Computational Aspects of Symbolic Dynamics

Part IV: SFT vs effective subshifts

E. Jeandel

LORIA (Nancy, France)

Introduction

- 2D SFTs and effective subshifts share many properties, as pointed out by the Aubrun-Sablik theorem.
- In this last part, we will speak about *differences* between SFTs and effective subshifts.
  - (Kolmogorov) Complexity
  - Periodic points

Complexity

- The complexity $p_x(n)$ of a point x is the number of different patterns of size n.

Theorem (Durand-Levin-Shen [DLS08])

Every k-dimensional SFT contains a point x of complexity $2^{O(n^{k-1})}$.
The bound is tight.
There are k-dimensional effective shifts where every point is of complexity $2^{\Omega(n^k)}$

The theorem is still valid for sofic shifts rather than SFT (obvious). The second part proves that the bound is tight in the first part, by Aubrun-Sablik.

Kolmogorov complexity

- The Kolmogorov complexity $K(w)$ of a word w is the size of the smallest program that outputs w.
- The exact definition does not really matter here.

If every pattern of size n of x is of Kolmogorov complexity less than c, then $p_x(n) \leq 2^c$

(there are at most 2^c programs of size less than c)

If x is computable, then for every pattern w of size n, $K(w) \leq \log p_x(n) + O(\log n)$.

The program contains an integer $i \leq c$ and outputs the i-th pattern of size n of x.

The theorem

Theorem (Durand-Levin-Shen [DLS08])

Every 2-dimensional SFT contains a point x so that for all patterns w of size $n \times n$, $K(w) = O(n)$.

How to prove it ?

Lemma

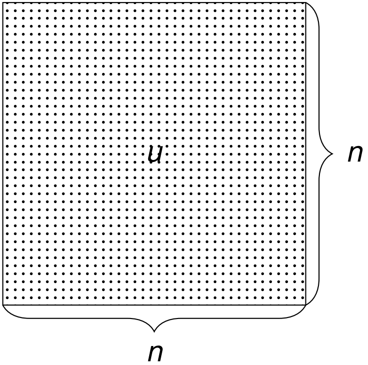
For every n, there exists a globally admissible pattern w of size $n \times n$ and of complexity $K(w) \leq cn + O(1)$ (where c depends only on the SFT)

Proof :

Lemma

For every n, there exists a globally admissible pattern w of size $n \times n$ and of complexity $K(w) \leq cn + O(1)$ (where c depends only on the SFT)

Proof :

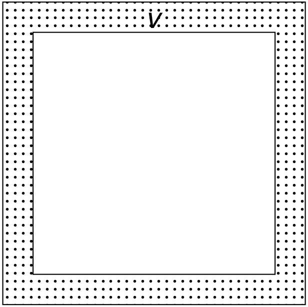


Start from any globally admissible pattern u of size $n \times n$

Lemma

For every n, there exists a globally admissible pattern w of size $n \times n$ and of complexity $K(w) \leq cn + O(1)$ (where c depends only on the SFT)

Proof :

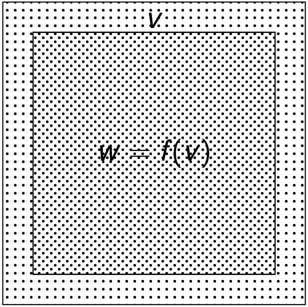


Let v be the boundary of u.
It is a word of size $4n - 4$.

Lemma

For every n, there exists a globally admissible pattern w of size $n \times n$ and of complexity $K(w) \leq cn + O(1)$ (where c depends only on the SFT)

Proof :



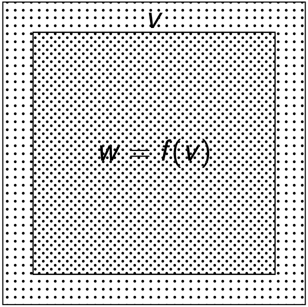
Given only v, we can find algorithmically a globally admissible w.

(Use your favorite deterministic algorithm to fill the inside with valid letters)

Lemma

For every n, there exists a globally admissible pattern w of size $n \times n$ and of complexity $K(w) \leq cn + O(1)$ (where c depends only on the SFT)

Proof :



Clearly $K(w) \leq K(v) + O(1)$.

And $K(v) \leq (4n - 4) \log A + O(1)$

That ends the lemma. However subpatterns of w might be of big complexity.

