ARTICLE IN PRESS

Theoretical Computer Science ▮ (▮▮▮▮) ▮▮▮–▮▮▮

Contents lists available at ScienceDirect

Theoretical Computer Science

journal homepage: www.elsevier.com/locate/tcs

Entropy sensitivity of languages defined by infinite automata, via Markov chains with forbidden transitions

Wilfried Huss, Ecaterina Sava *, Wolfgang Woess

Institut für Mathematische Strukturtheorie, Technische Universität Graz, Steyrergasse 30, 8010 Graz, Austria

ARTICLE INFO

Article history:
Received 8 December 2009
Received in revised form 11 June 2010
Accepted 21 July 2010
Communicated by B. Durand

Keywords:
Formal language
Oriented graph
Infinite sofic system
Growth sensitivity
Entropy
Irreducible Markov chain
Spectral radius

ABSTRACT

A language L over a finite alphabet $\mathbf{\Sigma}$ is growth sensitive (or entropy sensitive) if forbidding any finite set of factors F of L yields a sublanguage L^F whose exponential growth rate (entropy) is smaller than that of L. Let (X, E, ℓ) be an infinite, oriented, edge-labelled graph with label alphabet $\mathbf{\Sigma}$. Considering the graph as an (infinite) automaton, we associate with any pair of vertices $x, y \in X$ the language $L_{x,y}$ consisting of all words that can be read as labels along some path from x to y. Under suitable general assumptions, we prove that these languages are growth sensitive. This is based on using Markov chains with forbidden transitions.

1. Introduction

Let $\mathbf{\Sigma}$ be a finite alphabet and $\mathbf{\Sigma}^*$ the set of all finite words over $\mathbf{\Sigma}$, including the empty word ϵ. A language L over $\mathbf{\Sigma}$ is a subset of $\mathbf{\Sigma}^*$. All our languages will be infinite. We denote by $|w|$ the length of the word w. A *factor* of a word $w = a_1 a_2 \dots a_n$ is a word of the form $a_i a_{i+1} \dots a_j$, with $1 \le i \le j \le n$. The *growth* or *entropy* of L is

$$\mathsf{h}(L) = \limsup_{n\to\infty} \frac{1}{n} \log\bigl|\{w \in L : \ |w| = n\}\bigr|.$$

For a finite, non-empty set $F \subset \mathbf{\Sigma}^+ = \mathbf{\Sigma}^* \setminus \{\epsilon\}$ consisting of factors of elements of L, we let

$$L^F = \{w \in L : \ \text{no } v \in F \text{ is a factor of } w\}.$$

The issue addressed here is to provide conditions under which, for a class of languages associated with infinite graphs, $\mathsf{h}(L^F) < \mathsf{h}(L)$. If this holds for *any* set F of *forbidden factors*, then the language L is called *growth sensitive* (or *entropy sensitive*).

Questions related to growth sensitivity have been considered in different contexts.

In *group theory*, in relation to regular normal forms of finitely generated groups, the study of growth sensitivity has been proposed by Grigorchuk and de la Harpe [9] as a tool for proving the Hopfianity of a given group or class of groups; see also [1,4].

In *symbolic dynamics*, the number $\mathsf{h}(L)$ associated with a regular language accepted by a finite automaton with suitable properties appears as the *topological entropy* of a *sofic system*; see [11, Chapters 3 and 4]. Entropy sensitivity appears as the strict inequality between the entropies of an irreducible sofic shift and a proper subshift [11, Cor. 4.4.9].

---

[*] Corresponding author. Tel.: +43 316 873 7132.
E-mail addresses: huss@finanz.math.tu-graz.ac.at (W. Huss), sava@TUGraz.at (E. Sava), woess@TUGraz.at (W. Woess).

0304-3975/$ – see front matter © 2010 Elsevier B.V. All rights reserved.
doi:10.1016/j.tcs.2010.07.020

Please cite this article in press as: W. Huss, et al., Entropy sensitivity of languages defined by infinite automata, via Markov chains with forbidden transitions, Theoretical Computer Science (2010), doi:10.1016/j.tcs.2010.07.020

Motivated by these bodies of work, Ceccherini-Silberstein and Woess [6,7,5] have elaborated practicable criteria that guarantee the growth sensitivity of *context-free* languages.

