Symbolic Dynamics

MPRI course by Olivier Carton and Valérie Berthé

Personal notes of Jean Abou Samra (work in progress!)

Contents

1. Subshift basics

1.a. Topology of infinite words

Let A be an alphabet, i.e., a finite nonempty set. To avoid degenerate cases, we assume A has at least two letters. We endow A with the discrete topology. Note that A is trivially compact (because all open covers are already finite).

An infinite word or ω-word on the alphabet A is just a sequence of letters, i.e., an element of $A^{\mathbb{N}}$ (sometimes denoted A^{ω}). A subset of $A^{\mathbb{N}}$ is sometimes called an ω-language.

We endow the space $A^{\mathbb{N}}$ with the product topology.

The topology can also be described more explicitly. By definition of a product topology, a prebase of this topology is made of sets $A^k P A^{\mathbb{N}}$, where k is an integer and P is a subset of letters (these are the inverse images of subsets of A by the canonical projections). One easily checks that the finite intersections of these sets are the sets which can be written as $w_1 A^{\mathbb{N}} \cup w_2 A^{\mathbb{N}} \cup \ldots \cup w_n A^{\mathbb{N}}$, where $w_1, \ldots, w_n$ are finite words. Therefore, the sets $wA^{\mathbb{N}}$, for w finite word, form a basis of the topology. Now, define $d : A^{\mathbb{N}} \times A^{\mathbb{N}} \to \mathbb{R}_+$ by setting, for all infinite words u, v, $d(u,v) = 0$ if $u = v$, else $d(u,v) = 2^{-n}$, where n is the smallest index where u and v differ. One easily checks that d is a distance, and even an ultrametric distance. Thus, endowed with this distance, $A^{\mathbb{N}}$ is an ultrametric space. It is easily seen that the open balls are exactly the sets $wA^{\mathbb{N}}$ with w a finite word, i.e., the elements of the basis we have found earlier. In conclusion, the topology induced by this distance is the same as the product topology. In the context of infinite words, the open balls $wA^{\mathbb{N}}$ are sometimes called cylinders.

In $A^{\mathbb{N}}$, the notion of convergence that the topology induces is simple (pointwise) convergence. This follows from the definition of the distance: if a sequence converges to a limit, the indices at which a word from the sequence start to differ from the limit grow to infinity.

We can see that $A^{\mathbb{N}}$ is compact in a number of ways: by Tychonoff's theorem; by proving the Borel-Lebesgue property with König's lemma; by showing every sequence has a converging subsequence, again with König's lemma.

Let us restate the characterization of open sets and dualize it for closed sets in a symmetric way:

Prefix-based characterization of open sets: A subset $U \subset A^{\mathbb{N}}$ is open iff there exists $P \subset A^*$ such that U is the set of infinite words having some finite prefix in P.

Prefix-based characterization of closed sets: A subset $C \subset A^{\mathbb{N}}$ is closed iff there exists $P \subset A^*$ such that C is the set of infinite words having *all* finite prefixes in P.

These is a nice visualization of these two characterizations in terms of infinite trees. The set of finite words A^* can be viewed as an infinite $|A|$-ary tree where each node is a finite word, and an infinite word is an infinite branch of this tree. A cylinder is a "cone": the set of infinite branches passing through a certain node. An open set is a union of such cones, and a closed set is the set of infinite branches of some subtree.

We can also say a few things about minimum and maximum P sets in these two characterizations:

- Given an open set U, we can find a maximum such P, by taking the set of finite words v such that all words starting with v are in U. In our visualization, we take all words v such that the cone rooted in v is included in U. This maximum P has the following property: for all $v \in A^*$, one has $v \in P$ iff for all $a \in A$, $va \in P$.

  Proof. Easy. □

  Conversely, given a P that has this property, then it is maximum for the corresponding U.

  Proof. Take v such that all words starting with v are in U, and let us show $v \in P$. If v has a proper prefix in P, this is immediate, by the assumption on P. Now assume v has no proper prefix in P. In particular, for all $w \in A^{\mathbb{N}}$, there is a prefix z of w such that $vz \in P$. Build a subtree of A^* by removing nodes z such that $vz \in P$. By contraposing König's lemma, this subtree is finite. By applying the assumption on P in a bottom-up fashion from the leaves to the root, we get $v \in P$. □

- Given an open set U, another P that works is the set of finite words wa (where $w \in A^*$ and $a \in A$) such that all infinite words starting with wa are in U, and some word starting with w is not in U, or the singleton with just the empty word if $U = A^{\mathbb{N}}$. In our visualization, this is the set of roots of the cones.

  Proof. Easy. □

- Given C, we can find a minimum such P, the set of prefixes of words of C. In the visualization, C, as a set of infinite branches, defines a subtree, and P is the set of nodes of this subtree. This minimum P has the following property: for all $v \in A^*$, one has $v \in P$ iff there exists $a \in A$ such that $va \in P$. Conversely, a P with this property is minimum for the corresponding U.

  Proof. Dualize the statement for open sets. □

- (We can also dualize the second statement, but it seems less interesting. In the visualization, the set P is the set of all nodes except the roots of cones.)

Let us move to clopens:

Prefix-based characterization of clopens: A subset $Y \subset A^{\mathbb{N}}$ is clopen iff there exists $P \subset A^*$ *finite* such that Y is the set of infinite words having some finite prefix in P. (This is the same statement as for open sets, but with P required to be finite.)

Proof.

- Let P be a set of finite words and let Y be the set of infinite words with some prefix in P. Since Y is open, let us show it is closed. Because P is finite, the length of words in P is bounded by some integer n. Let Q be the set of infinite words which either have length $> n$, or have length $\leq n$ and are not in P. Then Y is the set of infinite words with all their prefixes in Q, so Y is closed.
- Let Y be a clopen set. Because it is open, it can be written as a union of cylinders. Because it is closed and $A^{\mathbb{N}}$ is compact, it is compact, therefore this open cover has a finite subcover.

□

1.b. Subshifts

The shift map, denoted $T : A^{\mathbb{N}} \to A^{\mathbb{N}}$ is just the map $u_0 u_1 u_2 u_3 \ldots \mapsto u_1 u_2 u_3 u_4 \ldots$ which erases the first letter. Trivially, the shift map is continuous. The orbit of an infinite word u under shift, noted $O(u)$ is the set $\{T^n(u), n \in \mathbb{N}\}$, or in words: the set of suffixes of u.

A subshift (or shift space, or just shift) is a closed subset X of $A^{\mathbb{N}}$ which is closed under shift (i.e., $\forall u \in X, T(u) \in X$). Viewing a closed set of infinite words as the set of infinite paths in an infinite tree, subshifts correspond to autosimilar trees: what can be observed on any node can also be observed from the root.

Because a subshift is a closed set of a compact space, any subshift is compact.

Since a subshift is closed, it can be characterized as the set of infinite words with all prefixes in P, for some minimum set P of finite words. One easily checks that such a set is closed under shift iff P is suffix-closed. In other words, instead of a characterization by prefixes, we now have a characterization by factors:

Factor-based characterization of subshifts: A subset $X \subset A^{\mathbb{N}}$ is a subshift iff there exists $P \subset A^*$ such that X is the set of infinite words with all their factors in P. In the visualization, X is the set of nodes of an autosimilar tree, i.e., a tree such that the subtree at any node is equal to the whole tree. Given X, we can find a minimum such P, which is the set of factors of words of X, denoted Fact(X).

Some fundamental examples: The set of infinite words on $A = \{a, b\}$ without two consecutive a letters is a subshift, called the golden subshift. The full shift is just all of $A^{\mathbb{N}}$, and the empty shift is just the empty set.

The set of *bi-infinite* words on $A = \{a, b\}$ where all maximal blocks of consecutive a letters have even length is a subshift, called the even subshift. Note that the even subshift is not defined on infinite words but on bi-infinite words, i.e., not on $A^{\mathbb{N}}$ but on $A^{\mathbb{Z}}$. One can endow $A^{\mathbb{Z}}$ with the product topology just like $A^{\mathbb{N}}$, recovering a notion of subshift (adapting all the characterizations is straightforward though tedious).

Instead of characterizing a subshift through allowed factors, we can of course complement and use forbidden factors. A subshift of finite type is one which can be described through a finite set of forbidden factors.

A sofic subshift is one which can be described through a set of factors P which is a regular language. If X is sofic, then the minimal P, namely $\mathrm{Fact}(X)$ is also regular.

Proof. Let R be a regular language. Take a (complete) deterministic finite automaton $\mathcal{A}$ that recognizes R. Let X be the subshift of infinite words with all their factors in R. Then X is the set of infinite words whose path in $\mathcal{A}$ only passes through final states. By removing non-final states, then removing states that cannot lead to a cycle, one gets an automaton which recognizes $\mathrm{Fact}(X)$. Therefore, $\mathrm{Fact}(X)$ is regular. □

Trivially, every subshift of finite type is sofic.

Examples: The golden shift is of finite type: it can be described by a single forbidden factor aa. The even shift can be described by the forbidden factors $b(aa)^{2n+1}b$ for $n \in \mathbb{N}$. On the other hand, it is easily seen not to be of finite type.