Lemma

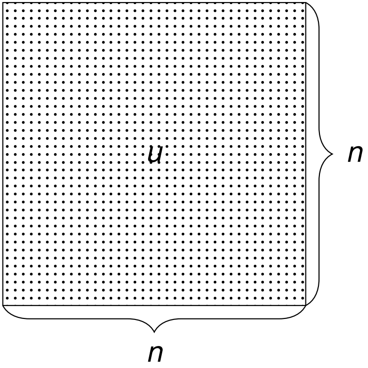
For every n, there exists a globally admissible pattern w of size $n \times n$ so that for all subpatterns p of w, $K(p) \leq c|p| + O(1)$ (where c depends only on the SFT)

Proof :

Lemma

For every n, there exists a globally admissible pattern w of size $n \times n$ so that for all subpatterns p of w, $K(p) \leq c|p| + O(1)$ (where c depends only on the SFT)

Proof :

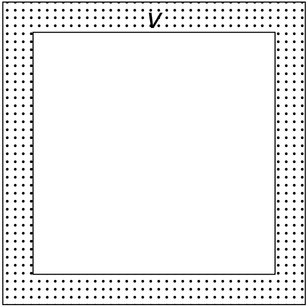


Start from any globally admissible pattern u of size $n \times n$

Lemma

For every n, there exists a globally admissible pattern w of size $n \times n$ so that for all subpatterns p of w, $K(p) \leq c|p| + O(1)$ (where c depends only on the SFT)

Proof :

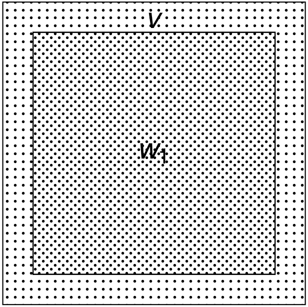


Let v be the boundary of u.
It is a word of size $4n - 4$.

Lemma

For every n, there exists a globally admissible pattern w of size $n \times n$ so that for all subpatterns p of w, $K(p) \leq c|p| + O(1)$ (where c depends only on the SFT)

Proof :

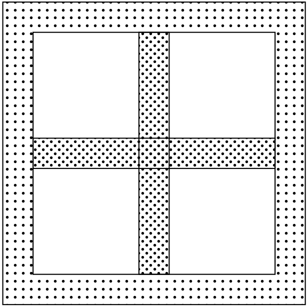


Find algorithmically some globally admissible w_1

Lemma

For every n, there exists a globally admissible pattern w of size $n \times n$ so that for all subpatterns p of w, $K(p) \leq c|p| + O(1)$ (where c depends only on the SFT)

Proof :

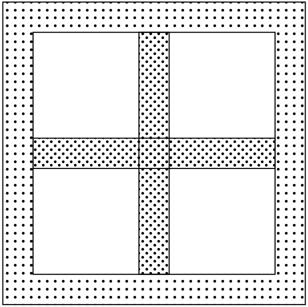


Keep only the backbone

Lemma

For every n, there exists a globally admissible pattern w of size $n \times n$ so that for all subpatterns p of w, $K(p) \leq c|p| + O(1)$ (where c depends only on the SFT)

Proof :

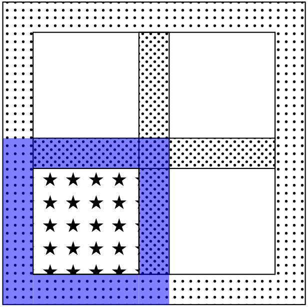


Call recursively f on all four boundaries

Lemma

For every n, there exists a globally admissible pattern w of size $n \times n$ *so that for all subpatterns p of w,* $K(p) \leq c|p| + O(1)$ *(where c depends only on the SFT)*

Proof :

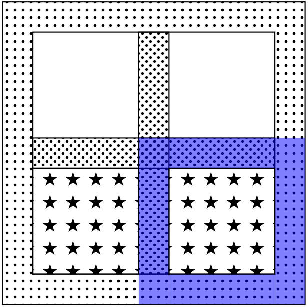


Call recursively f on all four boundaries

Lemma

For every n, there exists a globally admissible pattern w of size $n \times n$ so that for all subpatterns p of w, $K(p) \leq c|p| + O(1)$ (where c depends only on the SFT)

Proof :

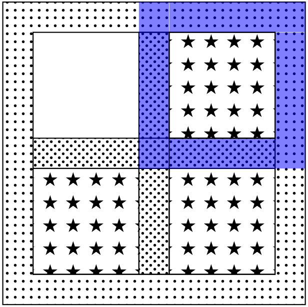


Call recursively f on all four boundaries

Lemma

For every n, there exists a globally admissible pattern w of size $n \times n$ so that for all subpatterns p of w, $K(p) \leq c|p| + O(1)$ (where c depends only on the SFT)