The main result of the present note can be seen as a direct extension of [11, Cor. 4.4.9] to the entropies of infinite sofic systems; see below for further comments and references.

Our basic object is an infinite oriented graph (X, E, ℓ) whose edges are labelled by elements of a finite alphabet $\mathbf{\Sigma}$. Each edge has the form $e = (x, a, y)$, where $e^- = x$ and $e^+ = y \in X$ are the initial and the terminal vertices of e, and $\ell(e) = a \in \mathbf{\Sigma}$ is its label. We will also write $x \xrightarrow{a} y$ for the edge $e = (x, a, y)$, or just $x \to y$ in situations where we do not care about the label. Multiple edges and loops are allowed, but two edges with the same end vertices must have distinct labels.

A *path* of length n in (X, E, ℓ) is a sequence $\pi = e_1 e_2 \dots e_n$ of edges such that $e_i^+ = e_{i+1}^-$, for $i = 1, 2, \dots n-1$. We say that it is a path from x to y if $e_1^- = x$ and $e_n^+ = y$. The label $l(\pi)$ of π is the word $\ell(\pi) = \ell(e_1)\ell(e_2)\dots\ell(e_n) \in \mathbf{\Sigma}^*$ that we read along the path. We also allow the empty path from x to x, whose label is the *empty word* $\epsilon \in \mathbf{\Sigma}^*$. For $x, y \in X$, denote by $\Pi_{x,y}$ the set of all paths π from x to y in (X, E, ℓ).

The languages which we consider here are

$$L_{x,y} = \{\ell(\pi) \in \mathbf{\Sigma}^* : \pi \in \Pi_{x,y}\}, \text{ where } x, y \in X.$$

That is, we can interpret the edge-labelled graph (X, E, ℓ) as an infinite automaton (labelled digraph) with initial state x and terminal state y, so that $L_{x,y}$ is the language accepted by the automaton.

We say that (X, E, ℓ) is *deterministic* if, for every vertex x and every $a \in \mathbf{\Sigma}$, there is at most one edge with initial point x and label a. Any automaton (finite or infinite) can be transformed into a deterministic one that accepts the same language, by the well-known powerset construction. See, for example, [2, Prop. 1.4.1].

As in the finite case, we need an irreducibility assumption. The graph (X, E, ℓ) is called *strongly connected* if, for every pair of vertices x, y, there is an (oriented) path from x to y. Furthermore, we say that it is *uniformly connected* if, in addition, the following holds.

- There is a constant K such that for every edge $x \to y$ there is a path from y to x with length at most K.

In the finite case, the two notions coincide, as one can take $K = |X|$. The *forward distance* $d^+(x, y)$ of $x, y \in X$ is the minimum length of a path from x to y. We write

$$\mathsf{h}(X) = \mathsf{h}(X, E, \ell) = \sup_{x,y \in X} \mathsf{h}(L_{x,y}),$$

and call this the entropy of our oriented, labelled graph. It is a well-known and easy to prove fact that, for a strongly connected graph, $\mathsf{h}(L_{x,y}) = \mathsf{h}(X)$ for all $x, y \in X$.

We also need a reasonable assumption on the set of forbidden factors.

We say that a finite set $F \subset \mathbf{\Sigma}^+$ is *relatively dense* in the graph (X, E, ℓ) if there is a constant D such that, for every $x \in X$, there are $y \in X$ and $w \in F$ such that $d^+(x, y) \le D$, and there is a path starting at y which has label w.

Note that the assumptions of uniformly connectedness and relatively denseness cannot be avoided, since they play an important role in the proof of the main result. This fails without these assumptions.