It is an easy observation that any intersection of subshifts is a subshift. This allows to define generated subshifts: the subshift generated by $X \subset A^{\mathbb{N}}$ is the smallest subshift containing X, i.e., the intersection of all subshifts containing X (the empty intersection being $A^{\mathbb{N}}$). We can also describe generated subshifts more explicitly. Observe that if Y is closed under shift, then $\mathrm{cl}(Y)$ is also closed under shift. Therefore, the subshift generated by X is $\mathrm{cl}(O(X))$ (because it must contain $\mathrm{cl}(O(X))$, and that is a subshift). One also checks that $\mathrm{Fact}(\mathrm{cl}(O(X))) = \mathrm{Fact}(X)$, so the subshift generated by X is the subshift defined by the set of factors of X.

1.c. Subshift morphisms

A subshift morphism between two subshifts $X \subset A^{\mathbb{N}}$ and $Y \subset B^{\mathbb{N}}$ (on two alphabets A and B, possibly different) is a map $\varphi : X \to Y$ such that:

- φ commutes with the shift operator: $\varphi \circ T_A = T_B \circ \varphi$.

  ("Commutes with *the* shift" is technically improper since the shift operators on the two alphabets, T_A and T_B, are not the same.)

- φ is continuous, in the topological sense. In other words, φ "commutes with the limit operator" (also improper): if (u_n) converges, then $\varphi(u_n)$ converges and $\lim \varphi(u_n) = \varphi(\lim u_n)$.

A local map between two subshifts $X \subset A^{\mathbb{N}}$ and $Y \subset B^{\mathbb{N}}$ is a map $\varphi : X \to Y$ for which there exist an integer k and a function $\hat{\varphi} : \mathrm{Fact}_k(X) \to B$ such that

$$\varphi(u_0 u_1 u_2 u_3 \ldots) = \hat{\varphi}(u_0 u_1 \ldots u_{k-1}) \hat{\varphi}(u_1 \ldots u_k) \hat{\varphi}(u_2 \ldots u_{k+1}) \ldots$$

In other words, the n-th output letter of φ only depends on the input letters from the n-th to the $(n+k)$-th (in a way that is "the same everywhere": it does not depend on n either).

A map is a subshift morphism iff it is a local map.

Proof. It is clear from the definitions that any local map is a subshift morphism.

For the converse, we use a topological argument. Take $\varphi : X \to Y$ a subshift morphism. Recall that X is compact, being a subshift. By Heine's theorem, φ is uniformly continuous. Therefore, there exists k such that for all $x, y \in X$, $d(x, y) < k$ implies $d(\varphi(x), \varphi(y)) < 1/2$, which means that any two words which agree on the k first letters are sent to words with the same first letter. Call $\hat{\varphi} : \mathrm{Fact}_k(X) \to Y$ the map which sends every word u of length k which appears as a prefix of words of X to the common first letter of all images of such words of X by φ. Now recall that φ commutes with shift. For all $x \in X$, we have

$$\begin{aligned}
\varphi(x_0 x_1 x_2 x_3 \ldots) &= \hat{\varphi}(x_0 x_1 \ldots x_k) \varphi(x_1 x_2 x_3 \ldots) \\
&= \hat{\varphi}(x_0 x_1 \ldots x_k) \hat{\varphi}(x_1 x_2 \ldots x_{k+1}) \varphi(x_2 x_3 x_4 \ldots) \\
&= \hat{\varphi}(x_0 x_1 \ldots x_k) \hat{\varphi}(x_1 x_2 \ldots x_{k+1}) \hat{\varphi}(x_2 x_3 \ldots x_{k+2}) \varphi(x_3 x_4 x_5 \ldots)
\end{aligned}$$

Continuing indefinitely, we get

$$\varphi(x_0 x_1 x_2 x_3 \ldots) = \hat{\varphi}(x_0 x_1 \ldots x_k) \hat{\varphi}(x_1 x_2 \ldots x_{k+1}) \ldots$$

which proves that φ is local. □

2. Subshifts and automata

2.a. Recognizing sofic subshifts with semiautomata

As a preliminary, recall that given a non-deterministic finite automaton $\mathcal{A}$, we can trim $\mathcal{A}$ by removing states q which are either not accessible (there is no path from some initial state to q) or not co-accessible (there is no path from q to some final state). Trivially, this does not change the regular language recognized by $\mathcal{A}$.

For subshifts, it will be more interesting to work with automaton structures: an semiautomaton is like an automaton, but without initial or final states. It can recognize a language, the set of labels of finite paths in the structure, which is always factor-closed. It can also recognize an ω-language, the set of labels of infinite paths in the structure. Note that an semiautomaton needs to be non-complete in order to recognize a non-trivial language. (One can see an semiautomaton as a Büchi automaton where all states are initial and final.)

If $\mathcal{A}$ is a trimmed automaton recognizing a factor-closed language, then the semiautomaton $\mathcal{A}'$ obtained by forgetting initial and final states in $\mathcal{A}$ recognizes the same language — no new words are accepted.

Proof. Take a path $p \to \ldots \to q$ in $\mathcal{A}'$. Since $\mathcal{A}$ is trimmed, this path can be extended to a path $i \to \ldots \to p \to \ldots \to q \to \ldots \to f$, where i is an initial state and f is a final state. This proves that the label of $p \to \ldots \to q$ is a factor of a word accepted by $\mathcal{A}$, and therefore it is accepted by $\mathcal{A}$, since the language of $\mathcal{A}$ is factor-closed. □

We can also "ω-trim" semiautomata, by removing states that cannot be part of an infinite path (this preserves the ω-language recognized by the structure, though it does not preserve the language).

We can now give an automata-based characterization of sofic subshifts: A set is a sofic subshift iff it is the ω-language recognized by some automaton structure. More precisely:

- If X is a sofic subshift, there exists an semiautomaton $\mathcal{A}$ recognizing the language $\mathrm{Fact}(X)$, and any such structure $\mathcal{A}$ recognizes the ω-language X.
- If $\mathcal{A}$ is an semiautomaton, the ω-language X it recognizes is a sofic subshift, and if $\mathcal{A}$ is additionally ω-trimmed, then the language it recognizes is $\mathrm{Fact}(X)$.

Proof.

- From semiautomata to sofic subshifts:

  Let $\mathcal{A}$ be an semiautomaton. We want to show that the ω-language of $\mathcal{A}$ is a sofic subshift X, and if $\mathcal{A}$ is additionally ω-trimmed then its language is $\mathrm{Fact}(X)$. Since ω-trimming does not change the ω-language, assume wlog that $\mathcal{A}$ is ω-trimmed.

  Let X be the ω-language of $\mathcal{A}$. Note that X is trivially stable under shift (since one can always remove the first step of an infinite path).

Let us show that X is topologically closed. Let (u_n) be a sequence in X which converges to an infinite word u. By convergence, for all n, the prefix $u_0...u_{n-1}$ of u is a prefix of all terms of (u_n) after a certain index. Since all terms of (u_n) are in X, there is a path labeled $u_0...u_{n-1}$ in $\mathcal{A}$. Build a tree by combining all these paths. König's lemma gives us an infinite branch of the tree, which is an infinite path in $\mathcal{A}$ labeled u. Therefore $u \in X$.

We have proved that X is a subshift. Let us show that $\mathcal{A}$ recognizes $\mathrm{Fact}(X)$. Obviously, every word of $\mathrm{Fact}(X)$ is recognized by $\mathcal{A}$ (a word of $\mathrm{Fact}(X)$ is a factor of a word which has an infinite path in $\mathcal{A}$, therefore it has a finite path in $\mathcal{A}$). Conversely, given v a word which has a finite path in $\mathcal{A}$, the finite path can be extended to an infinite path, because $\mathcal{A}$ is ω-trimmed, and therefore $v \in \mathrm{Fact}(X)$.

Finally, the subshift X is sofic because $\mathrm{Fact}(X)$ is regular. Indeed, $\mathrm{Fact}(X)$ is recognized by $\mathcal{A}$, which can be seen as a non-deterministic finite automaton where all states are initial and final.

- From sofic subshifts to semiautomata:

  Let X be a sofic subshift. Since $\mathrm{Fact}(X)$ is regular and factor-closed, it is recognized by some semiautomaton. Let $\mathcal{A}$ be any such structure.

  Let $u \in X$. Every prefix of u is accepted by $\mathcal{A}$, and thus is the label of some path in $\mathcal{A}$. Using König's lemma, we can find an infinite path in $\mathcal{A}$ labeled by u.

  Conversely, let u be the label of some infinite path in $\mathcal{A}$. Each factor of u is the label of some finite path in $\mathcal{A}$, and therefore belongs to $\mathrm{Fact}(X)$. By the factor-based characterization of subshifts, $u \in X$.

  Combining these facts, $\mathcal{A}$ recognizes the ω-language X.

□

2.b. De Brujin semiautomata for subshifts of finite type

For a subshift X of finite type on bi-infinite words, we can give a simple construction of an automaton which recognizes X, its de Brujin automaton (structure). Indeed, X can be characterized by a finite set of forbidden factors F. Let n be the maximum length of an element of F. The states of the de Brujin semiautomaton for F are words of length n which do not contain a forbidden factor from F, with all transitions $au_1u_2...u_{n-1} \xrightarrow{b} u_1u_2...u_{n-1}b$. This automaton is deterministic. The construction can be easily adapted (adding some states for the beginning of the computation on a word) to infinite instead of bi-infinite words.

