Preservation of normality
by unambiguous transducers

Olivier Carton

June 2, 2020

Abstract

We consider finite state non-determinisitic but unabiguous transducers with infinite inputs and infinite outputs, and we consider the property of Borel normality of sequences of symbols. When these transducers are strongly connected, and when the input is a Borel normal sequence, the output is a sequence in which every block has a frequency given by a weighted automaton over the rationals. We provide an algorithm that decides in cubic time whether a unambiguous transducer preserves normality.

1 Introduction

More than one hundred years ago Émile Borel [2] gave the definition of *normality*. A real number is normal to an integer base if, in its infinite expansion expressed in that base, all blocks of digits of the same length have the same limiting frequency. Borel proved that almost all real numbers are normal to all integer bases. However, very little is known on how to prove that a given number has the property.

The definition of normality was the first step towards a definition of randomness. Normality formalizes the least requirements about a random sequence. It is indeed expected that in a random sequence, all blocks of symbols with the same length occur with the same limiting frequency. Normality, however, is a much weaker notion than the one of purely random sequences defined by Martin-Löf [10].

The motivation of this work is the study of transformations preserving randomness, hence preserving normality. The paper is focused on very simple transformations, namely those that can be realized by finite-state machines. We consider automata with outputs, also known as sequential transducers, mapping infinite sequences of symbols to infinite sequences of symbols. Input deterministic transducers were considered in [6] where it was shown that preservation of normality can be checked in polynomial time for these transducers. This paper extends the results to unambiguous transducers, that is, transducers where each sequence is the input label of exaclty one accepting run. These machines are of great importance because they coincide with functional transducers in the following sense. Each unambiguous transducer is indeed functional as there is at

1

most one output for each input but is was shown conversely that each functional transducer is equivalent to some unambiguous one [8].

An auxiliary result involving weighted automata is introduced to obtain the main result. It states that if an unambiguous and strongly connected transducer is fed with a normal sequence then the frequency of each block in the output is given by a weighted automaton on rational numbers. It implies, in particular, that the frequency of each block in the output sequence does not depend on the input nor the run labeled with it as long as this input sequence is normal.

The paper is organized as follows. Notions of normal sequences and transducers are introduced in Section 2. Main results are stated in Section 3. Proofs of the results and algorithms are given in Section 6.

2 Basic Definitions

2.1 Normality

Before giving the formal definition of normality, let us introduce some simple definitions and notation. Let A be a finite set of *symbols* that we refer to as the *alphabet*. We write $A^{\mathbb{N}}$ for the set of all sequences on the alphabet A and A^* for the set of all (finite) words. Let us denote by μ the uniform measure on $A^{\mathbb{N}}$. The length of a finite word w is denoted by $|w|$. The positions of sequences and words are numbered starting from 1. To denote the symbol at position i of a sequence (respectively word) w we write $w[i]$, and to denote the substring of w from position i to j inclusive we write $w[i{:}j]$. The empty word is denoted by λ. The cardinality of a finite set E is denoted by $\#E$.

Given two words w and v in A^*, the number $|w|_v$ of occurrences of v in w is defined by

$$|w|_v = \#\{i : w[i{:}i + |v| - 1] = v\}.$$

For example, $|abbab|_{ab} = 2$. Given a word $w \in A^+$ and a sequence $x \in A^{\mathbb{N}}$, we refer to the *frequency of w in x* as

$$\mathrm{freq}(x, w) = \lim_{n \to \infty} \frac{|x[1{:}n]|_w}{n}$$

when this limit is well-defined.

A sequence $x \in A^{\mathbb{N}}$ is *normal* on the alphabet A if for every word $w \in A^*$:

$$\mathrm{freq}(x, w) = \frac{1}{(\#A)^{|w|}}$$

An occurrence of v is called *aligned* if its starting position i (as above) is such that $i - 1$ is a multiple of the length of v. An alternative definition of normality can be given by counting aligned occurrences, and it is well-known that they are equivalent (see for example [1]). We refer the reader to [4, Chap.4] for a complete introduction to normality.

The most famous example of a normal word is due to Champernowne [7], who showed in 1933 that the infinite word obtained from concatenating all the natural numbers (in their usual order):

0123456789101112131415161718192021222324252627282930 . . .

is normal on the alphabet $\{0, 1, \ldots, 9\}$.

2.2 Automata and transducers

In this paper we consider automata with outputs, also known as transducers. We refer the reader to [12] for a complete introduction to automata accepting sequences. Such finite-state machines are used to realize functions mapping words to words and especially sequences to sequences. Each transition of these transducers consumes exactly one symbol of their input and outputs a word which might be empty. As many reasoning ignore the outputs of the transitions, we first introduce automata.
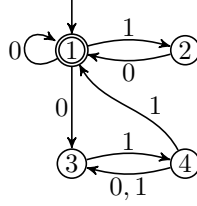

Figure 1: An unambiguous automaton

A *(Büchi) automaton* $\mathcal{A}$ is a tuple $\langle Q, A, \Delta, I, F \rangle$ where Q is the state set, A the alphabet, $\Delta \subseteq Q \times A \times Q$ the transition relation, $I \subseteq Q$ the set of initial states and F is the set of final states. A transition is a tuple $\langle p, a, q \rangle$ in $Q \times A \times Q$ and it is written $p \xrightarrow{a} q$. A *finite run* in $\mathcal{A}$ is a finite sequence of consecutive transitions,

$$q_0 \xrightarrow{a_1} q_1 \xrightarrow{a_2} q_2 \cdots q_{n-1} \xrightarrow{a_n} q_n$$

Its *input* is the word $a_1 a_2 \cdots a_n$. An *infinite run* in $\mathcal{A}$ is a sequence of consecutive transitions,

$$q_0 \xrightarrow{a_1} q_1 \xrightarrow{a_2} q_2 \xrightarrow{a_3} q_3 \cdots$$

A run is *initial* if its first state q_0 is initial, that is, belongs to I. A run is called *final* if it visits infinitely often a final state. Let us denote by $q \xrightarrow{x} \infty$ the existence of a final run labeled by x and starting from state q. An infinite run is *accepting* if it is both initial and final. An automaton is called *unambiguous* if each sequence is the label of at most one accepting run. By definition, deterministic automata automata are unambiguous but they are not the only ones as it is shown by Figure 1.

Each automaton $\mathcal{A}$ can be seen as a directed graph $\mathcal{G}$ by ignoring the labels of its transitions. We define the *strongly connected components* (SCC) of $\mathcal{A}$ as the strongly connected components of $\mathcal{G}$. An automaton $\mathcal{A}$ is called *strongly connected* if it has a single strongly connected component.

A transducer with input alphabet A and output alphabet B is informally an automaton whose labels of transitions are pairs (a, v) in $A \times B^*$. The pair (a, v) is usually written $a|v$ and a transition is thus written $p \xrightarrow{a|v} q$. The symbol a and the word v are respectively called the *input label* and the *output label* of the transition. More formally a *transducer* $\mathcal{T}$ is a tuple $\langle Q, A, B, \Delta, I, F \rangle$, where Q is a finite set of states, A and B are the input and output alphabets respectively, $\Delta \subseteq Q \times A \times B^* \times Q$ is a finite transition relation and $I \subseteq Q$ is the set of initial states and F is the set of final states of the Büchi acceptance condition. The *input automaton* of a transducer is the automaton obtained by ignoring the
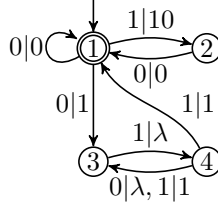
Figure 2: An unambiguous transducer

output label of each transition. The input automaton of the transducer pictured in Figure 2 is pictured in Figure 1. A transducer is called unambiguous if its input automaton is unambiguous.