Theorem 1.1. *Suppose that (X, E, ℓ) is uniformly connected and deterministic with label alphabet $\mathbf{\Sigma}$. Let $F \subset \mathbf{\Sigma}^+$ be a finite, non-empty set which is relatively dense in (X, E, ℓ). Then*

$$\sup_{x,y \in X} \mathsf{h}(L^F_{x,y}) < \mathsf{h}(X) \quad \textit{strictly}.$$

We say that (X, E, ℓ) is *fully deterministic* if, for every $x \in X$ and $a \in \mathbf{\Sigma}$, there is precisely one edge with initial point x and label a. We remark that, in automata theory, the classical terminology is deterministic and complete, instead of fully deterministic. Since in graph theory a complete graph is one in which every pair a distinct vertices is connected by an unique edge, we shall use the notion of fully deterministic graphs throughout this paper.

Corollary 1.2. *If (X, E, ℓ) is uniformly connected and fully deterministic, then $L_{x,y}$ is growth sensitive for all $x, y \in X$.*

Indeed, in this case, for every $x \in X$ and every $w \in \mathbf{\Sigma}^*$, there is precisely one path with label w starting at x.

With our edge-labelled graph (X, E, ℓ), we can consider the *full shift space* which consists of all bi-infinite words over $\mathbf{\Sigma}$ that can be read along the edges of some bi-infinite path in (X, E, ℓ). When (X, E, ℓ) is strongly connected, the entropy $\mathsf{h}(L_{x,y})$ is independent of x and y and equals the topological entropy of the full shift space of the graph. See, for example, [10,14] or [3] for a selection of related work and references, and also the discussion in [11, Section 13.9].

If we consider the shift space consisting of all those bi-infinite words as above that do not contain any factor in F, then the interpretation of Corollary 1.2 is that the associated entropy is strictly smaller than $\mathsf{h}(X)$.

The theorem, once approached in the right way, is not hard to prove. It is based on a classical tool, a version of the Perron–Frobenius theorem for infinite non-negative matrices; see e.g. [16]. We shall first reformulate things in terms of Markov chains and forbidden transitions.

Please cite this article in press as: W. Huss, et al., Entropy sensitivity of languages defined by infinite automata, via Markov chains with forbidden transitions, Theoretical Computer Science (2010), doi:10.1016/j.tcs.2010.07.020

2. Markov chains and forbidden transitions

We now equip the oriented, edge-labelled graph (X, E, ℓ) with additional data: with each edge $e = (x, a, y)$, we associate a probability $p(e) = p(x, a, y) \geq \alpha > 0$, where α is a fixed constant, such that

$$\sum_{e \in E : e^- = x} p(e) \leq 1 \quad \text{for every } x \in X. \tag{1}$$

Our assumption to have the uniform lower bound $p(e) \geq \alpha$ for each edge implies that the outdegree (number of outgoing edges) of each vertex is bounded by $1/\alpha$. We interpret $p(e)$ as the probability that a particle with current position $x = e^-$ moves in one (discrete) time unit along e to its end vertex $y = e^+$. Observing the successive random positions of the particle at the time instants $0, 1, 2, \ldots$, we obtain a Markov chain with state space X whose one-step transition probabilities are

$$p(x, y) = \sum_{a \in \boldsymbol{\Sigma} : (x,a,y) \in E} p(x, a, y).$$

We shall also want to record the edges and their labels used in each step, which means considering a Markov chain on a somewhat larger state space, but we will not need to formalise this in detail. In (1), we admit the possibility that $1 - \sum_y p(x, y) > 0$ for some x. This number is then interpreted as the probability that a particle positioned at x dies at the next step.

We write $p^{(n)}(x, y)$ for the probability that the particle starting at x is at position y after n steps. This is the (x, y)-element of the n-power P^n of the transition matrix $P = \big(p(x, y)\big)_{x,y \in X}$. If (X, E, ℓ) is strongly connected, then P is irreducible, and it is well-known that the number

$$\rho(P) = \limsup_{n \to \infty} p^{(n)}(x, y)^{1/n}$$

is independent of x and y. See once more [16]. Often, $\rho(P)$ is called the spectral radius of P. It is the parameter of exponential decay of the transition probabilities.

