbb{N}$ the sets $U_N(x)$ of all $x' = (y', \kappa')$ such that $\forall i \ \ |i| < N \implies y(i) = y'(i) \ \& \ \kappa(i) = \kappa'(i)$ are open and form a basis for the topology.

The proof that X is compact and F a homeomorphism then becomes completely standard and independent from the more technical parts of the argument.

What about minimality?

Minimality of (X, F) is assured in our construction by the following properties:

(FC) When $x \in X$ and $m < n$, then $k_m(x) = k_n(x) \mod T^+(m)$.

(XM) For all $m < n$ and all $\psi, \varphi \in \mathcal{X}_m$, every $\psi^+ \in \mathcal{X}_n$ contains a block of the form $\psi^\frown\varphi$.

Two open problems

Question 1: Does there exist a near-subshift system (Z, σ, D) such that

$$\liminf_{T\to\infty} \frac{\ln sep(Z, 2, D_T)}{T} < \limsup_{T\to\infty} \frac{\ln sep(Z, 2, D_T)}{T} ?$$

Question 2: Does there exists a minimal near-subshift system (Z, σ, D) such that

$$\liminf_{T\to\infty} \frac{\ln span(Z, 2, D_T)}{T} < \limsup_{T\to\infty} \frac{\ln span(Z, 2, D_T)}{T} ?$$

We know that examples as above exist if, and only if:

- There exist corresponding examples with X being a subset of $\mathbb{R}^n$ with the metric induced by the sup-norm.
- There exist such examples with the alphabet $A = \{0, 1, 2\}$ and D being based on the usual metric d given by $d(a, b) = |a - b|$.