Proof :

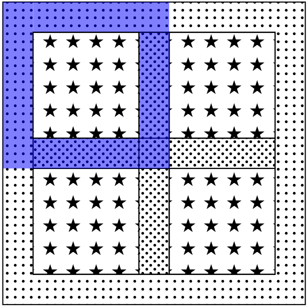


Call recursively f on all four boundaries

Lemma

For every n, there exists a globally admissible pattern w of size $n \times n$ so that for all subpatterns p of w, $K(p) \leq c|p| + O(1)$ (where c depends only on the SFT)

Proof :

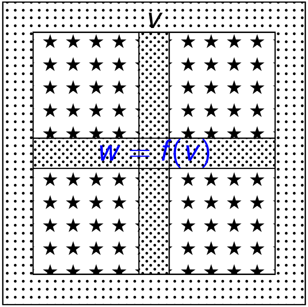


Call recursively f on all four boundaries

Lemma

For every n, there exists a globally admissible pattern w of size $n \times n$ so that for all subpatterns p of w, $K(p) \leq c|p| + O(1)$ (where c depends only on the SFT)

Proof :



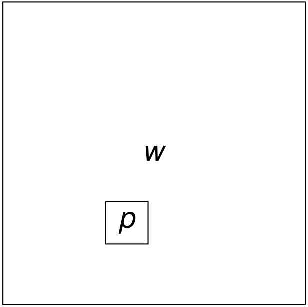
Call recursively f on all four boundaries

Let's call $w = f(v)$ the result

Lemma

For every n, there exists a globally admissible pattern w of size $n \times n$ so that for all subpatterns p of w, $K(p) \leq c|p| + O(1)$ (where c depends only on the SFT)

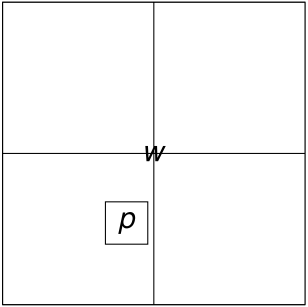
Let p be a subpattern of w of size $k \times k$



Lemma

For every n, there exists a globally admissible pattern w of size $n \times n$ so that for all subpatterns p of w, $K(p) \leq c|p| + O(1)$ (where c depends only on the SFT)

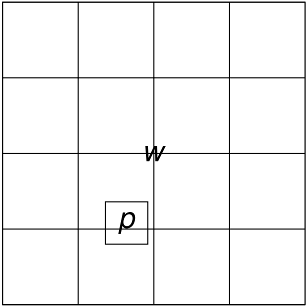
Let p be a subpattern of w of size $k \times k$
By construction, w contains a grid of patterns of complexity $O(n/2^i)$ for all i



Lemma

For every n, there exists a globally admissible pattern w of size $n \times n$ so that for all subpatterns p of w, $K(p) \leq c|p| + O(1)$ (where c depends only on the SFT)

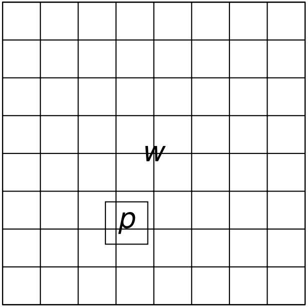
Let p be a subpattern of w of size $k \times k$
By construction, w contains a grid of patterns of complexity $O(n/2^i)$ for all i



Lemma

For every n, there exists a globally admissible pattern w of size $n \times n$ so that for all subpatterns p of w, $K(p) \leq c|p| + O(1)$ (where c depends only on the SFT)

Let p be a subpattern of w of size $k \times k$
By construction, w contains a grid of patterns of complexity $O(n/2^i)$ for all i



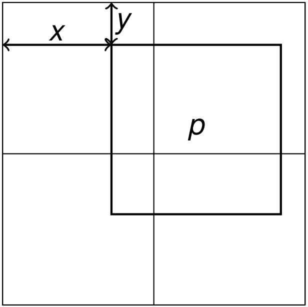
Lemma

For every n, there exists a globally admissible pattern w of size $n \times n$ so that for all subpatterns p of w, $K(p) \leq c|p| + O(1)$ (where c depends only on the SFT)

Let p be a subpattern of w of size $k \times k$
The four patterns are of size $k' \leq 2k$ for some k'.
To specify p, it is sufficient :

- To describe the four patterns. This is linear in their size by construction
- To describe x and y. This is logarithmic in x and y.
- To describe the size of p. This is logarithmic in k.