Let once more $F \subset \boldsymbol{\Sigma}^+$ be finite. We interpret the elements of F as sequences of *forbidden transitions*. That is, we restrict the motion of the particle: at no time is it allowed to traverse any path π with $\ell(\pi) \in F$ in k successive steps, where k is the length of π. We write $p_F^{(n)}(x, y)$ for the probability that the particle starting at x is at position y after n steps, without having made any such sequence of forbidden transitions. Let

$$\rho_{x,y}(P_F) = \limsup_{n \to \infty} p_F^{(n)}(x, y)^{1/n}, \quad x, y \in X.$$

These numbers are not necessarily independent of x and y, and they are not the elements of the n-matrix power of some substochastic matrix.

Recall that a transition matrix $Q = \big(q(x, y)\big)_{x,y \in X}$ on the state space X is called *substochastic* if there exists a constant $\varepsilon > 0$ such that, for all $x \in X$,

$$\sum_{y \in X} q(x, y) \leq 1 - \varepsilon.$$

That is, all row sums are bounded by $1-\varepsilon$. In order to give an upper bound for the restricted transition probabilities $p_F^{(n)}(x, y)$, we first show the following.

Lemma 2.1. *Suppose that (X, E, l) is strongly connected with label alphabet $\boldsymbol{\Sigma}$ and equipped with transition probabilities $p(e) \geq \alpha > 0$, $e \in E$. Let $F \subset \boldsymbol{\Sigma}^+$ be a finite, non-empty set which is relatively dense in (X, E, ℓ). Then there are $k \in \mathbb{N}$ and $\varepsilon_0 > 0$ such that*

$$\sum_{y \in X} p_F^{(k)}(x, y) \leq 1 - \varepsilon_0 \quad \textit{for all } x \in X.$$

In other words, the transition matrix $Q = \big(p_F^{(k)}(x, y)\big)_{x,y \in X}$ is strictly substochastic, with all row sums bounded by $1 - \varepsilon_0$.

Proof. Let $R = \max_{w \in F} |w|$, and let $D \in \mathbb{N}$ be the constant from the definition of relative denseness of F. Set $k = D + R$. For each $x \in X$, we can find a path π_1 from x to some $y \in X$ with length $d \leq D$ and a path π_2 starting at y which has label $w \in \boldsymbol{\Sigma}^*$. Let z be the endpoint of π_2, and choose any path π_3 that starts at z and has length $k - d - |w|$. (Such a path exists by strong connectedness.) Then let π be the path obtained by concatenating π_1, π_2 and π_3.

The probability that the Markov chain starting at x makes its first k steps along the edges of π is

$$\mathbb{P}(\pi) \geq \alpha^k = \varepsilon_0 > 0.$$

Hence

$$\sum_{y \in X} p_F^{(k)}(x, y) \leq \sum_{y \in X} p^{(k)}(x, y) - \mathbb{P}(\pi) \leq 1 - \varepsilon_0,$$

and this upper bound holds for every x. □

Please cite this article in press as: W. Huss, et al., Entropy sensitivity of languages defined by infinite automata, via Markov chains with forbidden transitions, Theoretical Computer Science (2010), doi:10.1016/j.tcs.2010.07.020

The matrix P acts on functions $h : X \to \mathbb{R}$ by $Ph(x) = \sum_y p(x,y)h(y)$. Next, we state two key results due to Pruitt [15, Lemma 1] and [15, Corollary to Theorem 2], which will be used in the proof of the main result.

Lemma 2.2. *If the transition matrix P is irreducible and $Ph \le sh$ for some $s > 0$ and $h \ne 0$, then $h > 0$.*

Lemma 2.3. *If the transition matrix $P = \{p(x,y)\}_{x,y\in X}$ is such that for every $x \in X$ the entries $p(x,y) = 0$ for all $y \in X$ except finitely many, then the equation*

$$Ph = sh$$

has a solution for all $s \ge \rho(P)$.

Using these lemmatas, we prove the following result on sensitivity of the Markov chain with respect to forbidding the transitions in F.

Theorem 2.4. *Suppose that (X, E, ℓ) is uniformly connected with label alphabet $\mathbf{\Sigma}$ and equipped with transition probabilities $p(e) \ge \alpha > 0$, $e \in E$. Let $F \subset \mathbf{\Sigma}^+$ be a finite, non-empty set which is relatively dense in (X, E, ℓ). Then*

$$\sup_{x,y\in X} \rho_{x,y}(P_F) < \rho(P) \quad \textit{strictly.}$$

Proof. We shall proceed in two steps.