2.c. Local semiautomata

An (n, d)-local semiautomaton is an semiautomaton where for all paths

$$p_0 \xrightarrow{a_1} p_1 \xrightarrow{a_2} p_2 \xrightarrow{a_3} \dots \xrightarrow{a_{n-1}} p_{n-1} \xrightarrow{a_n} p_n$$

$$p'_0 \xrightarrow{a_1} p'_1 \xrightarrow{a_2} p'_2 \xrightarrow{a_3} \dots \xrightarrow{a_{n-1}} p'_{n-1} \xrightarrow{a_n} p'_n$$

we have $p_d = p'_d$.

In other words, there exist a "left extension length" (d) and a "right extension length" ($n - d$) such that from two different states q, s, one can never build both reverse paths of the left extension length from q and s labeled by the same word, and paths of the right extension length from q and s labeled by the same word. When we write "n-local", it means "(n, d)-local for some d" (n is the sum of the left extension length and the right extension length).

Yet another formulation: an semiautomaton is local iff *a bi-infinite word is always accepted in at most one way*: there do not exist two distinct states p, q and a bi-infinite word u such that there is a bi-infinite path labeled u centered in p and a bi-infinite path labeled u centered in q. Schematically:

$$\mathcal{A} \text{ local} \iff \nexists \begin{array}{c} \cdots \xrightarrow{a_{-3}} \xrightarrow{a_{-2}} \xrightarrow{a_{-1}} p \xrightarrow{a_0} \xrightarrow{a_1} \xrightarrow{a_2} \cdots \\ \neq \\ \cdots \xrightarrow{a_{-3}} \xrightarrow{a_{-2}} \xrightarrow{a_{-1}} q \xrightarrow{a_0} \xrightarrow{a_1} \xrightarrow{a_2} \cdots \end{array}$$

Proof. $\mathcal{A}$ being local obviously implies there do not exist such p, q. Conversely, assume $\mathcal{A}$ is not local. For some pair of distinct states q, s, there do not exist left and right extension lengths as in the definition of a local semiautomaton (if extension lengths existed for all pairs of states, one could take the max to get extension lengths for the whole automaton structure). Take such a pair q, s. By construction, there exist arbitrarily long paths labeled by the same word starting from q and s or ending on q and s. König's lemma gives us infinite paths labeled by the same infinite word starting and ending on q and s. □

An unambiguous semiautomaton is an semiautomaton where there is never more than one path to go from a given state to a given state with a given word as label. Schematically:

$$\mathcal{A} \text{ unambiguous} \iff \nexists \begin{tikzcd} & r \arrow[dr, "v"] & \\ p \arrow[ur, "u"] \arrow[dr, "u"'] & \neq & q \\ & s \arrow[ur, "v"'] & \end{tikzcd}$$

Obviously, a deterministic or co-deterministic semiautomaton is unambiguous.

If a strongly connected semiautomaton is local, then it is unambiguous.

Proof. Let $\mathcal{A}$ strongly connected and local. Assume $\mathcal{A}$ is ambiguous. Let p, q, r, s states and u, v words as in the diagram above. By strong connectivity, there exists a word w such that $q \xrightarrow{w} p$:

$$\begin{tikzcd} & r \arrow[dr, "v"] & \\ p \arrow[ur, "u"] \arrow[dr, "u"'] & \neq & q \arrow[ll, bend left=60, "w"] \\ & s \arrow[ur, "v"'] & \end{tikzcd}$$

This contradicts locality, since there are bi-infinite paths

$$\begin{array}{c} \cdots \xrightarrow{u} r \xrightarrow{v} q \xrightarrow{w} p \xrightarrow{u} r \xrightarrow{v} q \xrightarrow{w} p \xrightarrow{u} r \xrightarrow{v} \cdots \\ \neq \\ \cdots \xrightarrow{u} s \xrightarrow{v} q \xrightarrow{w} p \xrightarrow{u} s \xrightarrow{v} q \xrightarrow{w} p \xrightarrow{u} s \xrightarrow{v} \cdots \end{array}$$

□

Let $\mathcal{A}$ be a strongly connected semiautomaton. Then $\mathcal{A}$ is local iff there do not exist two different, non-empty, identically labeled cycles. Schematically:

$$\mathcal{A} \text{ strongly connected} \Rightarrow$$

$$\mathcal{A} \text{ local} \iff \nexists \begin{tikzcd} p \arrow[loop below, "u"] \arrow[r, phantom, "\neq"] & q \arrow[loop below, "u"] \end{tikzcd}$$

Proof. Let $\mathcal{A}$ strongly connected.

If there exist p, q, u as on the diagram, then obviously $\mathcal{A}$ is not local, since there are bi-infinite paths

$$\cdots \xrightarrow{u} p \xrightarrow{u} p \xrightarrow{u} p \xrightarrow{u} \cdots$$
$$\neq$$
$$\cdots \xrightarrow{u} q \xrightarrow{u} q \xrightarrow{u} q \xrightarrow{u} \cdots$$

For the converse, assume $\mathcal{A}$ is not local. Take bi-infinite paths

$$\cdots \xrightarrow{a_{-2}} p_{-1} \xrightarrow{a_{-1}} p_0 \xrightarrow{a_0} p_1 \xrightarrow{a_1} \cdots$$
$$\neq$$
$$\cdots \xrightarrow{a_{-2}} q_{-1} \xrightarrow{a_{-1}} q_0 \xrightarrow{a_0} q_1 \xrightarrow{a_1} \cdots$$

We distinguish two cases. The first case is when there are infinitely many $i < 0$ such that $p_i \neq q_i$. In this case, by the pigeonhole principle, there exist indices $i < j < 0$ such that $p_i \neq q_i$, $p_j \neq q_j$ and $(p_i, q_i) = (p_j, q_j)$. Then we are done, since we have found cycles

$$p_i \xrightarrow{a_i} \cdots \xrightarrow{a_j} p_j = p_i$$
$$\neq$$
$$q_i \xrightarrow{a_i} \cdots \xrightarrow{a_j} q_j = q_i$$

The second case is when there are only finitely many $i < 0$ such that $p_i \neq q_i$. Then, for some low enough i, we have $p_i = q_i$. We distinguish again two cases for positive indices, depending on whether there finitely or infinitely many $i' > 0$ such that $p_i = q_i$. For "infinitely", we are done like for negative indices, and for "finitely", we get $i' > 0$ such that $p_{i'} = q_{i'}$. Using the strong connectivity of $\mathcal{A}$, we find a path $p_{i'} \xrightarrow{b_0} \cdots \xrightarrow{b_n} p_i$. This concludes, since we have cycles

$$\begin{tikzcd}
p_i \arrow[r, "a_i"] & \cdots \arrow[r, "a_{-1}"] & p_0 \arrow[r, "a_0"] & \cdots \arrow[r, "a_{i'}"] & p_{i'} \arrow[llll, bend right, "b"'] \\
= & & \neq & & = \\
q_i \arrow[r, "a_i"] & \cdots \arrow[r, "a_{-1}"] & q_0 \arrow[r, "a_0"] & \cdots \arrow[r, "a_{i'}"] & q_{i'} \arrow[llll, bend left, "b"]
\end{tikzcd}$$

□

Locality test: Using the previous proposition, we can test whether a strongly connected semiautomaton $\mathcal{A}$ is local. It suffices to build the structure $\mathcal{A} \times \mathcal{A}$ and to check for the presence of a non-empty cycle which passes through at least one non-diagonal state (where "diagonal" means "of the form (q, q)"). This can be done in quadratic time.

Another corollary is that if $\mathcal{A}$ is a local, strongly connected automaton structure, then $\mathcal{A}$ is $|\mathcal{A}|^2$-local (where $|A|$ is the number of states in $\mathcal{A}$). (TODO: Olivier Carton's notes claim $|\mathcal{A}|^2 - |\mathcal{A}|$ but I'm not sure how this is deduced.)

Finally, we give a link between local semiautomata and subshifts: The (sofic) subshift recognized by a local semiautomaton is of finite type.

Proof. Let $\mathcal{A}$ be an n-local automaton. Wlog, assume $\mathcal{A}$ is ω-trimmed. (If it is not, we can always ω-trim it, which preserves the language and yields an automaton that is still n-local.) Let X be the ω-language recognized by $\mathcal{A}$. By the automata-based characterization of sofic subshifts, X is a sofic subshift and $\mathrm{Fact}(X)$ is the language of $\mathcal{A}$. We wish to prove that every word not in $\mathrm{Fact}(X)$ has a "small" factor not in $\mathrm{Fact}(X)$, where "small" means "below a certain constant". We claim that $n + 1$ is such a constant. We prove this by contraposition. Take a word u of length $|u| \geq n + 1$ such that all length $n + 1$ factors of u are in $\mathrm{Fact}(X)$. Since $u_0 \ldots u_{n-1} \in \mathrm{Fact}(X)$, there is a path

$$q_0^{(0)} \xrightarrow{u_0} \ldots \xrightarrow{u_n} q_{n+1}^{(0)}$$

Likewise, there is a path

$$q_0^{(1)} \overset{u_1}{\rightarrow}\rightarrow \ldots \overset{u_{n+1}}{\rightarrow} q_{n+1}^{(1)}$$

By locality, we get $q_{d+1}^{(0)} = q_d^{(1)}$. Iterating, we get a path

$$q_0^0 \overset{u_0}{\rightarrow} \ldots \underset{d}{\overset{u}{\rightarrow}} q_{d+1}^{(0)} = q_d^{(1)} \overset{u_{d+1}}{\rightarrow} q_{d+1}^{(1)} = q_d^{(2)} \overset{u_{d+2}}{\rightarrow} \ldots$$

and therefore $u \in \text{Fact}(X)$. □

Note that the converse holds: if a subshift is of finite type, then it is recognized by some local semiautomaton. This is because a de Brujin automaton is local (when we have a sufficiently long path arriving to a state, the n last transitions suffice to tell us what that state is). In a sense, local automata are "de Brujin-like" automata.