$K(p) \leq 4 * O(k') + 2 \log k' + \log k = O(k)$



The Theorem

Theorem (Durand-Levin-Shen [DLS08])

Every 2-dimensional SFT contains a point x so that for all patterns w of size $n \times n$, $K(w) = O(n)$.

For some c, we can build arbitrary large patterns w so that $K(p) \leq c.|p|$ for every pattern p in w. The result follows by compactness.

The bound is tight

Theorem (Durand-Levin-Shen [DLS08])

There exists a k-dimensional effective shift so that for all points x, for all patterns w of size $n \times n$, $K(w) = \Omega(n^k)$.

Proof in 1D : Let S_p be the subshift that forbids all patterns w so that $K(w) < |w|/4$ for $|w| \geq p$

- S_p is effective.
- S_p is nonempty if p is big enough. Indeed, fix $c = \frac{1}{\sqrt{2}}$. Then

$$\sum 2^{k/4} c^k < \infty$$

So that $\sum_{k \geq p} 2^{k/4} c^k < 2c - 1$ for some p.

Another difference between SFTs and effective shifts

- Periodic points are another classical conjugacy invariant
- In dimension 1, a point x is periodic of period p if

$$\{q | \forall i, x_{q+i} = x_i\} = p\mathbb{Z}$$

For a 1D SFT S, the set $L = \{p | \exists x \text{ of period } p\}$ is well understood

- L is always nonempty if S is nonemtpy.
- L is always a semilinear set (finite unions of linear sets, of the form $c + b\mathbb{N}$)
- comes for automata theory : Unary regular languages are semilinear

Effective shifts

What about effective shifts ?

- There is a semialgorithm that halts if p is not a period
- Test, for all (primitive) words w of size n, if ${}^{\omega}w^{\omega} \notin S$
- ${}^{\omega}w^{\omega} \notin S$ iff it contains one of the forbidden patterns (hence we have only a semi algorithm)

Recall that a set L of integers is co-recursively-enumerable if

- There is an algorithm that halts on input n iff $n \notin L$
- Equivalently, there is an algorithm f that enumerates its complement, ${}^{c}L = \{f(n), n \in \mathbb{N}\}$

The set of periods of a (1D) effective shift is co-recursively-enumerable.

Effective shifts

Theorem

For every co-recursively enumerable set L, there exists a (1D) effective shift S so that the set of periods of S is exactly L.

Proof

We start from S_1 over $\{0, 1, \sharp\}$ that forbids :

- $\mathcal{F}_1 = \{xyxyx, x \in \{a, b\}^+, y \in \{a, b\}^\star\}$
- $\mathcal{F}_2 = \{\sharp u \sharp v \sharp, u \neq v\}$

S_1 is clearly effective.
What are periodic points for S_1 ?

- A periodic point of S_1 must contain a $\sharp$ symbol, due to $\mathcal{F}_1$.
- A periodic point of S_1 must be of the form $\dots \sharp w \sharp w \sharp w \dots$, due to $\mathcal{F}_2$.
- The period of a word is exactly the distance between two consecutive $\sharp$ symbols.

Furthermore, there exist words in S_1 of any period p : consider $\dots \sharp w \sharp w \dots$ where w is the first $p - 1$ letters of the Thue-Morse word.

Proof

Now start from L a co-recursively-enumerable set, given by an enumeration f of its complement

$$\mathcal{F}_3 = \{\sharp w \sharp, \exists n \in \mathbb{N}, |w| = f(n) - 1\}$$

(we forbid two $\sharp$ symbols to be at a distance $m \notin L$)

Then the set of periods of S forbidding $\mathcal{F}_1, \mathcal{F}_2, \mathcal{F}_3$ is exactly L.

Higher dimensions

What happens in higher dimensions ?

- First we have to define what is the period of a point p
- The period of a point x is a lattice

$$\Gamma_x = \{q \in \mathbb{Z}^d | \forall i, x_{q+i} = x_i\}$$

- Hard to deal with.
- The period of a point x is p if

$$\Gamma_x = p\mathbb{Z}^d$$

Multidimensional effective shifts

Theorem

The set of periods of any d-dimensional sofic (resp. effective) shift may be any co-recursively enumerable set.

Uses a combination of the same idea and Aubrun-Sablik's result.

What about SFTs ?

Multidimensional SFTs

Let S be a SFT. To test if p is a period :

- Test, for all patterns w of size p, whether a periodic filling by w is in S
- This can be done, for a given pattern, in polynomial time.