Step 1. *We assume that $P = \big(p(x,y)\big)_{x,y\in X}$ is stochastic and that $\rho(P) = 1$.*

Consider the matrix Q of Lemma 2.1. Let $Q^n = \big(q^{(n)}(x,y)\big)_{x,y\in X}$ be its n-th matrix power. $q^{(n)}(x,y)$ is the probability that the Markov chain starting at x is in y at time nk and does not make any forbidden sequence of transitions in each of the discrete time intervals $[(j-1)k, jk]$ for $j \in \{1,\dots,n\}$. Therefore

$$p_F^{(nk)}(x,y) \le q^{(n)}(x,y),$$

and also, by the same reasoning, for $i = 0,\dots,k-1$,

$$p_F^{(nk+i)}(x,y) \le \sum_{z\in X} q^{(n)}(x,z) p_F^{(i)}(z,y), \quad i = 0\dots,k-1.$$

Therefore, for every $x \in X$ and $i = 0,\dots,k-1$,

$$\sum_{y\in X} p_F^{(nk+i)}(x,y) \le \sum_{z\in X} q^{(n)}(x,z) \underbrace{\sum_{y\in X} p_F^{(i)}(z,y)}_{\le 1} \le (1-\varepsilon_0)^n,$$

since Lemma 2.1 implies that the row sums of the matrix power Q^n are bounded above by $(1-\varepsilon_0)^n$. We conclude that

$$\limsup_{n\to\infty} p_F^{(nk+i)}(x,y)^{1/(nk+i)} \le (1-\varepsilon_0)^{1/k},$$

so $\rho_{x,y}(P_F) \le (1-\varepsilon_0)^{1/k} = 1-\varepsilon$, where $\varepsilon > 0$.

Step 2. *General case.* We reduce this case to the previous one.

Since P is irreducible and every row of P has only finitely many non-zero entries, Lemmas 2.2 and 2.3 guarantee the existence of a strictly positive solution $h : X \to \mathbb{R}$ for the equation

$$Ph = \rho(P)\cdot h;$$

that is, h is $\rho(P)$-*harmonic.* Consider now the h-transform of the transition probabilities $p(e)$ of P, $e = (x,a,y) \in E$, given by

$$p^h(e) = p^h(x,a,y) = \frac{p(x,a,y)h(y)}{\rho(P)h(x)},$$

and the associated transition matrix P^h with entries

$$p^h(x,y) = \sum_{a:(x,a,y)\in E} p^h(x,a,y).$$

The Markov chain associated with P^h is called the h-*process.*

Then $\rho(P^h) = 1$. Using uniform connectedness, we show that there is a constant $\bar\alpha > 0$ such that $p^h(e) \ge \bar\alpha$ for each $e = (x,a,y) \in E$. Indeed, for such an edge, there is $k \le K$ such that $d^+(y,x) = k$, whence

$$\rho(P)^k h(y) = \sum_{z\in X} p^{(k)}(y,z)h(z) \ge \alpha^k h(x),$$

Please cite this article in press as: W. Huss, et al., Entropy sensitivity of languages defined by infinite automata, via Markov chains with forbidden transitions, Theoretical Computer Science (2010), doi:10.1016/j.tcs.2010.07.020

so

$$p^h(x,a,y) \geq \big(\alpha/\rho(P)\big)^{k+1}.$$

Recall that K is the constant used in the definition of the uniform connectedness. We can now choose $\bar{\alpha} = \big(\alpha/\rho(P)\big)^{K+1}$. We see that with P^h we are now in the situation of Step 1. Thus, forbidding the transitions of F for the Markov chain with transition matrix P^h, we get $\rho_{x,y}(P_F^h) \leq 1 - \varepsilon$ for all $x, y \in X$, where $\varepsilon > 0$.

We now show that $\rho_{x,y}(P_F^h) = \rho_{x,y}(P_F)/\rho(P)$, which will conclude the proof.