2.d. Determinization of semiautomata

TODO

2.e. Minimization of semiautomata

TODO

3. Factor complexity

3.a. Complexity function

Given an infinite word u, we note p_u the factor complexity function of u, defined by $p_u(n) = |\text{Fact}_u(n)|$, i.e., $p_u(n)$ is the number of distinct factors of length n in u. Obviously, $1 \leq p_u(n) \leq |A|^n$.

If w is a factor of u, we note w^+ the number of letters a such that wa is still a factor of u (note that w^+ implicitly depends on u). Likewise, w^- is the number of letters a such that aw is still a factor of u. Remark that for all infinite word u, for all n,

$$p_u(n+1) = \sum_{w \in \text{Fact}_u(n)} w^+ = \sum_{w \in \text{Fact}_u(n)} w^-$$

We can generalize factor complexity to subshifts ($p_X(n)$ is the number of different factors of length n in words of X). Note that for all infinite word u, $p_{\text{cl}(O(u))} = p_u$.

3.b. Periodicity

A periodic word is an infinite word that can be written as u^ω for some finite word u (where u^ω denotes uuu...).

An ultimately periodic word is an infinite word with a periodic suffix. An aperiodic word is an infinite word which is not ultimately periodic.

Characterizations of periodicity: Let u be an infinite word. The following are equivalent:

(1) u is ultimately periodic,
(2) p_u is bounded,
(3) $p_u(n) \leq n$ for some n
(4) p_u is not (strictly) increasing
(5) $O(u)$ is finite
(6) $\mathrm{cl}(O(u))$ is finite

The implication (4) ⇒ (1) is known as the Morse-Hedlund theorem, often stated as the contrapositive: an aperiodic word has increasing factor complexity.

Proof. (1) ⇒ (2) and (2) ⇒ (3) are obvious.

(3) ⇒ (4): If $p_u(1) = 1$, then u is constant, and obviously (4) holds. Otherwise, $p_u(1) = 2$, and there exists $n > 1$ such that $p_u(n) \leq n$, so p_u cannot be strictly increasing (as that would imply $p_u(n) > n$ for all n, by induction).

(4) ⇒ (1) There exists k such that $p_u(k+1) - p_u(k) = 0$. By the previous formula, for all $w \in \mathcal{L}_u(k)$, $w^+ = 1$. This entails that u is ultimately periodic, since there is a unique way to choose the next letter after k given letters.

(1) $\iff$ (5) $\iff$ (6) are clear. □

TODO: can we add (7) $O(u)$ is closed ?

3.c. Entropy

We recall Fekete's lemma: Let $(u_n)_{n\in\mathbb{N}}$ be a real sequence. Assume (u_n) is subadditive, i.e., for all m, n, it holds that $u_{m+n} \leq u_m + u_n$. Then (u_n/n) converges and

$$\lim_{n\to+\infty}\left(\frac{u_n}{n}\right) = \inf_{n\in\mathbb{N}}\left(\frac{u_n}{n}\right)$$

Let u be an infinite word. Observe that the sequence $\left(\log_{|A|}(p_u(n))\right)$ is subadditive. Indeed, for all m, n, noting that a factor of length $m+n$ is the concatenation of a factor of length m and a factor of length n, we have

$$\begin{aligned}\log_{|A|}(p_u(m+n)) &\leq \log_{|A|}(p_u(m)\cdot p_u(n)) \\ &= \log_{|A|}(p_u(m)) + \log_{|A|}(p_u(n))\end{aligned}$$

By Fekete's lemma, the sequence $\left(\left(\log_{|A|}(p_u(n))\right)/n\right)$ converges. We define the entropy of u as

$$H(u) = \lim \frac{\log_{|A|}(p_u(n))}{n} = \inf \frac{\log_{|A|}(p_u(n))}{n} \in [0,1]$$

Just like factor complexity, we can generalize entropy to subshifts, using the same definition. Again, $H(u) = H(\mathrm{cl}(O(u)))$.

Examples: The entropy of the full shift is obviously 1, and that of the empty shift is obviously 0. The entropy of the golden shift is $\log_2(\varphi)$, where φ is the golden ratio

$$\varphi = \frac{1+\sqrt{5}}{2}$$

Proof. Denote the golden shift by X, and for all $n \in \mathbb{N}$, denote by A_n, resp. B_n, the number of distinct factors of X ending with the letter a, resp. b. Observe that, by definition of the even shift, $B_{n+1} = A_n + B_n$, since the letter b can follow any factor without aa, and $A_{n+1} = B_n$, since the letter a can only follow a factor that ends in b. Hence, $B_{n+2} = A_{n+1} + B_{n+1} = B_n + B_{n+1}$. For $n \neq 0$, we have $p_X(n) = A_n + B_n$, and therefore $p_X(n) = B_{n+1}$, thus $p_X(n+2) = B_{n+3} = B_{n+2} + B_{n+1} = p_X(n+1) + p_X(n)$. We also note that $p_X(1) = 2$ and $p_X(2) = 3$. Therefore, $p_X(n) = F_{n+1}$, where F_n is the Fibonacci sequence (in fact this works for $n = 0$ too). We use the well-known formula $F_n = \frac{1}{\sqrt{5}}(\varphi^n + (-\varphi)^{-n})$ to get

$$\frac{1}{n}\log_2(p_X(n)) \sim \frac{1}{n}\log_2(\varphi^{n+1}) = \frac{n+1}{n}\log_2(\varphi) \sim \log_2(\varphi)$$

and therefore $H(X) = \log_2(\varphi)$. □

TODO:

Soit $\mathcal{A}$ un automate fortement connexe non-ambigu. Sa matrice associée est irréductible apériodique.

On calcule… entropie = $\ln \rho(A)$.

4. Recurrent words

A recurrent word is an infinite word u such that all factors of u appear infinitely often in u. If additionally all factors appear with bounded gaps, meaning that for all factor v, there is n such that every factor of length n contains v, then the word is said to be uniformly recurrent.

A minimal subshift is a subshift which does not have any non-empty proper subshift, i.e., a subshift X such that for all $u \in X$, one has $\mathrm{cl}(O(u)) = X$.

For all infinite word u, the subshift $\mathrm{cl}(O(u))$ is minimal iff u is uniformly recurrent.

Proof.

- Assume u is uniformly recurrent. Take a subshift X such that $\emptyset \subsetneq X \subseteq \mathrm{cl}(O(u))$ and let us prove $X = \mathrm{cl}(O(u))$. For this, we use the factor-based characterization of subshifts. Let w be a factor of u. We want to prove that w is a factor of some word of X. In fact, we will see that it is a factor of all words of X (which suffices because X is nonempty). Let $v \in X$. By definition of $\mathrm{cl}(O(u))$, v is the limit of a sequence (v_n) of factors of u of increasing lengths. Since u is uniformly recurrent, there is N such that every length N factor of u contains w. After a certain index, the first N letters of v_n stabilize to a factor of u, which contains w. Hence, w is a factor of v.
- Assume $\mathrm{cl}(O(u))$ is minimal. Suppose by contradiction that u is not uniformly recurrent. There is a factor v and a sequence (u_n) of factors of increasing lengths such that v is not a factor of any of the (u_n). By compactness, (u_n) has a subsequence which converges to a word w. Then $\mathrm{cl}(O(w))$ is a nonempty subshift included in $\mathrm{cl}(O(u))$, and the inclusion is strict because v is not a factor of w.

□

If a subshift X is minimal, then all the words it contains are uniformly recurrent. (The converse is false: take the union of the subshifts generated by two uniformly recurrent words on disjoint alphabets.)

Proof. Assume X is minimal. For all $u \in X$, $\mathrm{cl}(O(u))$ is a nonempty subshift of the minimal subshift X, so $\mathrm{cl}(O(u)) = X$ and $\mathrm{cl}(O(u))$ is minimal, thus u is uniformly recurrent. □

We can also define a linearly recurrent infinite word: one for which the gaps increase at most linearly with factor size. Given a uniformly recurrent word u, define, for every factor v, the number $g(v)$ as the minimum length such that every factor of u having this length contains v. Define g_n as the maximum of $g(v)$ among factors v of length n. Then u is said linearly recurrent when $g_n = \mathcal{O}(n)$.

