Trees in positive entropy subshifts

Ville Salo[*]
vosalo@utu.fi

April 29, 2022

Abstract

I give a simple proof for the fact that positive entropy subshifts contain infinite binary trees where branching happens synchronously in each branch, and that the branching times form a set with positive lower asymptotic density.

The proof is trivial! Just view it as a
rational metric space
whose elements are
countable combinatorial games

1 Introduction

In topological dynamics, the topological entropy of a dynamical system measures the information in orbits, by counting the exponential growth rate of different partial orbits with a certain of accuracy. One example of a dynamical system is a *subshift*, a dynamical system whose points are infinite words over a finite alphabet A, and the dynamics is the left shift. One way entropy can arise in a subshift is when, somewhere in partial orbits, we literally see all distinct words A^n up to n. One may ask to what extent entropy always arises from such "free choices".

Two formal notions that most directly capture this idea are the *independence entropy* of a subshift and *independence sets*. The former measures how much entropy comes from "set-theoretic hyperrectangles", namely sets of the form $A_1 \times A_2 \times \cdots A_n$ which are contained in the language of the subshift as $n \to \infty$, where $A_i \subseteq A$ are any subsets of the alphabet and we consider consecutive letters only, see [12], (p. 303) for the precise definition. Positive entropy does not imply positive independence entropy (see Proposition 1).

Independence sets, on the other hand, are defined similarly in [10, 3], but we allow gaps between the positions where the prescribed letters appear, and there is no constraint on the words seen in the gaps between the positions. More precisely, in the case of a subshift X with a binary alphabet $A = \{0, 1\}$, an *independence set* is a set $N \subseteq \mathbb{N}$ such that $X|_N = \{0, 1\}^N$. It turns out that the existence of an independence set of positive lower asymptotic density does capture positive entropy,

---

see [10, 3] for a proof in this binary subshift case, and [5, 6, 9] for analogous results in more general settings.

The following theorem gives a new proof of this result (see Corollary 1), but it is stronger, as there is an additional constraint on how the independent choices must be realized. It also applies in the case of a general alphabet.

Theorem 1. *A subshift $X \subseteq A^{\mathbb{N}}$ has positive entropy if and only if it contains a steadily branching binary tree.*

We can interpret $A^{\mathbb{N}}$ as the boundary of an $|A|$-ary tree and X a closed subset of this boundary. A closed subset of the boundary of a tree is determined by finite subpaths in the tree, and it is in this sense that X contains the binary tree. Steadily branching means that the embedded binary tree branches at a uniform sequence of times which form a set of positive lower asymptotic density. A formal definition is given in Section 3.

Positive entropy refers to topological entropy, and in the context of subshifts, means the exponential growth of the number of words. Countable subshifts can have a growth rate arbitrarily close to exponential (see Proposition 2), but clearly contain trees with only finitely many branchings. Thus, if we weaken the assumption of positive entropy, the existence of a steadily branching binary tree can fail in a very strong sense.

It is also easy to find examples with positive entropy where it is not possible to find a tree where branchings happen a syndetic set of times (see Proposition 3). There are also examples where no individual tree which branches at a uniform sequence of times "captures" all of the entropy, as can be seen from [12], (Example 2). These remarks show that we cannot expect to essentially improve the conclusion about the tree under the assumption of positive entropy.

The proof of Theorem 1 is straightforward using the standard results about the winning shift, which is the set of winning turn orders in a certain word-building game associated to X. The winning shift is known to have the same entropy as X. This implies it has a point with positive density, and this point is the branching structure of a steadily branching tree in X. This note is simply a self-contained elaboration of this deduction. I give a combinatorial proof directly from the definitions, and also include an ergodic theoretic proof.

It is easy to deduce a two-sided version of the result as well. For $x \in A^{-\mathbb{N}}$ and $y \in A^{\mathbb{N}}$, write $x \cdot y \in A^{\mathbb{Z}}$ for the configuration with x and y back to back.