For a path $\pi = e_1 \dots e_n$ from x to y, let (as above) $\mathbb{P}(\pi)$ be the probability that the original Markov chain traverses the edges of π in n successive steps, and let $\mathbb{P}^h(\pi)$ be the analogous probability with respect to the h-process. Then

$$\mathbb{P}^h(\pi) = \frac{\mathbb{P}(\pi)h(y)}{\rho(P)^n h(x)}.$$

Let us write $\Pi_{x,y}^n(\neg F)$ for the set of all paths π from x to y with length n for which $\ell(\pi)$ does not contain a factor in F. Then the n-step transition probabilities of the h-process with the transitions in F forbidden are

$$p^{h(n)}_{F}(x,y) = \sum_{\pi \in \Pi_{x,y}^n(\neg F)} \mathbb{P}^h(\pi) = \sum_{\pi \in \Pi_{x,y}^n(\neg F)} \frac{\mathbb{P}(\pi)h(y)}{\rho(P)^n h(x)} = \frac{p_F^{(n)}(x,y)h(y)}{\rho(P)^n h(x)}.$$

Taking n-th roots and passing to the upper limit, we obtain the required identity. □

With this result, it is now easy to deduce Theorem 1.1.

Proof of Theorem 1.1. Since (X, E, l) is deterministic with label alphabet $\mathbf{\Sigma}$, the outdegree of every $x \in X$ is at most $|\mathbf{\Sigma}|$. Equip the edges of (X, E, ℓ) with the transition probabilities $p(x, a, y) = 1/|\mathbf{\Sigma}|$ when $(x, a, y) \in E$. Then the n-step transition probabilities of the resulting Markov chain are given by

$$p^{(n)}(x,y) = \frac{\big|\{w \in L_{x,y} : |w| = n\}\big|}{|\mathbf{\Sigma}|^n}.$$

Therefore, because (X, E, ℓ) is uniformly connected, we have

$$\mathsf{h}(X) = \mathsf{h}(L_{x,y}) = \limsup_{n\to\infty} \frac{1}{n} \log\big(p^n(x,y)|\mathbf{\Sigma}|^n\big) = \log\big(\rho(P)\cdot|\mathbf{\Sigma}|\big).$$

Analogously,

$$\mathsf{h}(L_{x,y}^F) = \log\big(\rho_{x,y}(P_F)\cdot|\mathbf{\Sigma}|\big).$$

By Theorem 2.4,

$$\sup_{x,y\in X} \rho_{x,y}(P_F) < \rho(P),$$

and this implies that

$$\sup_{x,y\in X} \mathsf{h}(L_{x,y}^F) < \mathsf{h}(X)$$

strictly. □

Application to pairs of groups and their Schreier graphs

Let G be a finitely generated group and K a (not necessary finitely generated) subgroup. Also, let $\mathbf{\Sigma}$ be a finite alphabet and let $\psi : \mathbf{\Sigma} \to G$ be such that the set $\psi(\mathbf{\Sigma})$ generates G as a semigroup. We extend ψ to a monoid homomorphism from $\mathbf{\Sigma}^*$ to G by $\psi(w) = \psi(a_1)\cdots\psi(a_n)$ if $w = a_1 \dots a_n$ with $a_i \in \mathbf{\Sigma}$ (and $\psi(\epsilon) = 1_G$). The mapping ψ is called a *semigroup presentation* of G in [8].

The *Schreier graph* $X = X(G, K, \psi)$ has vertex set

$$X = \{Kg : g \in G\},$$

the set of all right K-cosets in G, and the set of all labelled, directed edges E is given by

$$E = \{e = (x,a,y) : x = Kg, y = Kg\psi(a),\ \text{where } g \in G,\ a \in \mathbf{\Sigma}\}.$$

Note that the graph X is fully deterministic and uniformly connected.