A *finite run* in $\mathcal{T}$ is a finite sequence of consecutive transitions,

$$ q_0 \xrightarrow{a_1|v_1} q_1 \xrightarrow{a_2|v_2} q_2 \cdots q_{n-1} \xrightarrow{a_n|v_n} q_n $$

Its *input* and *output labels* are the words $a_1 a_2 \cdots a_n$ and $v_1 v_2 \cdots v_n$ respectively.

An *infinite run* in $\mathcal{T}$ is an infinite sequence of consecutive transitions,

$$ q_0 \xrightarrow{a_1|v_1} q_1 \xrightarrow{a_2|v_2} q_2 \xrightarrow{a_3|v_3} q_3 \cdots $$

Its *input* and *output labels* are the sequences of symbols $a_1 a_2 a_3 \cdots$ and $v_1 v_2 v_3 \cdots$ respectively.

If $\mathcal{T}$ is a unambiguous transducer, each sequence x is the input label of at most one accepting run in $\mathcal{T}$. When this run does exist, its output is denoted by $\mathcal{T}(x)$. We say that a unambiguous transducer $\mathcal{T}$ *preserves normality* if for each normal word x, $\mathcal{T}(x)$ is also normal.

An automaton (respectively transducer) is said to be *trim* if each state occurs in an accepting run. Automata and transducers are always assumed to be trim since useless states can easily be removed.

We end this section by stating very easy but useful facts about unambiguous automata. If $\langle Q, A, B, \Delta, I, F \rangle$ is an unambiguous automaton then each automaton $\langle Q, A, B, \Delta, \{q\}, F \rangle$ obtained by taking state q as initial state is also unambiguous. Similarly, removing states or transitions from an unambiguous automaton yields an unambiguous automaton. Combining these two facts gives that each strongly connected component, seen as an automaton; of an unambiguous automaton is still an unambiguous automaton.

2.3 Weighted Automata

We now introduce weighted automata. In this paper we only consider weighted automata whose weights are rational numbers with the usual addition and multiplication (see [13, Chap. III] for a complete introduction).

A *weighted automaton* $\mathcal{A}$ is a tuple $\langle Q, B, \Delta, I, F \rangle$, where Q is the state set, B is the alphabet, $I : Q \to \mathbb{Q}$ and $F : Q \to \mathbb{Q}$ are the functions that assign to each state an initial and a final weight and $\Delta : Q \times B \times Q \to \mathbb{Q}$ is a function that assigns to each transition a weight.

4

As usual, the weight of a run is the product of the weights of its transitions times the initial weight of its first state and times the final weight of its last state. Furthermore, the weight of a word $w \in B^*$ is the sum of the weights of all runs with label w and it is denoted $\text{weight}_{\mathcal{A}}(w)$.
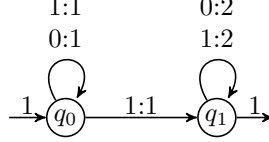

Figure 3: A weighted automaton

A transition $p \xrightarrow{a} q$ with weight x is pictured $p \xrightarrow{a:x} q$. Non-zero initial and final weights are given over small incoming and outgoing arrows. A weighted automaton is pictured in Figure 3. The weight of the run $q_0 \xrightarrow{1} q_1 \xrightarrow{0} q_1 \xrightarrow{1} q_1 \xrightarrow{0} q_1$ is $1 \cdot 1 \cdot 2 \cdot 2 \cdot 2 \cdot 1 = 8$. The weight of the word $w = 1010$ is $8 + 2 = 10$. More generally the weight of a word $w = a_1 \cdots a_k$ is the integer $n = \sum_{i=1}^{k} a_i 2^{k-i}$ (w is a binary expansion of n with possibly some leading zeros).

3   Results

We now state the main results of the paper. The first one states that when a transducer is strongly connected, unambiguous and complete, the frequency of each finite word w in the output of a run with a normal input label is given by a weighted automaton over $\mathbb{Q}$. The second one states that it can be checked whether an unambiguous transducer preserves normality.

Theorem 1. *Given an unambiguous and strongly connected transducer, there exists a weighted automaton $\mathcal{A}$ such that for each normal sequence x in the domain of $\mathcal{T}$ and for any finite word w, $\text{freq}(\mathcal{T}(x), w)$ is equal to $\text{weight}_{\mathcal{A}}(w)$.*

Furthermore, the weighted automaton $\mathcal{A}$ can be computed in cubic time with respect to the size of the transducer $\mathcal{T}$.

Theorem 1 only deals with strongly connected transducers, but Proposition 10 deals with the general case by showing that it suffices to apply Theorem 1 to some strongly connected components to check preservation of normality.

To illustrate Theorem 1 we give in Figure 4 two weighted automata which compute the frequency of each finite word w in $\mathcal{T}(x)$ for a normal input x and the transducer $\mathcal{T}$ pictured in Figure 2. The leftmost one is obtained by the procedure described in the next section. The rightmost one is obtained by removing useless states from the leftmost one.

Theorem 2. *It can be decided in cubic time whether an unambiguous transducer preserves normality or not.*

4   Adjacency matrix of the automaton

In this section, we introduce the adjacency matrix of an automaton. This matrix is particularly useful when the automaton is strongly connected and unambigu-
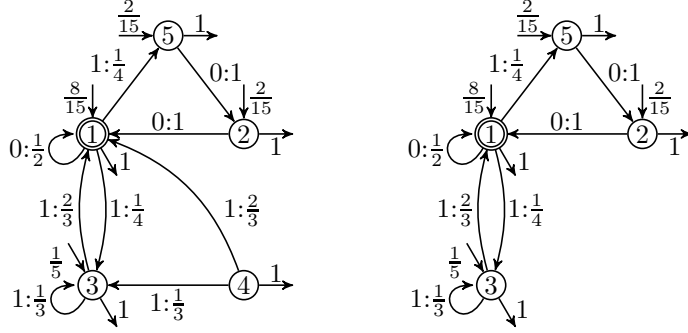
Figure 4: Two weighted automata

ous. Its spectral radius characterizes the fact that the automaton does accept or not a normal sequence as stated in Proposition 4 below.

Let $\mathcal{A}$ be an automaton with state set Q. The *adjacency matrix* of $\mathcal{A}$ is the $Q \times Q$-matrix M defined by $M_{p,q} = \#\{a \in A : p \xrightarrow{a} q\}/\#A$. Its entry $M_{p,q}$ is thus the number of transitions from p to q divided by the cardinality of the alphabet A. The factor $1/\#A$ is just a normalization factor to compare the spectral radius of this matrix to 1 rather than to the cardinality of the alphabet. By a slight abuse of notation, the spectral radius of the adjacency matrix, is called the spectral radius of the automaton.

The adjacency matrix of the unambiguous automaton pictured in Figure 1 is the matrix M given by

$$M = \frac{1}{2} \begin{pmatrix} 1 & 1 & 1 & 0 \\ 1 & 0 & 0 & 0 \\ 0 & 0 & 0 & 1 \\ 1 & 0 & 2 & 0 \end{pmatrix}$$

It can be checked that the spectral radius of this matrix is 1.