Theorem 2. *A subshift $X \subseteq A^{\mathbb{Z}}$ has positive entropy if and only if for some $x \in A^{-\mathbb{N}}$, the set $\{y \in A^{\mathbb{N}} : x \cdot y \in X\}$ contains a steadily branching binary tree.*

The winning shift was introduced in [18], and has also been studied in [14]. We found out while working on [14] that a concept equivalent to the winning shift was discovered in the set systems setting already in 2002 [1] (even if ostensibly only for binary words), and I found out while working on the present paper that [1] has been applied by dynamicists [10, 3] (this paper is independent from [18], but almost contemporary), and their proof that this gives large independence sets essentially boils down to the same proof as I give here. Nevertheless, I feel that though Theorem 1 is an interesting statement about subshifts, to our knowledge the statement does not appear in the literature, and its proof through winning shifts is worth making explicit. Furthermore, unlike Theorem 1, the statement about independence sets does not trivially generalize to non-binary alphabets.

As discussed, I obtain a new proof of the fact entropy implies the existence of certain types of independence sets in the case of a binary alphabet.

Corollary 1. *Let $X \subseteq \{0,1\}^{\mathbb{N}}$ be a subshift. Then, X has positive entropy if and only if it has an independence set of positive lower asymptotic density.*

2

Again, I mention that this result is proved in [10, 3], and that generalizations appear in [5, 6, 9]. To our knowledge, the first reference where the Sauer-Shelah lemma appears in a dynamics context is [5]. The most naive generalization of Corollary 1, the existence of a high density set N with $X|_N = A^N$, fails for a non-binary alphabet A (even for trivial reasons, by artificially increasing the alphabet of the subshift). However, one can state an equivalent condition for having higher entropy in this way [19], (Theorem 8.3), [3], (Theorem 2). (The non-trivial direction of this equivalence also appears in [18], (Proposition 5.9).)

As I show in this note, winning shifts say something stronger than Corollary 1 in the case of one-dimensional subshifts, but it is not clear how to generalize them to other settings, at least in a way that allows seeing them naturally as a dynamical system.

Question 1. *Does every expansive system of positive entropy contain a steadily branching binary tree (and what does that even mean)? More generally, can the notion of winning shift be extended to general (expansive) systems? Group/monoid actions?*

2 Definitions

Let $\mathbb{N} \ni 0$ be the natural numbers. An *alphabet* is a finite set A. *Words* are elements of the free monoid generated by A, denoted by A^*. Elements of A are seen as belonging to A^* when convenient, $u \cdot v$ denotes the multiplication in the free monoid (formal concatenation). The *length* $|w|$ of a word is defined in the obvious way. Write $A^n = \{w \in A^* : |w| = n\}$.

The set $A^{\mathbb{N}}$, for a finite alphabet A (or later $A^{\mathbb{Z}}$), always carries the product topology, and it is homeomorphic to Cantor space if $|A| \geq 2$. A *subshift* is a closed subset of $A^{\mathbb{N}}$ for a finite alphabet A (or later $A^{\mathbb{Z}}$) satisfying $\sigma(X) \subseteq X$ where σ is the *shift* $\sigma(x)_i = x_{i+1}$. The *words of a subshift* $X \subseteq A^{\mathbb{N}}$ are the words $w \in A^*$ such that $x|_{[0,|w|-1]} = w$ for some $x \in X$. The words of a subshift form its *language*. Writing L_n for the intersection of the language of X with A^n, the *(topological) entropy* of a subshift X is the exponential growth rate of the number of words, $h(X) = \lim_{n \to \infty} \frac{\log |L_n|}{n}$ (which exists by the subadditivity of $\log |L_n|$). The *lower asymptotic density* of $N \subseteq \mathbb{N}$ is $\liminf_k \frac{|N \cap [0,k-1]|}{k}$.

Two subshifts $X \subset A^{\mathbb{N}}, Y \subset B^{\mathbb{N}}$ are *topologically conjugate* if there is a shift-commuting homeomorphism $\phi : X \to Y$. A subshift is *countable* if it countably has many points. A set $N \subseteq \mathbb{N}$ is *syndetic* if $\mathbb{N} \subseteq N + [-n, n]$ where $+$ denotes the Minkowski sum.

3 The Winning Shift

For $z, b \in \mathbb{N}^{\mathbb{N}}$, denote $z \leq b \iff \forall i : z_i \leq b_i$.