The *word problem* of (G, K) with respect to ψ is the language

$$L(G,K,\psi) = \{w \in \mathbf{\Sigma}^* : \psi(w) \in K\}.$$

Please cite this article in press as: W. Huss, et al., Entropy sensitivity of languages defined by infinite automata, via Markov chains with forbidden transitions, Theoretical Computer Science (2010), doi:10.1016/j.tcs.2010.07.020

The word problem for a recursively presented group G is the algorithmic problem of deciding whether two words represent the same element. Also, this terminology is used in the context of formal language theory and goes back at least to the seminal paper of Muller and Schupp [12]. For additional information, see also [13]. In their work, for a finitely generated group G the **word problem** $W(G)$ is the set of all words on the generators and their inverses which represent the identity element of G.

If we consider the "root" vertex $o = K$ of the Schreier graph, then in the notation of the introduction, we have $L(G, K, \psi) = L_{o,o}$; compare with [8, Lemma 2.4].

We can therefore apply Theorem 1.1 and Corollary 1.2 to the graph $X(G, K, \psi)$ in order to deduce the following.

Corollary 2.5. *The word problem of the pair* (G, K) *with respect to any semigroup presentation* ψ *is growth sensitive (with respect to forbidding an arbitrary non-empty finite subset* $F \subset \mathbf{\Sigma}^*$*).*

Acknowledgement

W. Huss and W. Woess were supported by the Austrian Science Fund project FWF-P19115-N18. E. Sava was supported by the NAWI Graz project.

References

[1] G.N. Arzhantseva, I.G. Lysenok, Growth tightness for word hyperbolic groups, Math. Z. 241 (2002) 597–611.
[2] J. Berstel, D. Perrin, C. Reutenauer, Codes and Automata, Cambridge University Press, Cambridge, 2010.
[3] M. Boyle, Mike, J. Buzzi, R. Gómez, Almost isomorphism for countable state Markov shifts, J. Reine Angew. Math. 592 (2006) 23–47.
[4] T. Ceccherini-Silberstein, F. Scarabotti, Random walks, entropy and hopfianity of free groups, in: V.A. Kaimanovich (Ed.), Random Walks and Geometry, de Gruyter, Berlin, 2004, pp. 413–419.
[5] T. Ceccherini-Silberstein, Growth and ergodicity of context-free languages. II. The linear case, Trans. Amer. Math. Soc. 359 (2007) 605–618.
[6] T. Ceccherini-Silberstein, W. Woess, Growth and ergodicity of context-free languages, Trans. Amer. Math. Soc. 354 (2002) 4597–4625.
[7] T. Ceccherini-Silberstein, W. Woess, Growth-sensitivity of context-free languages, Theoret. Comput. Sci. 307 (2003) 103–116.
[8] T. Ceccherini-Silberstein, W. Woess, Context-free pairs of groups I - Context-free pairs and graphs (submitted), arXiv:0911.0090v1.
[9] R.I. Grigorchuk, P. de la Harpe, On problems related to growth, entropy, and spectrum in group theory, J. Dynam. Control Systems 3 (1997) 51–89.
[10] B.M. Gurevič, Topological entropy of a countable Markov chain, Dokl. Akad. Nauk SSSR 187 (1969) 715–718 (Russian). Engl. transl. in Soviet Math. Dokl. 10 (1969) 911–915.
[11] D. Lind, B. Marcus, An Introduction to Symbolic Dynamics and Coding, Cambridge University Press, Cambridge, 1995.
[12] D.E. Muller, P.E. Schupp, Groups, the theory of ends, and context-free languages, J. Comput. System Sci. 26 (1983) 295–310.
[13] D. E. Muller, P. E. Schupp, The theory of ends, pushdown automata, and second-order logic, Theoret. Comput. Sci. 37 (1985) 51–57.
[14] K. Petersen, Chains, entropy, coding, Ergodic Theory Dynam. Systems 6 (1986) 415–448.
[15] W. E. Pruitt, Eigenvalues of non-negative matrices, Ann. Math. Statist. 35 (4) (1964) 1797–1800.
[16] E. Seneta, Non-negative Matrices and Markov Chains, Revised reprint of the second edition, Springer, New York, 2006.

Please cite this article in press as: W. Huss, et al., Entropy sensitivity of languages defined by infinite automata, via Markov chains with forbidden transitions, Theoretical Computer Science (2010), doi:10.1016/j.tcs.2010.07.020