5. Purely substitutive words

5.a. Fixed points of substitutions

A substitution is a map $\sigma : A \to A^+$ which we extend as $\sigma : A^* \to A^*$ by $\sigma(a_1 a_2 ... a_n) = \sigma(a_1)...\sigma(a_n)$ and as $\sigma : A^{\mathbb{N}} \to A^{\mathbb{N}}$ by $\sigma(a_1 a_2 a_3 ...) = \sigma(a_1)\sigma(a_2)\sigma(a_3)....$ Substitutions equipped with composition have a monoid structure.

Note that the image of every letter must be nonempty (otherwise the extension as $A^{\mathbb{N}} \to A^{\mathbb{N}}$ would not work).

A substitution σ is right-prolongable in a letter a when $\sigma(a)$ can be written as $\sigma(a) = av$ with $v \neq \varepsilon$ (mind the latter condition: $\sigma(a) = a$ is not allowed). If this is the case, we remark that

$$\begin{aligned}
\sigma(a) &= a \cdot u \\
\sigma^2(a) &= a \cdot u \cdot \sigma(u) \\
\sigma^3(a) &= a \cdot u \cdot \sigma(u) \cdot \sigma^2(u) \\
&\cdots
\end{aligned}$$

Thus, the sequence $\sigma^n(a)$ has a limit, denoted $\sigma^\infty(a)$. It is a fixed point of σ.

Examples. The Fibonacci word is generated by the substitution

$$\begin{aligned}
a &\mapsto ab \\
b &\mapsto a
\end{aligned}$$

which is right-prolongable in a. One checks it has factor complexity $n \mapsto n+1$. It starts with $abaababaabaab....$

The Thue-Morse word is generated by

$$\begin{aligned}
a &\mapsto ab \\
b &\mapsto ba
\end{aligned}$$

It starts with $abbabaabbaababba....$ (TODO: factor complexity?)

Observe that if u is a fixed point of σ then there are only two possibilities.

- First case: u is made entirely of letters a such that $\sigma(a) = a$.
- Second case: there is a letter a in u such that σ is right-prolongable in a. All letters b that appear before satisfy $\sigma(b) = b$.

5.b. Primitive substitutions

The matrix of a substitution is defined for a substitution σ as $\mathrm{Mat}(\sigma)_{i,j} = |\sigma(a_j)|_{a_i}$, where $a_1, ..., a_n$ are the letters. It has size $n \times n$.

Let $\Psi : A^* \to \mathbb{N}^n$ be the map that counts occurrences of the letters, i.e., $\Psi(u) = \left(|u|_{a_1}, ..., |u|_{a_n}\right)$. Then $\mathrm{Mat}(\sigma)$ is the matrix (in the standard basis) of the unique linear map that sends $\Psi(u)$ to $\Psi(\sigma(u))$ for all u.

Observe that Mat is a monoid morphism from substitutions to matrices.

A substitution σ is said primitive when $\mathrm{Mat}(\sigma)$ is primitive. In other words: for some power of the substitution, the image of every letter must contain every letter. (Remark that if σ is primitive, then σ^k is primitive for all $k \in \mathbb{N}^*$.)

With general substitutions, it annoyingly does not hold that if u is a fixed point of σ starting with a, then σ is right-prolongable in a and $u = \sigma^\infty(a)$. For example, we can take the Fibonacci substitution σ and extend it to a three-letter alphabet by letting $\sigma(c) = c$. Then $(\sigma^n(ca))_n$ converges to a fixed point of σ which is the Fibonacci word with c prepended, but σ is not right-prolongable in c and $(\sigma^n(c))_n$ does not converge to an infinite word. Fortunately, with primitive substitutions, this technical annoyance is removed:

If a primitive substitution σ has a fixed point u starting with a, then σ is right-prolongable in a and $(\sigma^n(a))$ converges to u.

Proof. Simply remark that $\sigma(a)$ must start with a, and if it was reduced to a, then one would have $\sigma^n(a) = a$ for all n, and σ would not be primitive (because the alphabet has at least two letters by assumption). □

Any primitive substitution admits a power that has a fixed point.

Proof. Let σ be a primitive substitution. Let a be any letter. By the pigeonhole principle, there exist $k, l \in \mathbb{N}^*$ with $k < l$ such that $\sigma^k(a)$ and $\sigma^l(a)$ start with the same letter, which we call b. The word $\sigma^l(a)$, which is also $\sigma^{l-k}(\sigma^k(a))$, starts with $\sigma^{l-k}(b)$, which is non-empty (since all letters are sent by σ to non-empty words), thus $\sigma^{l-k}(b)$ starts with b. Now, σ^{l-k} is primitive, there exists m such that $\sigma^{(l-k)m}(b)$ contains all letters. This word starts with b, and because it contains all letters, it is not reduced to b (since we always assume the alphabet has at least two letters). Therefore, $\sigma^{(l-k)m}$ is right-prolongable in b. □

Fixed points of powers of a primitive substitution all generate the same subshift.

Proof. A fixed point of σ^l and a fixed point of σ^l are both fixed points of σ^{kl}, therefore it is enough to treat the case of two fixed points of the same primitive substitution σ (not of some powers of it). Let σ primitive, let u and v fixed points of σ starting with the letters a and b respectively. Let w be a factor of u. Recall that u is the limit of $(\sigma^n(a))_n$. We can choose n large enough so that $\sigma^n(a)$ contains w. Since σ is primitive, there exists p such that $\sigma^p(b)$ contains a, therefore v contains a, and since it is a fixed point of σ, it contains $\sigma^n(a)$, which contains w. We have proved that all factors of u are factors of v. By symmetry, all factors of v are factors of u. We conclude by the factor-based characterization of subshifts. □

By the previous two theorems, we can defined the subshift of a primitive substitution, which is the common subshift generated by all fixed points of some power of that substitution (such fixed points exist and all give the same subshift).

Lemma: Let σ be a primitive substitution. Denote by ρ the dominant eigenvalue of its matrix given by the Perron-Frobenius theorem. Then for all letter a, there exists a constant C such that $|\sigma^n(a)| \underset{n\to+\infty}{\sim} C\rho^n$.

Proof. Let i be the index of a in the ordering of the alphabet used to define M_σ. Recall that $|\sigma^n(a)|$ is the sum of the i-th column of M_{σ^n}, which is equal to $(M_\sigma)^n$. By the Perron-Frobenius theorem, $(M_\sigma/\rho)^n \to A$ for some positive matrix A. Setting C to be the sum of the i-th column of A, we get $|\sigma^n(a)| \sim C\rho^n$. □

A fixed point of a primitive substitution is linearly recurrent.

Proof.

Let σ be a primitive substitution. Let k be such that $(M_\sigma)^k > 0$. Let u be a fixed point of σ, starting with a letter a.

First, we prove that u is uniformly recurrent. Let v be a factor of u.

Let v be a factor of u. We have $u = \sigma^k(a)\sigma^k(u_1)\sigma^k(u_2)...$ and all the $\sigma^k(u_i)$ contain a, therefore a appears with gaps bounded by some constant n (explicitly, n can be chosen to be twice the maximum of $|\sigma^k(b)|$ among letters b). F

□

A fixed point of a primitive substitution is uniquely ergodic.

Proof. TODO □

Remark: Take $\sigma : 0 \mapsto 0010, 1 \mapsto 1$. Then $\sigma^\infty(0)$ is uniformly recurrent, however σ is not primitive.

Proof. TODO, easy. □

5.c. Factor complexity of substitutive words

A constant-length substitution is a substitution where the images of the letters all have the same length.

Let σ be a substitution right-prolongable in a. Let u be the fixed point of σ in a. Then $p_u(n) = \mathcal{O}(n^2)$. Furthermore, if σ is primitive or constant-length, then $p_u(n) = \mathcal{O}(n)$.

6. To be sorted

A measure-theoretic dynamical system is a measure space $(X, \mathcal{B}, \mu)$ endowed with a measurable map $T : X \to T$ such that μ is T-invariant, i.e., for all measurable set B, one has $\mu(T^{-1}(B)) = \mu(B)$.

It is said ergodic if X cannot be partitioned in two non-null measurable sets which are stable under T.

In an ergodic dynamical system, the orbit of x is dense for almost all x.

Proof. TODO □

Birkhoff Ergodic Theorem: Let $(X, \mathcal{B}, \mu, T)$ be an ergodic dynamical system. Let $f \in L^1(X, \mathbb{R})$. Then $\frac{1}{n}\sum_{k=0}^{n-1} f \circ T^k$ converges almost everywhere to $\int_X f\, \mathrm{d}\mu$.

In symbolic dynamics, the dynamical systems considered are over

An infinite word u is said **normal** iff every factor of u of length k has a frequency, which is $1/(|\Sigma|^k)$ (in other words, all factors of length k appear in an equally distributed way).

The **Champernowne word** on the alphabet $\{0, ..., n-1\}$ is the concatenation of base-n representations of 0, 1, 2, 3, … **Theorem:** The Champernowne word is normal. **Theorem:** The real number in $[0, 1]$ that the Champernowne represents is transcendental. (We do not know if converting this real to a representation in any base, possibly other than 2, necessarily gives a normal word.)

Remark: We know that π is transcendental, however we do not know if it is normal in base 10.

Let u be an infinite word. The **frequency** of a finite word v in u is the limit of $a_v \frac{n}{n}$, if it exists, where $a_v(n)$ is the number of times v appears in $u[0..n]$.