We implicitly suppose that the automaton $\mathcal{A}$ has at least one initial state and one final state. Otherwise, no sequence is accepted by it and nothing interested can be said about it. For each state q, let F_q be the *future* set, that is the set $F_q = \{x : q \xrightarrow{x} \infty\}$ of accepted sequences if q is taken as the only initial state. Let α_q be the measure of the set F_q. Note that the sum $\sum_{q \in Q} \alpha_q$ might be greater than 1 because the sets F_q might not be pairwise disjoint. We claim that the vector $\alpha = (\alpha_q)_{q \in Q}$ satisfies $M\alpha = \alpha$. This equality means that either α is the zero vector or that α is a right eigenvector of M for the eigenvalue 1. This equality comes from the following relations between the sets F_q. For each state p, one has

$$F_p = \biguplus_{p \xrightarrow{a} q} aF_q$$

where the symbol $\uplus$ stands for the union of pairwise disjoint sets. The fact that F_p is equal to the union of the sets aF_q for a and q ranging over all possible transitions $p \xrightarrow{a} q$ is true in any automaton accepting sequences. Furthermore, the unambiguity of $\mathcal{A}$ implies that the sets aF_q for different pairs (a, q) must be

pairwise disjoint. Therefore, for each state p,

$$\alpha_p = \mu(F_p) = \sum_{p \xrightarrow{a} q} \mu(aF_q) = \frac{1}{\#Q} \sum_{p \xrightarrow{a} q} \alpha_q.$$

By definition, the adjacency matrix is non-negative. By the Perron-Frobenius theorem, its spectral radius must be one of its eigenvalues. The following lemma states that if the automaton $\mathcal{A}$ is strongly connected and unambiguous, then the spectral radius of M is less than 1.

Lemma 3. *Let $\mathcal{A}$ be a strongly connected and unambiguous automaton. The maximum eigenvalue of its adjacency matrix M is less than 1.*

Proof. Consider the matrix $\#A \cdot M$. Its (p, q)-entry is the number of transitions from p to q. It follows that the (p, q)-entry of $(\#A \cdot M)^n$ is the number of runs of length n from p to q. Since $\mathcal{A}$ is unambiguous, each finite word is the label of at most one run from p to q. This yields that the entry $(\#A \cdot M)^n_{p,q}$ is bounded by the number $(\#A)^n$ of words of length n and that each entry of M^n is bounded by 1.

Since the automaton $\mathcal{A}$ is strongly connected, the matrix M is positive and irreducible. Let λ be its spectral radius which is a positive real number. Suppose that the period of M is the positive integer p. By Theorem 1.4 in [16], The matrix M^p can be decomposed as diagonal blocks of primitive matrices and at least one of this block M' has λ^p as eigenvalue. For a positive matrix M', there exists, by Theorem 1.2 in [16], a constant K such that each entry of the matrix M'^n satisfies $\lambda^n/K \leqslant M'^n_{p,q} \leqslant K\lambda^n$. This proves that $\lambda \leqslant 1$. $\square$

By the previous lemma, the spectral radius of the adjacency matrix of an unambiguous automaton is less than 1. The following proposition states when it is equal to 1 or strictly less than 1.

Proposition 4. *Let $\mathcal{A}$ be a strongly connected and unambiguous automaton and let λ be the spectral radius of its adjacency matrix. If $\lambda = 1$ then $\mathcal{A}$ accepts at least one normal sequence and each number α_q is positive. If $\lambda < 1$, then $\mathcal{A}$ accepts no normal sequence and each number α_q is equal to zero.*

Proof. Let M be the adjacency matrix of $\mathcal{A}$. Suppose first that its spectral radius satisfies $\lambda < 1$. The number of runs of length n is equal to the sum $(\#A)^n \sum_{p,q \in Q} M^n_{p,q}$ where $M^n_{p,q}$ is the (p, q)-entry of M^n. By the Perron-Frobenius theorem, there exists a constant K such that $M^n_{p,q} \leqslant K\lambda^n$ for each $p, q \in Q$. It follows that the number of words which are the label of some run in $\mathcal{A}$ is less than $K(\#Q)^2(\lambda\#A)^n$. For n great enough, $K(\#Q)^2\lambda^n$ is strictly less than 1 and some word of length n is the label of no run in $\mathcal{A}$. This implies that no normal sequence can be accepted by $\mathcal{A}$. The equality $M\alpha = \alpha$ shows that $\alpha = 0$ since 1 is not an eigenvalue of M.

We now suppose that the spectral radius λ of M is 1. The entropy of the sofic shift defined by $\mathcal{A}$ is $\log_2 \#A$, each finite word is the label of at least one run in $\mathcal{A}$. Otherwise, the entropy of the sofic shift would be strictly less than $\log_2 \#A$. Let $x = a_1a_2a_3 \cdots$ be a normal sequence. Each prefix $a_1 \cdots a_n$ of x is the label of a run in $\mathcal{A}$. By extraction, we get a run whose label is x. Note that this run might be neither initial nor final. To get an initial run we consider the new sequence ux where u is the label of a run from an initial state

to the starting state of the run labelled by x. To get a final run, we insert in x at positions of the form 2^k a word of length at most $2\#Q$ to make a small detour to a final state of $\mathcal{A}$. Since the inserted blocks have bounded lengths and they are inserted at sparse positions, the obtained sequence is still normal. We now prove that α is positive. We claim that almost all sequence are the label of a run which visiting infinitely often each state. It suffices to prove that each state p, almost all sequences are the label of a run visiting infinitely often p. Let $\mathcal{A}'$ be the automaton obtained by removing all transitions leaving p. By Theorem 1.5e in [16], the spectral radius of the adjacency matrix of $\mathcal{A}'$ is strictly less than 1. Therefore, by the previous case, the measures α'_q of the sets $F'_q = \{x : q \xrightarrow{x} \infty \text{ in } \mathcal{A}'\}$ are equal to zero. This proves that the set of sequences which are the label of a run never visiting p has measure 0. By the same reasoning, it can be shown that the set of sequences which are the label of a run visiting p finitely many times has also measure 0. This proves the claim. This shows that at least one entry of α must be positive. Since $M\alpha = \alpha$ and M is irreducible, all entries of α are positive. $\qquad\square$

The spectral radius of the adjacency matrix of the automaton pictured in Figure 1 is 1. The vector α is given by $\alpha_1 = \alpha_4 = 2/3$ and $\alpha_2 = \alpha_3 = 1/3$.

Suppose that the automaton $\mathcal{A}$ is strongly connected and unambiguous and that the spectral radius of its adjacency matrix M is 1. The vector α is then a right eigenvector of M for the eigenvalue 1. There is also a left eigenvector $\pi = (\pi_q)_{q \in Q}$ for the eigenvalue 1. This vector π is stricly positive and we normalize it in such a way that $\sum_{q \in Q} \pi_q \alpha_q = 1$.

The left eingenvector π of the adjacency matrix of the automaton pictured in Figure 1, normalized as explained just above, is given by $\pi_1 = \pi_3 = 2/3$ and $\pi_2 = \pi_4 = 1/3$. The vector of $(\pi_q \alpha_q)_{q \in Q}$ is then given $\pi_1 \alpha_1 = 4/9$, $\pi_2 \alpha_2 = \pi_3 \alpha_3 = 2/9$ and $\pi_4 \alpha_4 = 1/9$.

Figure 5: Unambiguous automata

We would like to emphasize that that the adjacency matrix M is not sufficient to compute the vector α. Said differently, two automata with the same adjacency matrix may have different vectors α. Consider the two automata pictured in Figure 5. The leftmost one is deterministic whereas the rightmost one is reverse deterministic. Both automata have the same matrix $M = \frac{1}{2}\left(\begin{smallmatrix} 1 & 1 \\ 1 & 1 \end{smallmatrix}\right)$ as adjacency matrix. For the leftmost automaton the sets F_1 and F_2 are both equal to $\{0,1\}^{\mathbb{N}}$ and thus $\alpha_1 = \alpha_2 = 1$. For the leftmost automaton the sets F_1 and F_2 are respectively equal to $0\{0,1\}^{\mathbb{N}}$ and $1\{0,1\}^{\mathbb{N}}$ and thus $\alpha_1 = \alpha_2 = 1/2$.

Note however that since α is the eigenvector of the irreducible matrix M for its Perron-Frobenius eigenvalue, it is unique up to a multiplicative factor. This means that the ratios α_q/α_p can be computed from the matrix M.