Definition 1. *Let A be a finite alphabet and let $X \subseteq A^{\mathbb{N}}$. A tree in X with branching structure $b \in \mathbb{N}^{\mathbb{N}}$ is a set of sequences $(x^z)_{\mathbb{N}^{\mathbb{N}} \ni z \leq b}$ where each x^z is an element of X, and:*

$$x^z_{[0,i)} = x^{z'}_{[0,i)} \iff z_{[0,i)} = z'_{[0,i)}.$$

A steadily branching binary tree is a tree with branching structure $b \in \{0,1\}^{\mathbb{N}}$ satisfying

$$\liminf_{i \to \infty} \frac{\sum_{j=0}^{i-1} b_j}{i} > 0.$$

In other words, for distinct sequences $z, z' \in \mathbb{N}^{\mathbb{N}}$, the first position where x^z and $x^{z'}$ differ is the same as the first position where z and z' differ. For binary b, this means that the nonzero positions

3

in b are the positions where our tree must branch in two, and more generally, $b_i = n$ means the tree must branch n times in position i, explaining why we call this sequence a branching structure. Note also that if $X \subseteq A^{\mathbb{N}}$ and X contains a tree with branching structure $b \in \mathbb{N}^{\mathbb{N}}$, then necessarily $b_i \le |A| - 1$ for all $i \in \mathbb{N}$.

Definition 2. *Let $X \subseteq A^{\mathbb{N}}$ be a subshift. Let $W(X)$, the* winning shift *of X, be the set of all branching structures of trees in X.*

This is defined in a game-theoretic framework in [18]: a tree in X with branching structure z can be interpreted as a winning strategy for the first player in a word-building game where on the ith turn, the first player picks a set of $z_i + 1$ symbols, then the second picks one of them, and the first player wins if the word obtained in the limit is in X.

It is shown in [18] that the winning shift is indeed a subshift. A subshift $Y \subseteq \mathbb{N}^{\mathbb{N}}$ is *hereditary* [9] if for all $y' \in \mathbb{N}^{\mathbb{N}}$ we have:

$$(y \in Y \wedge \forall i : y'_i \le y_i) \implies y' \in Y.$$

The following fact is Proposition 5.7 of [18] (the hereditarity claim is trivial). The notation differs slightly, in the reference $\tilde{W}(X)$ is used for what we call $W(X)$, and their $W(X)$ is the same as ours in the binary case, but differs in general.

Lemma 1. *The subshift $W(X)$ is hereditary and has the same number of words of each length as X.*

Proof sketch. The hereditarity claim is trivial. For the claim about words, one can define $W(L)$ for $L \subseteq A^n$ analogously to the subshift case (using finite trees). Letting $L \subseteq A^n$ be the language of X intersected with A^n, $W(L)$ is the language of $W(X)$ intersected with A^n. Now, we write $L = 0L_0 \sqcup 1L_1 \sqcup \cdots \sqcup (k-1)L_{k-1}$ and exchange a suitable sum:

$$
\begin{aligned}
|W(L)| &= |\{i \cdot w \ : \ |\{j \ : \ w \in W(L_j)\}| > i\}| \\
&= \sum_i \sum_{\substack{w \in A^{n-1} \\ |\{j \ : \ w \in W(L_j)\}| > i}} 1 \\
&= \sum_{w \in A^{n-1}} \sum_{i=0}^{|\{j \ : \ w \in W(L_j)\}| - 1} 1 \\
&= \sum_{w \in A^{n-1}} |\{j \ : \ w \in W(L_j)\}| \\
&= \sum_j |W(L_j)| = \sum_j |L_j| = |L|.
\end{aligned}
$$

The first equality is true by definition, and the penultimate one by induction. $\qquad\square$

For a finite word $w \in \mathbb{N}^*$, write $\sum w = \sum_i w_i$ for the sum of the symbols in w, and $|w|$ for the length of w as a word. The key to finding steadily branching trees is to study the *density* $\sum w/|w|$ of a word w. If $Y \subseteq \{0,1\}^{\mathbb{N}}$ is a subshift, by Y^k we mean the Cartesian power with the diagonal action, interpreted in an obvious way as an alphabet over the alphabet $\{0,1\}^k$.

We observe a simple combinatorial lemma.