This gives an exponential time algorithm, hence the set of periods is a *computable* set.

Multidimensional SFTs

We can say better. Recall that **NP** is the class of problems that are solvable in nondeterministic polynomial time.

If $L \subseteq \mathbb{N}$ is the set of periods of a d-dimensional SFT, then $\{a^n | n \in L\} \in$ **NP**

($\{a^n | n \in L\}$ is L coded in unary)

Multidimensional SFTs

Theorem (J., Vanier [JV10])

If $\{a^n | n \in L\} \in \mathbf{NP}$*, then there exists a multidimensional SFT with L as a set of periods.*

Sketch of the proof

Let's forget about **NP** for a moment.
We will try to build a d-dimensional SFT S that works as the effective subshifts S_1 we saw earlier

- All periodic points are of the form

$$
\begin{array}{l}
\dots \quad \dots \quad \dots \quad \dots \\
\dots\sharp\sharp\sharp\sharp\sharp\sharp\sharp\sharp\sharp\sharp\sharp\sharp\sharp\sharp\sharp\dots \\
\dots\sharp \qquad \sharp \qquad \sharp\dots \\
\dots\sharp \quad w \quad \sharp \quad w \quad \sharp\dots \\
\dots\sharp \qquad \sharp \qquad \sharp\dots \\
\dots\sharp\sharp\sharp\sharp\sharp\sharp\sharp\sharp\sharp\sharp\sharp\sharp\sharp\sharp\sharp\dots \\
\dots\sharp \qquad \sharp \qquad \sharp\dots \\
\dots\sharp \quad w \quad \sharp \quad w \quad \sharp\dots \\
\dots\sharp \qquad \sharp \qquad \sharp\dots \\
\dots\sharp\sharp\sharp\sharp\sharp\sharp\sharp\sharp\sharp\sharp\sharp\sharp\sharp\sharp\sharp\dots
\end{array}
$$

where the period is exactly the distance between two $\sharp$ symbols

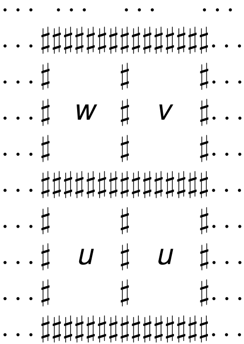
Idea of the proof (first step)

- Start from an *aperiodic* SFT, and add three symbols $+, -, |$.
- Add rules that force the three symbols to take the place of $\sharp$
- As the original SFT is aperiodic, any periodic point must have one of these symbols, and thus the square structure.
  - A lot of details to do that exactly. We don't want to have a periodic point with the symbols $|$ but no $-, +$.

This takes care of the $\mathcal{F}_1$ part of the SFT.

Idea of the proof (first step)

One problem remains. The point might be of the form



Idea of the proof (first step)

The aperiodic background must be the same in all the squares.

- Solution : Do not start from any aperiodic SFT, but for a *deterministic* aperiodic SFT.
  - In such a SFT (e.g. the Kari-Papasoglu SFT [KP99]), the filling of the square is entirely set by the symbols on the borders
  - Just ensure that two consecutive squares have the same border (easily done by adding an additional layer)

We are done for the first step. This generalizes in any dimension.

Idea of the proof (second step)

Now we need to keep only the squares of size n where a^n is accepted by some polynomial time nondeterministic Turing machine.

- The TM works in space-time n^k for some k
- To know if n is accepted can be done within a space-time diagram of size n^k.
- To know if n is accepted can be done in a square of size n^k.
- Problem : we only have a square of size n.

Idea of the proof (second step)

Do the construction in dimension $2k$ rather than in dimension 2.

- We now have a hypercube of size n, that can contain n^{2k} symbols.
- Enough to encode all the computation.
- *Fold* the $n^k \times n^k$ square into an hypercube of size n.

The proof is done, up to technical details (we also need to ensure that all hypercubes contain the same computation)

Bibliography I

- Bruno Durand, Leonid A. Levin, and Alexander Shen, *Complex tilings.*, Journal of Symbolic Logic **73** (2008), no. 2, 593–613.
- Emmanuel Jeandel and Pascal Vanier, *Periodicity in Tilings*, Developments in Language Theory (DLT), Lecture Notes in Computer Science, vol. 6224, Springer, 2010, pp. 243–254.
- Jarkko Kari and P. Papasoglu, *Deterministic Aperiodic Tile Sets*, Geometric And Functional Analysis **9** (1999), 353–369.