5   Markov chain of an unambiguous automaton

In this section, we introduce a Markov chain associated with an unambiguous automaton. The use of the ergodic theorem applied to this Markov chain is the main ingredient in the proof of Theorem 1. Let $\mathcal{A}$ be a strongly connected and unambiguous automaton and let p be one of its states. We also suppose that the spectral radius of its adjacency matrix M is 1. By Proposition 4, the measure α_q of each set $F_q = \{x : q \xrightarrow{x} \infty\}$ is non-zero.

We define a stochastic process $(X_n)_{n \geqslant 0}$ as follows. Its sample set is the set $F_p \subseteq A^{\mathbb{N}}$ equipped with the uniform measure μ. For each sequence $x = x_1 x_2 x_3 \cdots$ in F_p, there exists a unique accepting run

$$p = q_0 \xrightarrow{x_1} q_1 \xrightarrow{x_2} q_2 \xrightarrow{x_3} q_3 \cdots$$

The process is defined by setting $X_n(x) = q_n$ for each $x \in F_p$. The following proposition states the main propoerty of this process.

Proposition 5. *The process $(X_n)_{n \geqslant 0}$ is a Markov chain.*

Proof. We prove that this process is actually a Markov chain. A sequence x satisfies $X_n(x) = q_n$ if and only if, when factorizing x as $x = wy$ with $w = x[1{:}n]$, the word w is the label of a finite run $p \xrightarrow{w} q_n$ and the sequence y belongs to the set F_q. This remark allows us to compute the probability that $X_n = q_n$ for a given state q_n.

$$\mathrm{Prob}(X_n = q_n) = \mu(\{w \in A^n : p \xrightarrow{w} q_n\})\alpha_{q_n}$$
$$= \#\{w \in A^n : p \xrightarrow{w} q_n\}\alpha_{q_n}/(\#A)^n$$

A similar reasoning allows us to compute the probability that $X_n = q_n$ and $X_{n+1} = q_{n+1}$ for two given states q_n and q_{n+1}.

$$\mathrm{Prob}(X_{n+1} = q_{n+1}, X_n = q_n) = \mu(\{wa \in A^{n+1} : p \xrightarrow{w} q_n \xrightarrow{a} q_{n+1}\})\alpha_{q_{n+1}}$$

Using the definition of conditional probability, we get

$$\mathrm{Prob}(X_{n+1} = q_{n+1} | X_n = q_n) = \frac{\#\{a \in A : q_n \xrightarrow{a} q_{n+1}\}\alpha_{q_{n+1}}}{(\#A)\alpha_{q_n}}$$

Let $q_0, \ldots, q_n$ be $n+1$ states of the automaton $\mathcal{A}$ such that $q_0 = p$. A sequence x satisfies $X_n(x) = q_n, \ldots, X_0(x) = q_0$ if and only if the sequence x can be factorized $x = wx'$ where the word $w = a_1 \cdots a_n$ is the prefix of length n of x, there is a finite run $q_0 \xrightarrow{a_1} q_1 \cdots q_{n-1} \xrightarrow{a_n} q_n$ and x' belongs to the set F_{q_n}.

$$\mathrm{Prob}(X_n = q_n, \ldots, X_0 = q_0) = \mu(\{a_1 \cdots a_n \in A^n : q_0 \xrightarrow{a_1} q_1 \cdots q_{n-1} \xrightarrow{a_n} q_n\})\alpha_{q_n}$$

Using again the definition of conditional probability, we get

$$\mathrm{Prob}(X_{n+1} = q_{n+1} | X_n = q_n, \ldots, X_0 = q_0) = \frac{\#\{a \in A : q_n \xrightarrow{a} q_{n+1}\}\alpha_{q_{n+1}}}{(\#A)\alpha_{q_n}}$$

Since $\mathrm{Prob}(X_{n+1} = q_{n+1} | X_n = q_n)$ and $\mathrm{Prob}(X_{n+1} = q_{n+1} | X_n = q_n, \ldots, X_0 = q_0)$ have the same value, the process is indeed a Markov chain. $\qquad\square$

Let P the $Q \times Q$-matrix of probabilities for the introduced Markov chain. For each states $p, q \in Q$, the (p, q)-entry of P is given by $P_{p,q} = \#\{a \in A : p \xrightarrow{a} q\}\alpha_q/(\#A)\alpha_p$. Note that the matrix P and the adjacency matrix M of $\mathcal{A}$ are related by the equalities $P_{p,q} = M_{p,q}\alpha_q/\alpha_p$ for each states $p, q \in Q$. We claim that the stationary distribution of the stochastic matrix P is the vector $(\pi_q\alpha_q)_{q\in Q}$ where $\pi = (\pi_q)_{q\in Q}$ is the left eigenvector of the matrix M for the eigenvalue 1. Let us recall that π has been normalized such that $\sum_{q\in Q} \pi_q\alpha_q = 1$.

$$\sum_{p\in Q} \pi_p\alpha_p P_{p,q} = \alpha_q \sum_{p\in Q} \pi_p M_{p,q}$$
$$= \pi_q\alpha_q$$

Runs are defined as sequences of consecutive transitions, and can be considered as words over the alphabet made of all transitions. Therefore, the notion of *frequency* $\text{freq}(\rho, \gamma)$ of a finite run γ in an infinite run ρ is defined as in Section 2. Note that $\text{freq}(\rho, \gamma)$ is a limit and might not exist. As a run can merely be regarded as a sequence of states, $\text{freq}(\rho, q)$ is defined similarly when q is a state. Note that $\text{freq}(\rho, q)$ could equivalently be defined as the sum of all $\text{freq}(\rho, \tau)$ where τ ranges over all transitions (seen as runs of length 1) leaving q.

The application of the ergodic theorem to the previous Markov chain is used to prove the following proposition. It states that in a run whose label is a normal sequence, each state has a limiting frequency and that this frequency is given by the stationary distribution $(\pi_q\alpha_q)_{q\in Q}$. This statement is an extension to unambiguous automata of Lemma 4.5 in [15] which is only stated for deterministic automata.

Proposition 6. *Let $\mathcal{A}$ be a strongly connected and unambiguous automaton such that the spectral radius of its adjacency matrix is 1. Let ρ be an accepting run whose label is a normal sequence. Then, for any state r*

$$\lim_{n\to\infty} \frac{|\rho[1{:}n]|_r}{n} = \pi_r\alpha_r$$

where $\rho[1{:}n]$ is the finite run made of the first n transitions of ρ.

Note that the result of Proposition 6 implies that the frequencies of states do not depend on the input as long as this input is normal. Note also that the result is void if the spectral radius of the adjacency matrix is less than 1 because, by Proposition 4, no accepting run is labeled by a normal sequence. This assumption could be removed because the statement remains true but this is our choice to mention explicitly the assumption for clarity.

The proof of the proposition is based on the following lemma. Since the automaton $\mathcal{A}$ in unambiguous, there is, for two given states p and q and a given word w, a unique run from p to q labeled by w. This run is written $p \xrightarrow{w} q$ as usual.

Lemma 7. *Let $\mathcal{A}$ be a strongly connected and unambiguous automaton such that the spectral radius of its adjacency matrix is 1. Let r be a fixed state of $\mathcal{A}$. For any positive real numbers $\delta, \varepsilon > 0$, there exists an integer n such that for each integer $k \geqslant n$,*

$$\#\{w \in A^k : \exists p, q \in Q^2 \ ||p \xrightarrow{w} q|_r/k - \pi_r\alpha_r| > \delta\} < \varepsilon(\#A)^k$$

Note that for a triple (p, w, q), the finite run $p \xrightarrow{w} q$ might not exist. When we write $||p \xrightarrow{w} q|_r/k - \pi_r \alpha_r| > \delta$, it should be understood as follows. The run $p \xrightarrow{w} q$ does exist and it satisfies $||p \xrightarrow{w} q|_r/k - \pi_r \alpha_r| > \delta$.