Lemma 2. *If a subshift Y has positive entropy and is hereditary, then for some $\beta > 0$ there exist arbitrarily long words w of $Y \cap \{0,1\}^{\mathbb{N}}$ with $\sum w/|w| \ge \beta$.*

Proof. Let $\{0, 1, ..., k-1\}$ be the alphabet of Y. Let:

$$b(Y) = \{x \in \{0,1\}^{\mathbb{N}} \ : \ \exists y \in Y : \forall i : y_i \neq 0 \iff x_i \neq 0\}.$$

Since Y is hereditary, it is easy to see that the number of words in Y of length n is at most the number of words in $b(Y)^k$ of length n. Thus, if Y has positive entropy, so does $b(Y)^k$, and thus so does $b(Y)$. Suppose thus that the number of words in $b(Y)$ of every length n is at least α^n for some $\alpha > 0$, as is clearly implied by positive entropy.

The number of words of length n with at most k many 1s is at most:

$$\binom{n}{k} 2^k < \left(\frac{n \cdot e}{k}\right)^k 2^k.$$

Setting $k = \beta n$, this becomes $((\frac{n \cdot e}{\beta n})^\beta 2^\beta)^n$ and we observe that as $\beta \to 0$, we have $2^\beta \to 1$ and $(\frac{n \cdot e}{\beta n})^\beta = (\frac{e}{\beta})^\beta \to 1$, so for small enough $\beta > 0$, we have $(\frac{n \cdot e}{\beta n})^\beta 2^\beta < \alpha$, and thus there must be words of length n in $b(Y)$ with at least βn many 1s, for arbitrarily large n. These are words of $Y \cap \{0,1\}^{\mathbb{N}}$ since Y is hereditary. $\square$

This argument appears also in [11] (Theorem 4.8) (this theorem already includes the statement about lower asymptotic density, which we deduce in the next section).

4   The Proofs

Proof of Theorem 1. Obviously, a steadily branching tree implies positive entropy.

For the other direction, let $A \subseteq \mathbb{N}$ be a finite alphabet and $Y \subseteq A^{\mathbb{N}}$ a subshift. Write s_n for the maximal sum $\sum w$ of a word w of length n in Y. This sequence is clearly subadditive, so $\lim_{n\to\infty} s_n/n$ exists, say $\lim_n s_n/n = \alpha$.

We outline the usual addendum that there must be a configuration $y \in Y$ such that every prefix w of y satisfies $\sum w/|w| \geq \alpha$. Supposing that this is not the case, then every point X has a prefix w such that $\sum w/|w| < \alpha$, and this must happen after a bounded number of steps by compactness, thus there exists $\epsilon > 0$ such that for some m, we always find a prefix w of length at most m in any point $y \in Y$, such that $\sum w/|w| < \alpha - \epsilon$. Now, given any long word w of Y, we can split it as $w = w^0 w^1 \cdots w^k$ (where i in w^i denotes a superscript, not a power) with $|w^k| \leq m$ and $\sum w^i/|w^i| < \alpha - \epsilon$ for all $i < k$. If w is long enough, then since $\sum uv/|uv| \leq \max(\sum u/|u|, \sum v/|v|)$ for all words u, v, we have:

$$\sum w/|w| < \alpha - \epsilon/2$$

for all long enough words, a contradiction to $\lim_n s_n/n = \alpha$.

Now, we set $Y = W(X)$ and apply Lemma 2. The lemma implies that $\alpha > 0$ in the above. The point $y \in W(X)$ whose density stays above α gives the branching structure of a steadily branching tree in X. $\square$

The "usual addendum" is well known. The author learned it from [8, 7] and P. Guillon; Jeandel attributes it to [13, 4], and it also appears in [11, 3] where it is attributed to [15, 16].

The point of this note was to provide a combinatorial proof of Theorem 1 from the first principles, as this is easy to do. However, I have included also the proof using ergodic theory (see [2] for a basic reference).