The frequency of a factor w in u is defined by $f_u(w) = \lim_{n\to+\infty} \frac{1}{n}|u_0...u_{n-1}|_w$, if this limit exists. It is a uniform frequency if $\frac{1}{n}\left|u_k...u_{n+k-1}\right|_w$ has a limit uniformly in k. A uniquely ergodic infinite word is an infinite word which has a uniform frequency for all factors.

An infinite word u is said C-balanced if for all factors v, w of u having the same length, and for all letter a, one has $\left||v|_a - |w|_a\right| \leq C$.

The discrepancy of u is $\Delta(u) = \lim_{n\to+\infty} \sup_{a\in A}\left||u_0...u_{n-1}|_a - nf_u(a)\right|$.

An infinite word u is balanced iff it has uniform letter frequencies and there exists a constant B such that for all factor w of u, for all letter a, one has $\left||w|_i - |w|f_u(w)\right| \leq B$. Moreover, if u has letter frequencies, then u is balanced iff its discrepancy is finite.

7. Sturmian words

A sturmian word is an infinite word u such that for all n, $p_u(n) = n + 1$.

Let u be a sturmian word. We have $p_u(1) = 2$. Therefore, all sturmian words contain exactly two distinct letters, so we can assume that the alphabet has two letters.

A sturmian word is never ultimately periodic, by the Morse-Hedlund theorem.

A sturmian word is always recurrent.

Proof. Let u a sturmian word. Take w a factor of u, and note n its length. Assume w appears a finite number of times in u. There exists k such that w is not a factor of $T^k(u)$. Then $|\mathcal{L}_{T^k(u)}(n)| \leq (n+1) - 1 = n$, since every length-$n$ factor of $T^k(u)$ is a length-n factor of u, and w is not a length-n factor of $T^k(u)$. Therefore, by the Morse-Hedlund theorem, $T^k(u)$ is ultimately periodic, and therefore u is ultimately periodic, contradicting the assumption that it is sturmian. □

A right-special factor (resp. left-special factor) of u is a factor w of u such that $w^+ \geq 2$ (resp. $w^- \geq 2$). We similarly define right-special and left-special factors in subshifts. We can also speak of a left-special *infinite word* in a subshift.

Let u be a sturmian word, and let n. From

$$1 = p_u(n+1) - p_u(n) = \sum_{w \in \mathcal{L}_u(n)} (w^+ - 1)$$

we deduce that u has a *unique* right special factor of length n, noted $r_n(u)$. Likewise, there is a unique left special factor $l_n(u)$.

Rem: u sturmian $\iff \forall n, \exists! w, w^+ \geq n \iff$???

Observe that for all n, $l_n(u)$ is a prefix of $l_{n+1}(u)$ (easy). Therefore, $\lim_{n \to +\infty} l_n(u)$ is well-defined (each $l_{n+1}(u)$ adds a new letter).

Likewise, for all n, $r_n(u)$ is a suffix of $r_{n+1}(u)$.

Also observe that, noting $\Sigma = \{a, b\}$, we have $a \lim l_n(u) \in \mathrm{cl}(O(u))$ and $b \lim l_n(u) \in \mathrm{cl}(O(u))$, by the factor-based characterization of subshifts. Therefore, $\lim l_n(u)$ is left special in $\mathrm{cl}(O(u))$. Conversely, it is easy to see that if v is left special in $\mathrm{cl}(O(u))$ then $v = \lim l_n(u)$. We obtain:

Proposition: If u is sturmian, then there is a unique left special word in $\mathrm{cl}(O(u))$.

Remark: the converse is false. Exercise!

The n-th **Rauzy graph** of an infinite word u is a digraph whose vertices are length-n factors of u (i.e., $V = \mathcal{L}_u(n)$) and an edge $v \to w$ whenever there is a word z and two letters x, y such that $v = xy$, $w = yz$ and xyz is a factor of u.

Theorem: For all infinite word u, the following are equivalent:

(1) u is recurrent,
(2) All factors of u appear at least twice,
(3) For all factor w of u, $w^- \geq 1$,
(4) All Rauzy graphs of u are strongly connected.

Proof. (1) ⇒ (4) Easy

(4) ⇒ (3) Easy

(3) ⇒ (1) Assume u is not recurrent. It has a prefix which does not reappear, contradicting (3).

(1) ⟺ (2) Easy □

We now wonder: how do Rauzy graphs of sturmian look like?

$l_n(u) \to r_n(u)$ [two arrows back to $l_n(u)$]

Proof: basic case analysis using strong connectivity (because any sturmian word is recurrent).

Spoiler: the sequence of loops taken is, itself, sturmian (!!).

Theorem: Any sturmian word is **uniformly** recurrent.

Proof. Take u sturmian. Observe that the $r_u(n)$ appear with bounded gaps, because of the shape of the Rauzy graphs (another argument is that if $r_u(n)$ never appears after some time, one sees that u becomes ultimately periodic).

We conclude by proving that every factor w of u is a factor of some $r_u(n)$. $\square$

8. Appendix: Succinct topology reminders

8.a. General topology

A **topological space** is a set X endowed with a set of subsets of X, called a **topology**, which is stable under arbitrary unions and finite intersections.

In particular, any topology on X contains the empty set (the empty union) and X itself (the empty intersection).

The **trivial topology** on X is $\{\emptyset, X\}$. It is the smallest possible topology. The **discrete topology** is $\mathcal{P}(X)$ (the set of all subsets of X) and is the largest possible topology. These are trivially topologies.

Elements of the topology are called **open sets**. Their complements are **closed sets**. One checks that closed sets are closed under arbitrary intersection and finite union, and that $\emptyset$ and X are always closed. The **clopen** sets are those which are both open and closed.

An **open neighborhood** of a point $x \in X$ is an open set that contains x. A **neighborhood** of x is a set that contains an open neighborhood of x. Neighborhoods and open neighborhoods of x are closed under arbitrary unions and finite intersections (like open sets).

The **interior** $\operatorname{int}(Y)$ of a subset Y of X is the largest open set contained in Y, i.e., the union of all open sets contained in Y, or the set of points $y \in X$ such that some neighborhood of y is contained in Y. The **closure** (French: *adhérence*) of Y is the smallest closed set containing Y, i.e., the intersection of all closed sets containing Y, or the set of points $y \in X$ such that all neighborhoods of y intersect Y.

A **dense** subset of X is a subset A such that $\operatorname{cl}(A) = X$. A **nowhere dense** subset is a subset A such that $\operatorname{int}(\operatorname{cl}(A)) = \emptyset$.

One checks that an arbitrary intersection of topologies is a topology. Therefore, we can define a **generated topology**: any set $S \subset \mathcal{P}(X)$ generates a topology on X which is the smallest topology containing S, or the intersection of all topologies containing S. When S generates $\mathcal{T}$, one says S is a **subbasis** of $\mathcal{T}$ (French: *prébase*).

One also checks that the topology generated by S is the set of arbitrary unions of finite intersections of elements of S. This justifies the notion of **basis**: a basis of a topology $\mathcal{T}$ is a set B of open sets of $\mathcal{T}$ such that every open set can be written as an arbitrary union of sets from B. Therefore, if S is a subbasis of $\mathcal{T}$, then the set of finite intersections of elements of S is a basis of $\mathcal{T}$. (In particular, if S is stable under finite intersections, then it is a basis of its generated topology.)

A map between two topological spaces is **continuous** if the inverse image of any open set is an open set. One checks that $f : X \to Y$ is continuous iff for every point $x \in X$, f is continuous in x, meaning that for every neighborhood B of $f(x)$, there exists a neighborhood A of x such that $f(A) \subset B$.

A **homeomorphism** of topological spaces is a bijection which is continuous and whose inverse is also continuous. (Homeomorphisms are isomorphisms in the category of topological spaces.)

Take a family of topological spaces $(X_i)_{i \in I}$. The **product topology** on the set product $\Pi_{i \in I} X_i$ is the smallest topology that makes all canonical projections $\pi_j : (\Pi_{i \in I} X_i) \to X_j$ continuous.

A topological space is **Kolmogorov** (or T_0) iff no two distinct points are contained in exactly the same open sets. It is **Fréchet** (or T_1) if for all x, y, there is an open set containing x but not y and an open set containing y but not x, or equivalently, if all singletons are closed. It is **Hausdorff** (or **separated**, or T_2) iff for all distinct points y and y', there exist disjoint open sets Y and Y' such that $y \in Y$ and $y' \in Y'$. Every Hausdorff space is trivially Kolmogorov.

A sequence $(x_n)_{n \in \mathbb{N}}$ in a topological space is said to **converge** to a **limit** x iff for all neighborhood U of x, all terms of (u_n) after a certain index are in U. One checks that in a Kolmogorov space, the limit (when it exists) is unique.

A point x is an **accumulation point** (French: *valeur d'adhérence*) of a sequence (x_n) iff every neighborhood of x contains infinitely many terms of (x_n). In other words, the set of accumulation points of (x_n) is $\cap_{N \in \mathbb{N}} \operatorname{cl}(\{x_n, n \geq N\})$.

A point x is a **subsequential limit** of (x_n) iff some subsequence of (x_n) converges to x. Beware: every subsequential limit is an accumulation point, but the converse is false!