Proof. Since there are finitely many pairs (p, q) in Q^2, it suffices to prove

$$\#\{w \in A^k : ||p \xrightarrow{w} q|_r/k - \pi_r \alpha_r| > \delta\} < \varepsilon (\#A)^k/(\#Q)^2$$

for each pair (p, q) in Q^2. Therefore, we assume that a pair (p, q) is fixed. We consider the Markov chain $(X_n)_{n \geqslant 0}$ introduced above with initial state p. This Markov chain is irreducible because $\mathcal{A}$ is strongly connected. We apply the ergodic theorem for Markov chains [3, Thm 4.1] to the function $f = \mathbb{1}_r$ defined by $\mathbb{1}_r(s) = 1$ if $s = r$ and $\mathbb{1}_r(s) = 0$ otherwise. It follows that $\lim_{n \to \infty} S_n = \pi_r \alpha_r$ for almost all sequences in F_p where $S_n = \frac{1}{n} \sum_{i=1}^n \mathbb{1}_r(X_i)$. The positive numbers δ and ε being fixed, there is an integer n such that, for each $k \geqslant n$, the measure of the set $\{x : |S_k - \pi_r \alpha_r| > \delta\}$ is less than ε. Consider now a set F of the form $F = wF_q$ where w is a word of length $k \geqslant n$ and q is the state that has been fixed. The measure of F is $\alpha_q/(\#A)^k$. If the run $p \xrightarrow{w} q$ does exist, then S_k is constant on the set F because the number of occurrences of r in the first k positions of the run only depends on the finite run $p \xrightarrow{w} q$. It follows that the number of words w of length k such that $p \xrightarrow{w} q$ does exist and satisfies $||p \xrightarrow{w} q|_r/k = \pi_r \alpha_r| > \delta$ is bounded by $\varepsilon (\#A)^k/\alpha_q$. The result is then obtained by replacing ε by $\varepsilon \min\{\alpha_q : q \in Q\}$ which is positive because all entries of α are positive. $\square$

We now come to the proof of Proposition 6.

Proof of Proposition 6. Let ρ be an accepting run in $\mathcal{A}$ whose label is a normal sequence x. Since $\sum_{q \in Q} \pi_q \alpha_q = 1$, it suffices to prove that $\liminf_{n \to \infty} |\rho[1{:}n]|_r/n \geqslant \pi_r \alpha_r$ for each state r. We fix an arbitrary positive real number ε. Applying Lemma 7 with $\delta = \varepsilon$, we get an integer k such that the set $B \subseteq A^k$ defined by

$$B = \{w \in A^k : \exists p, q \in Q^2 \ ||p \xrightarrow{w} q|_r/k - \pi_r \alpha_r| > \varepsilon\}$$

satisfies $\#B < \varepsilon(\#A)^k$. The run ρ is then factorised

$$\rho = q_0 \xrightarrow{w_1} q_1 \xrightarrow{w_2} q_2 \xrightarrow{w_3} q_3 \cdots$$

where each word w_i has length k and $x = w_1 w_2 w_3 \cdots$ is a factorization of x in blocks of length k. Since the sequence x is normal, there is an integer N such that for each $n \geqslant N$ and each word w of length k, the cardinality of the set $\{1 \leqslant i \leqslant n : w_i = w\}$ satisfies $\#\{1 \leqslant i \leqslant n : w_i = w\} \geqslant n(1 - \varepsilon)/(\#A)^k$.

$$
\begin{aligned}
\liminf_{n \to \infty} \frac{|\rho[1{:}n]|_r}{n} &= \liminf_{n \to \infty} \frac{1}{nk} \sum_{i=1}^n |q_{i-1} \xrightarrow{w_i} q_i|_r \\
&\geqslant \liminf_{n \to \infty} \sum_{w \in A^k} \frac{\#\{1 \leqslant i \leqslant n : w_i = w\}}{n} \min_{p, q \in Q^2} \frac{|p \xrightarrow{w} q|_r}{k} \\
&\geqslant \sum_{w \notin B} (1 - \varepsilon)(\pi_r \alpha_r - \varepsilon)/(\#A)^k \\
&\geqslant (1 - \varepsilon)^2 (\pi_r \alpha_r - \varepsilon)
\end{aligned}
$$

Since this is true for any $\varepsilon > 0$, $\liminf_{n \to \infty} |\rho[1{:}n]|_r/n \geqslant \pi_r \alpha_r$ holds for each state r. This completes the proof of the proposition. $\square$

11

Proposition 6 states that each state has a frequency in a run whose label is normal. This result can be extended to finite runs as follows. The stochastic matrix P and its stationary distribution $(\pi_q \alpha_q)_{q \in Q}$ induce a canonical distribution on finite runs in the automaton $\mathcal{A}$. This distribution is defined as follows.

$$\mathrm{Prob}(q_0 \xrightarrow{a_1} q_1 \cdots q_{n-1} \xrightarrow{a_n} q_n) = \frac{\pi_{q_0} \alpha_{q_n}}{(\#A)^n}.$$

We claim that for each integer n, this is indeed a distribution on runs of length n. This means that

$$\sum_{p,q \in Q^2, w \in A^n} \mathrm{Prob}(p \xrightarrow{w} q) = 1.$$

For each symbol $a \in A$, define the $Q \times Q$-matrix P_a by

$$(P_a)_{p,q} = \begin{cases} \frac{\alpha_q}{(\#A)\alpha_p} & \text{if } p \xrightarrow{a} q \text{ is a transition of } \mathcal{A} \\ 0 & \text{otherwise.} \end{cases}$$

Note that the stochastic matrix P is equal to the sum $\sum_{a \in A} P_a$. Let $w = a_1 \cdots a_n$ be a word of length n. Let us write P_w for the product $P_{a_1} \cdots P_{a_n}$. This notation is consistent with the notation P_a for words of length 1. The (p,q)-entry of the matrix $P_w = P_{a_1} \cdots P_{a_n}$ is equal to $\alpha_q/(\#A)^n \alpha_p = \mathrm{Prob}(p \xrightarrow{w} q)/\pi_p \alpha_p$ if the run $p \xrightarrow{w} q$ does exist and to 0 otherwise.

$$\begin{aligned} \sum_{p,q \in Q^2, w \in A^n} \mathrm{Prob}(p \xrightarrow{w} q) &= \sum_{p,q \in Q^2} \pi_p \alpha_p \sum_{w \in A^n} (P_w)_{p,q} \\ &= \sum_{p,q \in Q^2} \pi_p \alpha_p (P^n)_{p,q} \\ &= \sum_{p \in Q} \pi_p \alpha_p = 1 \end{aligned}$$

The following proposition extends to finite runs the statement of Proposition 6 about states.

Proposition 8. *Let $\mathcal{A}$ be a strongly connected and unambiguous automaton such that the spectral radius of its adjacency matrix is 1. Let ρ be an accepting run whose label is a normal sequence. For any finite run $\gamma = q_0 \xrightarrow{a_1} q_1 \cdots q_{n-1} \xrightarrow{a_n} q_n$ of length n, one has*

$$\lim_{n \to \infty} \frac{|\rho[1:n]|_\gamma}{n} = \mathrm{Prob}(\gamma) = \frac{\pi_{q_0} \alpha_{q_n}}{(\#A)^n}$$

where $\rho[1:n]$ is the finite run made of the first n transitions of ρ.