Ergodic theoretic proof of Theorem 1. Obviously a steadily branching tree implies positive entropy. For the other direction, since the winning shift has positive entropy, it admits an invariant measure μ with positive entropy, thus $\mu([a]) > 0$ for some $a > 0$. By the ergodic decomposition there exists such an ergodic measure, and by the pointwise ergodic theorem, there is a point $y \in W(X)$ where the lower asymptotic density of as is positive, giving the result. $\qquad\square$

We conclude with the proofs of Theorem 2 and Corollary 1.

Proof of Theorem 2. Let $X_R \subseteq A^{\mathbb{N}}$ be the subshift of the right tails of points in X and apply Theorem 1. Let $y \in \{0,1\}^{\mathbb{N}}$ be the branching structure of some steadily branching tree and α the lower asymptotic density. It is easy to see that for any n, we can find a finite prefix w of $y = wx$ of a length of at least n such that for all prefixes u of x of length at most n, $\sum u/|u| \geq \alpha$. Otherwise, by cutting long words into ones of length at most n, as in the previous argument, we see that the lower asymptotic density is less than α.

The fact we can find such wx as a branching structure implies that the tree corresponding to x can follow at least one word of length $|w|$. Letting n tend to infinity, by compactness, we obtain a steadily branching tree that can follow some left tail. $\qquad\square$

Proof of Corollary 1. Supposing X has positive entropy, let $y \in W(X) \cap \{0,1\}^{\mathbb{N}}$ have positive lower asymptotic density. The set $N = \{n \ : \ y_n = 1\}$ is an independence set (which by assumption has positive lower asymptotic density): Suppose $(x^z)_{\mathbb{N}^{\mathbb{N}} \ni z \leq y}$ form a tree with branching structure y, i.e., we have:
$$x^z_{[0,i)} = x^{z'}_{[0,i)} \iff z_{[0,i)} = z'_{[0,i)}.$$

Supposing that $(m_i)_{i \in \mathbb{N}}$ enumerates N in order, let $N_n = \{m_0, ..., m_{n-1}\}$, and for $w \in \{0,1\}^n$ let $z^w \in \{0,1\}^{\mathbb{N}}$ be the characteristic sequence of $\{m_i \ : \ i < |w|, w_i = 1\}$. Clearly $z^w \leq y$ for all $w \in \{0,1\}^n$. The map:
$$\phi : 2^n \to \{x^z|_{N_n} \ : \ z \leq b\}$$

defined by $\phi(w) = x^{z^w}|_{N_n}$ is injective, because $\phi(u0v)_{m_{|u|}} \neq \phi(u1v)_{m_{|u|}}$. Thus, this map is surjective, giving $X|_{N_n} = \{0,1\}^{N_n}$. By compactness, we have $X|_N = \{0,1\}^N$. $\qquad\square$

5 Proofs of supplementary claims

In this section, I provide brief proofs for the claims made in the introduction.

Proposition 1. *There exists a subshift $X \subset \{0,1\}^{\mathbb{N}}$ with positive entropy, such that every topologically conjugate subshift $Y \cong X$ has zero independence entropy.*

Proof. By a compactness argument, any subshift X with positive independence entropy must contain two points which differ in only one position, i.e., $\exists x, y \in X : x_0 \neq y_0 \land \forall i \neq 0 : x_i = y_i$. Positive entropy does not imply the existence of a pair of points with finitely many differences, see [?], (Theorem 1.3), and clearly this property is preserved under conjugacy. $\qquad\square$

Proposition 2. *For any function $f : \mathbb{N} \to \mathbb{N}$ satisfying $f(n) = o(a^n)$ for all $a > 1$, there exists a countable subshift $X \subset \{0,1\}^{\mathbb{N}}$ such that for all large enough n, the number of words of length n in X is at least $f(n)$.*

Proof. The forbidden patterns of X are described as follows: if a finite subword contains n many 1s, they must be pairwise separated by a distance of at least m_n. Here, $(m_n)_n$ is some nondecreasing sequence. The number of words of X of length n is at least $2^{\lfloor n/m_n \rfloor}$, and if m_n grows slowly enough this stays above f. By the assumption on f we can have $m_n \to \infty$, and it is easy to see that X is then countable, since all its points have finite sum and finite subsets of $\mathbb{N}$ form a countable set. $\square$

Proposition 3. *There exists a subshift $X \subset \{0,1\}^{\mathbb{N}}$ with positive entropy such that every point in $W(X)$ contains arbitrarily long subwords of the form 0^n.*