However, there is a fairly large class of spaces on which it holds. A space is **first-countable** iff for each point x, there is a countable neighborhood basis of x, where a **neighborhood basis** of x is a set of open neighborhoods of x such that any neighborhood of x contains a neighborhood from the set. When there exists a countable neighborhood basis of x, there also exists a countable **nested neighborhood basis**, i.e., a countable neighborhood basis which is a weakly decreasing sequence for inclusion (from a countable neighborhood basis $(N_n)_{n \in \mathbb{N}}$, one builds $\left(\cap_{k \leq n} N_k\right)_{n \in \mathbb{N}}$). In a first-countable space, accumulation points are also subsequential limits.

Proof. Let $(x_n)_{n \in \mathbb{N}}$ a sequence and x an accumulation point of (x_n). Let $(V_n)_{n \in \mathbb{N}}$ a countable nested neighborhood basis of x. Choose $\varphi(0)$ such that $x_{\varphi(0)} \in V_0$ (which is possible by definition of an accumulation point) then choose $\varphi(1) > \varphi(0)$ such that $x_{\varphi(1)} \in V_1$, and so on. This builds a subsequence $\left(x_{\varphi(n)}\right)_{n \in \mathbb{N}}$ of $(x_n)_{n \in \mathbb{N}}$, and this subsequence converges to x. To see this, take a neighborhood V of x. By definition of a neighborhood basis, it contains V_N for some n. Since the basis is nested, $x_n \in V_n \subset V$ for all $n \geq N$. $\square$

A map $f : X \to Y$ between topological spaces is said to be **sequentially continuous** iff for any sequence (x_n) which converges to a limit x, the sequence $f(x_n)$ converges to $f(x)$. We have a situation similar to limits vs. subsequential limits: any continuous function is sequentially continuous, and the converse fails, but works if X is first-countable.

Proof. Assume f is continuous and let (x_n) be a sequence which converges to x in X. Let V be a neighborhood of $f(x)$ in Y. By continuity, $f^{-1}(V)$ is an open set of X. Since it contains x, it is a neighborhood of x. Hence, it contains all terms of (x_n) after a certain index, which proves that V contains all terms of $f(x_n)$ after a certain index.

Conversely, assume f is not continuous. There exists $x \in X$ at which f is discontinuous: there is a neighborhood B of $f(x)$ such that for all neighborhood A of x, one has $f(A) \not\subset B$. Take a countable nested neighborhood basis $(N_n)_{n\in\mathbb{N}}$ of x. For each n, choose $x_n \in N_n$ such that $f(x_n) \notin B$, and consider the sequence (x_n). This sequence converges to x, because any neighborhood of x contains some N_m, which contains all terms of (x_n) for $n \geq m$. However, $(f(x_n))$ does not converge to $f(x)$, because B is a neighborhood of $f(x)$ which contains none of the $f(x_n)$. Therefore, f is not sequentially continuous. □

Similarly to first-countability, a space is said to be **second-countable** iff it has a countable basis.

Compactness is a little complex because there are several different, non-equivalent notions of a compact space:

- A space X is **compact** if it satisfies the following equivalent properties:

  (1) Every open cover of X has a finite subcover (i.e., for all family $(U_i)_{i\in I}$ of open sets such that $\cup_{i\in I} U_i = X$, there exists $J \subset I$ finite such that $\cup_{j\in J} U_j = X$).

  (2) Any set of closed sets with the finite intersection property has nonempty intersection. (The finite intersection property means that every finite subset of this set has nonempty intersection.)

  Proof. These are trivially equivalent since closed sets are the complements of closed sets. □

  Beware! In French, *espace compact* is a different notion! Namely, it refers to a compact *and Hausdorff* space. The equivalent of "compact space" is "espace quasi-compact". (This is horrible terminology.)

  Property (1) is often called the "Borel-Lebesgue property".

- A space X is **countably compact** if it satisfies the following equivalent properties:

  (1) "Countable Borel-Lebesgue property": Every *countable* open cover contains a finite subcover.

  (2) Any *countable* set of closed sets with the finite intersection property has nonempty intersection.

  (3) Every weakly decreasing sequence (French: *suite décroissante*) of nonempty closed sets has nonempty intersection.

  (4) Every sequence has an accumulation point.

  Proof. (1) $\iff$ (2) is trivial as before.

  (2) ⇒ (3): A weakly decreasing sequence of nonempty closed sets trivially has the finite intersection property since each finite intersection is an element of the sequence.

  (3) ⇒ (2): Take a countable family of closed sets $(C_n)_{n\in\mathbb{N}}$ with the finite intersection property. Consider the sequence (C'_n) defined by $C'_n = \cap_{i\leq n} C_i$. Then (C'_n) is a weakly decreasing sequence of nonempty closed sets, so it has nonempty intersection, and the intersection of (C_i) is the same as the intersection of (C'_i).

(3) ⇒ (4): The set of accumulation points of (x_n) is $\bigcap_{N\in\mathbb{N}} \mathrm{cl}(\{x_n, n \geq N\})$, which is a weakly sequence of nonempty closed sets.

(4) ⇒ (3): Let (C_n) be a weakly decreasing sequence of nonempty closed sets. Using the axiom of choice[1], take a sequence (x_n) where each x_n belongs to C_n. This sequence has an accumulation point x. We claim that $x \in \bigcap_{n\in\mathbb{N}} C_n$ and therefore $\bigcap_{n\in\mathbb{N}} C_n$ is nonempty. Suppose contrarily that $x \notin C_N$ for some N. Since (C_n) is weakly decreasing, we have $x \notin C_n$ for all $n \geq N$, contradicting the fact that x is an accumulation point of (x_n). □

- A space X is **sequentially compact** if every sequence has a subsequential limit.

We have the implications: compact ⇒ countably compact (because if every open cover has a finite subcover, then in particular, every countable open cover has a finite subcover), and sequentially compact ⇒ countably compact (because if every sequence has a subsequential limit, every sequence has an accumulation point). Being countably compact neither implies being compact nor sequentially compact, and between being compact and sequentially compact, neither one implies the other.

There are many other notions of compactness to tear your hair with, e.g.: Lindelöf, paracompact, precompact, pseudocompact, metacompact, orthocompact, realcompact, …

Any closed set of a compact, resp. countably compact space is compact, resp. countably compact.

Proof. Take a quasi-compact space X, a closed set $Y \subset X$, and a cover $(U_i)_{i\in I}$ of Y by sets which are closed in Y. Since Y is closed, $X \setminus Y$ is open. Thus, for every $i \in I$, $U_i \cup X \setminus Y$ is open. We have an open cover $(U_i \cup X \setminus Y)_{i\in I}$ of X, from which we can extract a finite subcover $(U_j \cup X \setminus Y)_{j\in J}$ by quasi-compactness of X, and then $(U_j)_{j\in J}$ covers Y. Hence, Y is quasi-compact. The proof for countably compact spaces is similar. □

The celebrated and non-trivial **Tychonoff theorem** asserts that an arbitrary product of compact spaces is compact.

8.b. Metric spaces

A **metric space** is a set X endowed with a function $d : X \times X \to \mathbb{R}^+$ called **distance**, such that:

- (Positivity) $\forall xy \in X, d(x,y) = 0 \iff x = y$
- (Symmetry) $\forall xy \in X, d(x,y) = d(y,x)$
- (Triangular inequality) $\forall xyz \in X, d(x,z) \leq d(x,y) + d(y,z)$

For $x \in X$ and $\varepsilon > 0$, we denote by $B(x,\varepsilon)$ the **open ball** around x with radius ε: $B(x,\varepsilon) = \{y \in X, d(x,y) < \varepsilon\}$.

An **ultrametric space** is a metric space whose distance verifies the ultrametric inequality, which is stronger than the triangular inequality: $\forall xyz \in X, d(x,z) \leq \max(d(x,y), d(y,z))$.

[1]No, I will not make the effort to mark theorems that require the axiom of choice.

The **canonical topology** on a metric space X is the one generated by open balls. In fact, open balls are not only a subbasis, but even a basis of this topology. The open sets are exactly subsets $Y \subset X$ such that for all $y \in Y$, there is an open ball centered on y contained in Y.

Proof. Denote by $\mathcal{T}$ the topology generated by open balls, and $\mathcal{T}'$ the set of subsets $Y \subset X$ such that $\forall y \in Y, \exists \varepsilon > 0, B(y, \varepsilon) \subset Y$.

Trivially, $\mathcal{T}'$ is closed under arbitrary union. Let us show that it is closed under finite intersection. The empty intersection is trivial. Let $Y, Y \in \mathcal{T}'$, and let $y \in Y \cap Y'$. Let $\varepsilon, \varepsilon' > 0$ such that $B(y, \varepsilon) \subset Y$ and $B(y, \varepsilon') \subset Y'$. Then $B(y, \min(\varepsilon, \varepsilon')) \subset Y \cap Y'$, which concludes. Therefore, $\mathcal{T}'$ is a topology.

Let us show $\mathcal{T} \subset \mathcal{T}'$. Since $\mathcal{T}$ is the topology generated by open balls and $\mathcal{T}'$ is a topology, it suffices to prove that open balls belong to $\mathcal{T}'$. Take an open ball $B(x, \varepsilon)$. Let $y \in B(x, \varepsilon)$. Define $\varepsilon' = \varepsilon - d(x, y) > 0$. Then $D(y, \varepsilon') \subset B(x, \varepsilon)$. Indeed, if $z \in B(y, \varepsilon')$, then $d(x, z) \leq d(x, y) + d(y, z) < d(x, y) + \varepsilon - d(x, y) = \varepsilon$, thus $z \in B(x, \varepsilon)$.