Proof. We now define an automaton whose states are the runs of length n in $\mathcal{A}$. We let $\mathcal{A}^n$ denote the automaton whose state set is $\{p \xrightarrow{w} q : p, q \in Q, w \in A^n\}$ and whose set of transitions is defined by

$$\left\{ \gamma \xrightarrow{a} \gamma' : \begin{array}{l} \gamma = p \xrightarrow{b} p' \xrightarrow{w} q \\ \gamma' = p' \xrightarrow{w} q \xrightarrow{a} q' \end{array} \text{ with } a, b \in A \text{ and } w \in A^{n-1} \right\}$$

The Markov chain associated with the automaton $\mathcal{A}^n$ is called the *snake* Markov chain. See Problems 2.2.4, 2.4.6 and 2.5.2 (page 90) in [3] for more details. It is pure routine to check that the stationary distribution ξ of $\mathcal{A}^n$ is given by $\xi_{p \xrightarrow{w} q} = \mathrm{Prob}(p \xrightarrow{w} q) = \pi_p \alpha_q/(\#A)^n$ for each finite run $p \xrightarrow{w} q$ of length n in $\mathcal{A}$. To prove the statement, apply Proposition 6 to the automaton $\mathcal{A}^n$. $\square$

It should be pointed out that the distribution on finite path which is defined above is the Parry measure of the edge shift of the automaton. This shift is the shift of finite type whose symbols are the edges of the automaton [11, Thm 6.2.20].

Let γ be a finite run whose first state is p and let ρ be an infinite run. We call *conditional frequency* of γ in ρ the ratio $\mathrm{freq}(\rho, \gamma)/\mathrm{freq}(\rho, p)$. It is defined as soon as both frequencies $\mathrm{freq}(\rho, \gamma)$ and $\mathrm{freq}(\rho, p)$ do exist. The corollary of Propositions 6 and 8 is the following.

Corollary 9. *The conditional frequency of a finite run $p \xrightarrow{w} q$ of length n in an accepting run whose label is normal is $\alpha_q/(\#A)^n \alpha_p$.*

6 Algorithms and Proofs

In this section we provide the proofs for Theorems 1 and 2. The proofs are organized in three parts. First, the transducer $\mathcal{T}$ is normalized into another transducer $\mathcal{T}'$ realizing the same function. Then this latter transducer is used to define a weighted automaton $\mathcal{A}$. Second, the proof that the construction of $\mathcal{A}$ is correct is carried out. Third, the algorithms computing $\mathcal{A}$ and checking whether $\mathcal{T}$ preserves normality or not are given.

By Proposition 10 below, it suffices to analyze preservation of normality in some of the strongly connected components.

Proposition 10. *An unambiguous transducer $\mathcal{T}$ preserves normality if and only each strongly connnected component of $\mathcal{T}$ with a final state and spectral radius 1 preserves normality.*

Proof. Let ρ be an accepting run of $\mathcal{T}$ whose label is a normal sequence. This run ends in some strongly connected C, that is, all states which are visited infinitely often by the run belong to the same strongly connected component C. This component C must contain a final state because ρ is accepting and by Proposition 4, its spectral radius must be one.

Conversely, let C be a strongly component with a final state and spectral radius 1. By Proposition 4, there is a final run ρ contained in C and whose label is a normal sequence x. Suppose that ρ starts from state q in C. Let u be the label of a run from an initial state to q. The unique accepting run labeled by ux ends in C and C must preserve normality. $\qquad\square$

In what follows we only consider strongly connected transducers. Propositions 6 and 8 have the following consequence. Let $\mathcal{T}$ be a unambiguous and strongly connected transducer. If each transition has an empty output label, the output of any run is empty and then $\mathcal{T}$ does not preserve normality. Therefore, we assume that transducers have at least one transition with a non empty output label. By Propositions 6 and 8, this transition is visited infinitely often if the input is normal because the stationary distribution $(\pi_q \alpha_q)_{q \in Q}$ is positive. This guarantees that if the input sequence is normal, then the output sequence is infinite and $\mathcal{T}(x)$ is well-defined.

Note that the output labels of the transitions in $\mathcal{T}$ from Theorem 1 may have arbitrary lengths. We first describe the construction of an equivalent transducer $\mathcal{T}'$ such that all output labels in $\mathcal{T}'$ have length at most 1. We call

13

this transformation *normalization* and it consists in *replacing* each transition $p \xrightarrow{a|v} q$ in $\mathcal{T}$ such that $|v| \geqslant 2$ by n transitions:

$$p \xrightarrow{a|b_1} q_1 \xrightarrow{\lambda|b_2} q_2 \cdots q_{n-1} \xrightarrow{\lambda|b_n} q$$

where $q_1, q_2, \ldots, q_{n-1}$ are new states and $v = b_1 \cdots b_n$. We refer to p as the parent of $q_1, \cdots, q_{n-1}$.
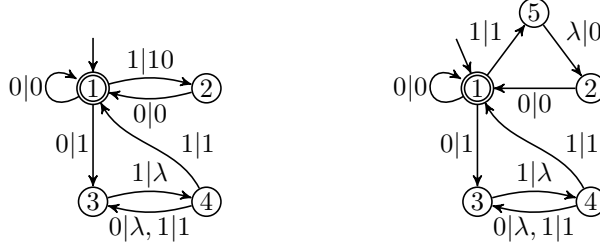

Figure 6: The transducer $\mathcal{T}$ and its normalization $\mathcal{T}'$

To illustrate the construction, the normalized transducer obtained from the transducer of Figure 2 is pictured in Figure 6. State 5 has been added to split the transition $1 \xrightarrow{1|10} 2$ into the finite run $1 \xrightarrow{1|1} 5 \xrightarrow{\lambda|0} 2$.

The main property of $\mathcal{T}'$ is stated in the following lemma which follows directly from the definition of normalization.

Lemma 11. *Both transducers $\mathcal{T}$ and $\mathcal{T}'$ realize the same function, that is, $\mathcal{T}(x) = \mathcal{T}'(x)$ for each sequence x in the domain of $\mathcal{T}$.*

From the normalized transducer $\mathcal{T}'$ we construct a weighted automaton $\mathcal{A}$ with the same state set as $\mathcal{T}'$. For all states p, q and for every symbol $b \in B$ the transition $p \xrightarrow{b} q$ is defined in $\mathcal{A}$. To assign weights to transitions in $\mathcal{A}$, we auxiliary assign weights to transitions in $\mathcal{T}'$ as follows. Each transition in $\mathcal{T}'$ of the form $p \xrightarrow{a|v} q$ where v is either a symbol or the empty word has weight $\alpha_q/(\#A)\alpha_p$. Each transition in $\mathcal{T}'$ of the form $p \xrightarrow{\lambda|b} q$ (and starting from a newly added state) has weight 1. The sum of weights of transitions starting from each state p is 1. Indeed, if p is a state of $\mathcal{T}$, the weights of transitions leaving p are the entries in line indexed by p of the stochastic matrix P. If p is a newly added state, there is only one transition leaving p which has weight 1. We now consider separately transitions that generate empty output from those that do not.

Consider the $Q \times Q$ matrix E whose (p, q)-entry is given for each pair (p, q) of states by

$$E_{p,q} = \sum_{a \in A} \text{weight}_{\mathcal{T}'}(p \xrightarrow{a|\lambda} q).$$

Let E^* be the matrix defined by $E^* = \sum_{k \geqslant 0} E^k$, where E^0 is the identity matrix. By convention, there is indeed an empty run from p to p for each state p and this run has weight 1. The entry $E^*_{p,q}$ is the sum of weights of all finite runs with empty output going from p to q. The matrix E^* can be computed because it is the solution of the linear equation $E^* = EE^* + I$ where I is the identity matrix. This proves in particular that all its entries are rational numbers.