Proof. If X itself is hereditary, then $W(W(X)) = W(X)$, so it suffices to find a hereditary X with positive entropy such that every point $x \in X$ contains arbitrarily long subwords of the form 0^n. For this, it clearly suffices to find a binary subshift Y where every point contains the subword 0^n for arbitrarily large n, but some configuration contains 1s with positive lower asymptotic density, as the smallest hereditary subshift containing Y then has the claimed property: entropy comes from independently flipping the positive density sequence of 1s, and taking the hereditary closure can clearly only add 0^n-subwords to points. The existence of such Y is folklore, see e.g. [17], (Example 7.3), for an explicit construction. $\square$

References

[1] Richard P. Anstee, Lajos Rónyai, and Attila Sali. Shattering news. *Graphs and Combinatorics*, 18(1):59–73, 2002.

[2] Manfred Denker, Christian Grillenberger, and Karl Sigmund. *Ergodic theory on compact spaces*. Lecture Notes in Mathematics, Vol. 527. Springer-Verlag, Berlin, 1976.

[3] Fryderyk Falniowski, Marcin Kulczycki, Dominik Kwietniak, and Jian Li. Two results on entropy, chaos and independence in symbolic dynamics. *Discrete Contin. Dyn. Syst. Ser. B*, 20(10):3487–3505, 2015.

[4] De-Jun Feng and Wen Huang. Lyapunov spectrum of asymptotically sub-additive potentials. *Comm. Math. Phys.*, 297(1):1–43, 2010.

[5] Eli Glasner and Benjamin Weiss. Quasi-factors of zero-entropy systems. *J. Amer. Math. Soc.*, 8(3):665–686, 1995.

[6] Wen Huang and Xiangdong Ye. A local variational relation and applications. *Israel J. Math.*, 151:237–279, 2006.

[7] Emmanuel Jeandel. Computability of the entropy of one-tape Turing machines. In *31st International Symposium on Theoretical Aspects of Computer Science*, volume 25 of *LIPIcs. Leibniz Int. Proc. Inform.*, pages 421–432. Schloss Dagstuhl. Leibniz-Zent. Inform., Wadern, 2014.

[8] Teturo Kamae and Michael Keane. A simple proof of the ratio ergodic theorem. *Osaka J. Math.*, 34(3):653–657, 1997.

[9] David Kerr and Hanfeng Li. Independence in topological and C^*-dynamics. *Mathematische Annalen*, 338(4):869–926, 2007.

[10] Marcin Kulczycki, Dominik Kwietniak, and Jian Li. Entropy of subordinate shift spaces. *Amer. Math. Monthly*, 125(2):141–148, 2018.

[11] Dominik Kwietniak. Topological entropy and distributional chaos in hereditary shifts with applications to spacing shifts and beta shifts. *Discrete Contin. Dyn. Syst.*, 33(6):2451–2467, 2013.

[12] Erez Louidor, Brian Marcus, and Ronnie Pavlov. Independence entropy of $\mathbb{Z}^d$-shift spaces. *Acta Applicandae Mathematicae*, pages 1–21, 2013.

[13] László Máté. On infinite composition of affine mappings. *Fund. Math.*, 159(1):85–90, 1999.

[14] Peltomäki, Jarkko and Salo, Ville. On winning shifts of marked uniform substitutions. *RAIRO-Theor. Inf. Appl.*, 53(1–2):51–66, 2019.

[15] Yuval Peres. A combinatorial application of the maximal ergodic theorem. *Bull. London Math. Soc.*, 20(3):248–252, 1988.

[16] I. Z. Ruzsa. On difference sets. *Studia Sci. Math. Hungar.*, 13(3-4):319–326 (1981), 1978.

[17] Ville Salo. Subshifts with sparse traces. *Studia Math.*, 255(2):159–207, 2020.

[18] Ville Salo and Ilkka Törmä. Playing with subshifts. *Fundam. Inform.*, 132(1):131–152, 2014.

[19] Benjamin Weiss. *Single orbit dynamics*, volume 95 of *CBMS Regional Conference Series in Mathematics*. American Mathematical Society, Providence, RI, 2000.