Finally, let us show $\mathcal{T}' \subset \mathcal{T}$. Let $Y \in \mathcal{T}'$. For every $y \in Y$, pick $\varepsilon_y > 0$ such that $B(y, \varepsilon_y) \subset Y$. Then $Y = \cup_{y \in Y} B(y, \varepsilon_y)$. Since Y can be written as a union of open balls, it belongs to $\mathcal{T}$. This argument also show that open balls form a basis of $\mathcal{T} = \mathcal{T}'$. □

Metric spaces are much "better behaved" than general topological spaces, and some notions have simple characterizations.

- Every metric space is Hausdorff. In particular, limits are unique, and the English and French meanings of "compact" coincide.

  Proof. Any two distinct points x and y are separated by $B(x, d(x, y)/2)$ and $B(y, d(x, y)/2)$. □

- A sequence (x_n) converges to a limit x iff for all $\varepsilon > 0$, $d(x, x_n) \leq \varepsilon$ holds for sufficiently large n.

  Proof. In the forward direction, take the neighborhoods $B(x, 1/n)$ of x. For the converse, use the fact that any neighborhood of x contains an open ball centered on x. □

- A subset U is open iff every convergent sequence whose limit is in U has all its terms in U after a certain index.

  Proof. Use the previous point. □

- A subset C is closed iff every convergent sequence whose terms are in C after a certain index has its limit in C.

  Proof. Contrapose the previous point. □

- Every metric space is first-countable.

  Proof.

  For any point x, take $(B(x, 1/n))_{n \in \mathbb{N}}$ as a countable neighborhood basis of x. □

  As a corollary, accumulation points and subsequential limits are the same thing in metric spaces. In particular, a metric space is countably compact (every sequence has an accumulation point) iff it is sequentially compact (every sequence has a subsequential limit).

- A function $f : X \to Y$ between two metric spaces is continuous iff $\forall x \in X, \forall \varepsilon > 0, \exists \eta > 0, \forall y \in X, d(x, y) < \eta \Rightarrow d(f(x), f(y)) < \varepsilon$.

  Proof. Just apply the definition of continuity and the fact that open balls form a basis. □

Also, it is continuous iff it is sequentially continuous, since metric spaces are first countable.

A **uniformly continuous function** $f : X \to Y$ between two metric spaces is a function that satisfies a stronger condition than continuity: $\forall \varepsilon > 0, \exists \eta > 0, \forall x, y \in X, d(x, y) < \eta \Rightarrow d(f(x), f(y)) < \varepsilon$. **Heine-Cantor theorem:** If X is compact, then any continuous function $f : X \to Y$ is uniformly continuous.

Proof. TODO □

Theorem: A metric space is compact iff it is sequentially compact, iff it is countably compact.

Proof. TODO. □

A **Cauchy sequence** in a metric space is a sequence $(x_n)_{n \in \mathbb{N}}$ such that for all $\varepsilon > 0$, it holds that $d(x_m, x_n) \leq \varepsilon$ for sufficiently large m, n. A metric space is **complete** if every Cauchy sequence converges.

Every compact metric space is complete.

Proof. In a compact metric space, a Cauchy sequence has a subsequential limit, which is easily seen to be a limit for the whole sequence. □

A function $f : X \to Y$ between two metric spaces X and Y is said to be **uniformly continuous** (TODO)

8.c. Examples and counterexamples

Define $f : [0, 1[\cup [2, 3[\to [0, 2[$ by $f(x) = x$ for $x \in [0, 1]$ and $f(x) = x - 1$ for $x \in [2, 3]$. Then f is bijective and continuous but not a homeomorphism.

The cofinite topology on $\mathbb{N}$ is the topology where closed sets are $\mathbb{N}$ and finite sets. It is Kolmogorov and Fréchet but not Hausdorff.

Given a poset X, the order topology on X is generated by "strict cones" $\{a \mid a < x\}$ and $\{a \mid a > x\}$. It is always Kolmogorov. If X is totally ordered, then it is Hausdorff. On $\mathbb{N}$, it is the discrete topology. On $\mathbb{R}$, it is the usual metric topology.

The left order topology is generated only by left cones $\{a \mid a < x\}$ and the right order topology by right cones $\{a \mid a > x\}$. On $\mathbb{N}$, the left order topology is made of $\mathbb{N}$ and subsets of the form $[\![0, n[\![$; it is not Fréchet. The right order topology is made of subsets of the form $[\![n, +\infty[\![$, and is also not Fréchet.

The Sierpiński space, which is $\{0, 1\}$ with the topology $\{\emptyset, \{0\}, \{0, 1\}\}$, is Kolmogorov but not Fréchet.

The Alexandrov topology on a poset X is formed by upward-closed sets. It has the interesting property that an *arbitrary* intersection of open sets is open.

The metric space $\mathbb{R}$ is second-countable.

The space of sequences $\mathbb{N} \to [0, 1]$ with the infinite norm is not compact. Consequently, $\mathcal{C}([0, 1], \mathbb{R})$ with the infinite norm is not compact. Both spaces are not second-countable (though they are first-countable, like all metric spaces).

TODO: compare compactness properties

9. Appendix: Perron-Frobenius theorem

Let M be an $n \times n$ matrix. It is is said to be primitive if $M \geq 0$ (the comparison is coefficient by coefficient) and $M^k > 0$ for some $k \in \mathbb{N}$.

Note that if M is primitive then M^l is primitive for all $l \in \mathbb{N}^*$.

Proof. Take k such that $M^k > 0$. Then $\left(M^l\right)^k = \left(M^k\right)^l > 0$. □

Also, given an $n \times n$ matrix M, build the digraph G_M which has indices $1, \ldots, n$, and an edge $i \to j$ iff $A_{i,j} \neq 0$. One says M is irreducible iff this digraph is strongly connected. If this is the case, one says M is aperiodic if the PGCD of the lengths of all cycles in G_M is 1.

A weakly positive matrix is primitive iff it is irreducible aperiodic. (No proof here, but this is not hard to see.)

Perron-Frobenius theorem for primitive matrices: Let M be a primitive matrix.

- The maximum modulus of an eigenvalue of M is attained for a single eigenvalue, which is called the dominant eigenvalue. Let us denote it $\rho(M)$.
- $\rho(M)$ is real and positive.
- $\rho(M)$ is a simple eigenvalue (both in the algebraic and geometric senses): its multiplicity in the characteristic polynomial is 1. Consequently, its multiplicity in the minimal polynomial is also 1, and its eigenspace has dimension 1 and is equal to the generalized eigenspace.
- The eigenspace for $\rho(M)$ is spanned by a positive eigenvector.
- Let x be a positive eigenvector of M for $\rho(M)$, and let y be a positive eigenvector of M^T for $\rho(M^T) = \rho(M)$. Then

$$M^n \underset{n \to +\infty}{\sim} \rho(M)^n \frac{x y^T}{\langle x, y \rangle}$$

Furthermore, the function ρ is increasing: if M and N are two primitive matrices such that $M \leq N$ implies $\rho(M) \leq \rho(N)$, and if additionally $M \neq N$ then $\rho(M) < \rho(N)$. (One could write "$M < N \Rightarrow \rho(M) < \rho(N)$" but $M < N$ is not to be interpreted coefficient-by-coefficient here, rather it means that $M_{i,j} \leq N_{i,j}$ for all i, j and $M_{i,j} < N_{i,j}$ for *some* i, j.)

10. Exercises from Valérie Berthé's exercise sheet

10.a. Exercise 1

1. If u is not recurrent, some factor (let k be its length) no longer occurs in some $v = T^n(u)$, therefore $p_v(k) \leq k + 1 - 1 = k$, thus v is ultimately periodic, and so is u.

2. a. v is ultimately periodic by the same argument as the previous question. Thus v has a suffix w^ω. By the factor-based characterization of subshifts, w^n is a factor of u for all n. The number of occurrences of a in factors of length $|w|$ admit two values, say C and $C + 1$. Wlog, assume $|w|_a = C$ (a symmetric argument would hold if $|w|_a = C + 1$). Consider an occurrence of w in u, which is not in the first run of consecutive occurrences of w (the first exists, since all the w^n appear as factors, and is finite, since u is aperiodic, so we can skip to right after it and find a second run). In a "sliding window" fashion, advance in the word until a letter which makes the suffix differ from w^ω is found (which again terminates because the run is finite). At this point, the letter read must be a and the letter dropped must be b, so that the current factor in the window has $C + 1$ times the letter a, otherwise the current factor in the window would have $C - 1$ times the letter a. We can do the same on the left (we don't reach the start of the word, because we started in the second run of w occurrences), and the same phenomenon happens. Comparing with an occurrence of 2 more consecutive w, we get a contradiction with 1-balance.

   b. Clearly, if u is recurrent, then all its Rauzy graphs must be strongly connected (take an occurrence of a factor, then a later occurrence of a second factor).

   Conversely, suppose all Rauzy graphs of u are strongly connected.