For each symbol $b \in B$ consider the $Q \times Q$ matrix D_b whose (p,q)-entry is given for each pair (p,q) of states by

$$(D_b)_{p,q} = \sum_{a \in A \uplus \{\lambda\}} \text{weight}_{\mathcal{T}'}(p \xrightarrow{a|b} q).$$

We define the weight of a transition $p \xrightarrow{b} q$ in $\mathcal{A}$ as

$$\text{weight}_{\mathcal{A}}(p \xrightarrow{b} q) = (E^* D_b)_{p,q}. \tag{1}$$

To assign initial weights to states we consider the matrix $\hat{P} = \sum_{b \in B} E^* D_b$. It is proved below in Lemma 15 that this matrix is stochastic. The initial vector of $\mathcal{A}$ is the stationary distribution $\hat{\pi}$ of $\hat{P}$, that is, the line vector $\hat{\pi}$ such that $\hat{\pi}\hat{P} = \hat{\pi}$. We assign to each state q the initial weight $\hat{\pi}_q$. Finally we assign final weight 1 to all states.

We give below the matrices E, E^*, D_a, D_b and $\hat{P}$ and the initial vector $\hat{\pi}$ of the weighted automaton obtained from the transducer pictured in Figure 6.

Proposition 12. *The automaton $\mathcal{A}$ computes frequencies, that is, for every normal word x and any finite word w in B^*, $\text{weight}_{\mathcal{A}}(w) = \text{freq}(\mathcal{T}(x), w)$.*

The proof of the proposition requires some preliminary results.

Let us recall that a set of words L is called *prefix-free* if no word in L is a proper prefix of another word in L. As runs are defined as sequences of (consecutive) transitions, this latter definition also applies when L is a set of runs. We define $\text{freq}(\rho, \Gamma)$ when Γ is a set of finite runs as follows. Suppose that ρ is the sequence $\tau_1 \tau_2 \tau_3 \cdots$ of transitions. Then $\text{freq}(\rho, \Gamma)$ is defined by

$$\text{freq}(\rho, \Gamma) = \lim_{n \to \infty} \frac{\#\{i < n : \exists k \geqslant 0 \ \ \tau_i \cdots \tau_{i+k} \in \Gamma\}}{n}.$$

If Γ is prefix-free (not to count twice the same start position i), the following equality holds

$$\text{freq}(\rho, \Gamma) = \sum_{\gamma \in \Gamma} \text{freq}(\rho, \gamma)$$

assuming that each limit of the right-hand sum does exist. If Γ is a set of finite runs starting from the same state p, the *conditional frequency* of Γ in a run ρ is defined as the ratio between the frequency of Γ in ρ and the frequency of p in ρ, that is, $\text{freq}(\rho, \Gamma)/\text{freq}(\rho, p)$. Furthermore if Γ is prefix-free, the conditional frequency of Γ is the sum of the conditional frequencies of its elements.

Let x be a fixed normal word and let ρ and ρ' be respectively the runs in $\mathcal{T}$ and $\mathcal{T}'$ with label x. By Proposition 6, the frequency $\text{freq}(\rho, q)$ of each state q is $\pi_q \alpha_q$ where π and α and the left and right eigenvectors of the adjacency matrix of $\mathcal{T}$ for the eigen value 1. The following lemma gives the frequency of states in ρ'.

Lemma 13. *There exists a constant C such that if r is a state of $\mathcal{T}$, then $\text{freq}(\rho', r) = \text{freq}(\rho, r)/C$ and if r is newly created, then $\text{freq}(\rho', r)$ is equal to $\text{freq}(\rho', p)\alpha_q/(\#A)\alpha_p$ where r comes from the splitting of a transition $p \xrightarrow{a|v} q$ in $\mathcal{T}$.*

Proof. Observe that there is a one-to-one relation between runs labeled with normal words in $\mathcal{T}$ and in $\mathcal{T}'$. More precisely, each transition τ in ρ is replaced by $\max(1, |v_\tau|)$ transitions in ρ' (where v_τ is the output label of τ).

By Proposition 8, each transition of $\mathcal{T}$ has a frequency in ρ. The first result follows by taking $C = \sum_\tau \text{freq}(\rho, \tau) \cdot \max(1, |v_\tau|)$ where the summation is taken over all transitions τ of $\mathcal{T}$ and v_τ is implicitly the output label of τ. The second result follows from Corollary 9 stating that each transition $p \xrightarrow{a|v} q$ has a conditional frequency of $\alpha_q/(\#A)\alpha_p$ in ρ. □

For each pair (p, q) of states and each symbol $b \in B$, consider the set $\Gamma_{p,b,q}$ of runs from p to q in $\mathcal{T}'$ that have empty output labels for all their transitions but the last one, which has b as output label.

$$\Gamma_{p,b,q} = \{p \xrightarrow{a_1|\lambda} \cdots \xrightarrow{a_n|\lambda} q_n \xrightarrow{a_{n+1}|b} q : n \geqslant 0, q_i \in Q, a_i \in A \cup \{\lambda\}\}$$

and let Γ be the union $\bigcup_{p,q \in Q, b \in B} \Gamma_{p,b,q}$. Note that the set Γ is prefix-free. Therefore, the run ρ' has a unique factorization $\rho = \gamma_0 \gamma_1 \gamma_2 \cdots$ where each γ_i is a finite run in Γ and the ending state of γ_i is the starting state of γ_{i+1}. Let $(p_i)_{i \geqslant 0}$ and $(b_i)_{i \geqslant 0}$ be respectively the sequence of states and the sequence of symbols such that γ_i belongs to $\Gamma_{p_i, b_i, p_{i+1}}$ for each $i \geqslant 0$. Let us call ρ'' the sequence $p_0 p_1 p_2 \cdots$ of states of $\mathcal{T}'$.

Lemma 14. *For each state q of $\mathcal{T}'$, the frequency $\text{freq}(\rho'', q)$ does exist.*

Proof. The sequence ρ'' is a subsequence of the sequence of states in the run ρ'. An occurrence of a state q in ρ' is removed whenever the output of the previous transition is empty.

Consider the transducer $\hat{\mathcal{T}}$ obtained by splitting each state q of $\mathcal{T}$ into two states q^λ and q^o in such a way that transitions with an empty output label end in a state q^λ and other transitions end in a state q^o. Then each transition $p \xrightarrow{a|v} q$ is replaced by either the two transitions $p^\lambda \xrightarrow{a|v} q^\lambda$ and $p^o \xrightarrow{a|v} q^\lambda$ if v is empty or by the two transitions $p^\lambda \xrightarrow{a|v} q^o$ and $p^o \xrightarrow{a|v} q^o$ otherwise. The state q_0^λ becomes the new initial state and non reachable states are removed. Let $\hat{\rho}$ be the run in $\hat{\mathcal{T}}$ labeled with x. By Proposition 6, the frequencies $\text{freq}(\hat{\rho}, q^\lambda)$ and $\text{freq}(\hat{\rho}, q^o)$ do exist. Now consider the normalization $\hat{\mathcal{T}}'$ of $\hat{\mathcal{T}}$ and the run $\hat{\rho}'$ in $\hat{\mathcal{T}}'$ labeled with x. It can be shown that the frequencies $\text{freq}(\hat{\rho}', q^\lambda)$ and $\text{freq}(\hat{\rho}', q^o)$ do exist by an argument similar to the proof of Lemma 13. The sequence ρ'' is obtained from $\hat{\rho}'$ by removing each occurrence of states q^λ and keeping occurrences of states q^o. It follows that the frequency of each state does exist in ρ''. □

Proof of Proposition 12. By Corollary 9, the conditional frequency in ρ of each finite run γ of length n from p to q is $\alpha_q/(\#A)^n \alpha_p$. It follows that the conditional frequency of each finite run γ' in ρ' is equal to its assigned weight in $\mathcal{T}'$ while defining $\mathcal{A}$. By Equation (1), the weight of the transition $p \xrightarrow{b} q$ in $\mathcal{A}$ is exactly the conditional frequency of the set $\Gamma_{p,b,q}$ for each triple (p, b, q) in $Q \times B \times Q$. More generally, the product of the weights of the transitions $p_0 \xrightarrow{b_1} p_1 \cdots p_{n-1} \xrightarrow{b_n} p_n$ is equal to the conditional frequency of the set $\Gamma_{p_0, b_1, p_1} \cdots \Gamma_{p_n, b_n, p_{n+1}}$ in ρ'.

It remains to prove that the frequency of each state q in ρ'' is indeed its initial weight in the automaton $\mathcal{A}$. Let us recall that the initial vector of $\mathcal{A}$ is

the stationary distribution of the stochastic matrix $\hat{P}$ whose (p, q)-entry is the sum $\sum_{b \in B} \text{weight}_{\mathcal{A}}(p \xrightarrow{b} q)$, which is the conditional frequency of pq (as a word of length 2) in ρ''. It follows that the frequencies of states in ρ'' must be the stationary of the matrix P.

Since the frequency of a word $v = b_1 \cdots b_n$ in $\mathcal{T}'(x)$ is the same as the sum over all sequences $p_0, p_1 \ldots, p_{n+1}$ of the frequencies of $\Gamma_{p_0, b_1, p_1} \cdots \Gamma_{p_n, b_n, p_{n+1}}$ in ρ', it is the weight of the word v in the automaton $\mathcal{A}$. $\qquad\square$

Lemma 15. *The matrix $P = \sum_{b \in B} E^* D_b$ which has been used to define the initial weights is stochastic.*

Proof. It is an easy observation that if P_1 and P_2 are two square matrices with non-negative coefficients such that $P_1 + P_2$ is stochastic, then $P_1^* P_2 = \sum_{n \geqslant 0} P_1^n P_2$ is stochastic. In our case, the matrices P_1 and P_2 come from the splitting of transitions of $\mathcal{T}'$ into the ones with empty output and the ones with non-empty output. They are respectively equal to $P_1 = E$ and $P_2 = \sum_{b \in B} D_b$. $\qquad\square$

Proofs of Theorems 1 and 2. To complete the proof of Theorems 1 and 2, we exhibit an algorithm deciding in cubic time whether an input deterministic transducer preserves normality. Let $\mathcal{T}$ be an unambiguous transducer $\langle Q, A, B, \Delta, I, F \rangle$. By definition, its size is the sum $\sum_{\tau \in \Delta} |\tau|$, where the size of a single transition $\tau = p \xrightarrow{a|v} q$ is $|\tau| = |av|$. We consider the alphabets to be fixed so they are not taken into account when calculating complexity.
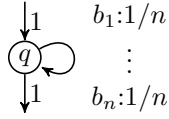

Figure 7: Weighted automaton $\mathcal{B}$ such that $\text{weight}_{\mathcal{B}}(w) = 1/(\#A)^{|w|}$

By Proposition 10, the algorithm decomposes the transducer into strongly connected components and make a list of all strongly connected components with a final state and spectral radius 1. This latter condition is checked in cubic time by computing the determinant of its adjacency matrix minus the identity matrix. For each strongly connected component in the list, the algorithm checks whether it preserves normality or not. This is achieved by computing the weighted automaton $\mathcal{A}$ and checking that the weight of each word w is $1/(\#A)^{|w|}$. This latter step is performed by comparing $\mathcal{A}$ with the weighted automaton $\mathcal{B}$ such that $\text{weight}_{\mathcal{B}}(w) = 1/(\#A)^{|w|}$. The automaton $\mathcal{B}$ is pictured in Figure 7.

> **Input:** $\mathcal{T} = \langle Q, A, B, \Delta, I, F \rangle$ an input deterministic complete transducer.
> **Output:** True if $\mathcal{T}$ preserves normality and False otherwise.
> **Procedure:**
>
>   I. Compute the strongly connected components of $\mathcal{T}$
>
>   II. For each strongly connected component S_i $\mathcal{T}$:

1. Compute the normalized transducer $\mathcal{T}'$, equivalent to S_i.
2. Use $\mathcal{T}'$ to build the weighted automaton $\mathcal{A}$:
   a. Compute the weights of the transitions of $\mathcal{A}$.
      Compute the matrix E
      Compute the matrix E^* solving $(I - E)X = I$
      For each $b \in B$, for each $p, q \in Q$:
         compute the matrix D_b
         define the transition $p \xrightarrow{b} q$ with weight $(E^* D_b)_{p,q}$.
   b. Compute the stationary distribution π of the Markov chain induced by $\mathcal{A}$.
   c. Assign initial weight $\pi[i]$ to each state i, and let final weight be 1 for all states.
3. Compare $\mathcal{A}$ against the automaton $\mathcal{B}$ using Schützenberger's algorithm [5, 14] to check whether they realize the same function.
4. If they do not compute the same function, return `False`.

III. Return `True`

Now we analyze the complexity of the algorithm. Computing recurrent strongly connected components can be done in time $O(\#Q^2) \leqslant O(n^2)$ using Kosaraju's algorithm if the transducer is implemented with an adjacency matrix [9, Section 22.5].

We refer to the size of the component S_i as n_i. The cost of normalizing the component is $O(n_i^2)$, mainly from filling the new adjacency matrix. The most expensive step when computing the transitions and their weight is to compute E^*. The cost is $O(n_i^3)$ to solve the system of linear equations. To compute the weights of the states we have $O(n_i^3)$ to solve the system of equations to find the stationary distribution. Comparing the automaton to the one computing the expected frequencies can be done in time $O(n_i^3)$ [5] since the coefficients of both automata are in $\mathbb{Q}$.

□

Conclusion

In this paper, we deal with normality which is defined through the uniform measure. Another extension is to analyze more general measures like, for instance, all Markovian measures.

Acknowledgements

References

[1] V. Becher and O. Carton. Normal numbers and computer science. In *Sequences, Groups, and Number Theory*, pages 233–269. Springer, 2018.

[2] É. Borel. Les probabilités dénombrables et leurs applications arithmétiques. *Rend. Circ. Mat. Palermo*, 27(2):247–271, 1909.

[3] P. Brémaud. *Markov Chains: Gibbs Fields, Monte Carlo Simulation, and Queues.* Springer, 2008.

[4] Y. Bugeaud. *Distribution modulo one and Diophantine approximation*, volume 193 of *Cambridge Tracts in Mathematics.* Cambridge University Press, Cambridge, 2012.

[5] A. Cardon and M. Crochemore. Détermination de la représentation standard d'une série reconnaissable. *ITA*, 14(4):371–379, 1980.

[6] O. Carton and E. Orduna. Preservation of normality by transducers. *CoRR*, abs/1904.09133, 2019.

[7] D. G. Champernowne. The construction of decimals normal in the scale of ten. *Journal of the London Mathematical Society*, 1(4):254–260, 1933.

[8] Ch. Choffrut and S. Grigorieff. *Uniformization of Rational Relations*, pages 59–71. Springer, 1999.

[9] T. H. Cormen, Ch. E. Leiserson, R. L. Rivest, and C. Stein. *Introduction to Algorithms, Third Edition.* The MIT Press, 3rd edition, 2009.

[10] R. G. Downey and D. Hirschfeldt. *Algorithmic randomness and complexity.* Theory and Applications of Computability. New York, NY: Springer. xxvi, 855 p., 2010.

[11] B. P. Kitchens. *Symbolic Dynamics.* Springer, 1998.

[12] D. Perrin and J.-É. Pin. *Infinite Words.* Elsevier, 2004.

[13] J. Sakarovitch. *Elements of Automata Theory.* Cambridge University Press, 2009.

[14] J. Sakarovitch. Rational and recognisable power series. In M. Droste, W. Kuich, and H. Vogler, editors, *Handbook of Weighted Automata*, chapter 4, pages 105—174. Springer, 2009.

[15] C. P. Schnorr and H. Stimm. Endliche Automaten und Zufallsfolgen. *Acta Informatica*, 1:345–359, 1972.

[16] E. Senata. *Non-negative Matrices ans Markov Chains.* Springer, 2006.