Unambiguous automata

Marie-Pierre Béal, Eugen Czeizler, Jarkko Kari and Dominique Perrin

Abstract. We give a new presentation of two results concerning synchronized automata. The first one gives a linear bound on the synchronization delay of complete local automata. The second one gives a cubic bound for the minimal length of a synchronizing pair in a complete synchronized unambiguous automaton. The proofs are based on results on unambiguous monoids of relations.

Mathematics Subject Classification (2000). 68Q45;20M35;94A45;68R15.

Keywords. automata, local automata, synchronized automata, unambiguous automata, unambiguous monoids of relations.

1. Introduction

A finite word w is said to be synchronizing if all paths labeled by w lead to a unique state. A finite deterministic automaton that has a synchronizing word is called synchronized. It is known that an n-state synchronized complete automaton has a synchronizing word of length $O(n^3)$. It is Cerny's Conjecture that there is always one of length at most $(n-1)^2$ [15].

The definition of a synchronizing word can be extended to non-deterministic automata as follows. A pair of words (x, y) is called synchronizing if all paths (p, x, q, y, r), where p, q, r are states, use the same state q. The automaton is again called synchronizing if there is a synchronizing pair. An unambiguous automaton is such that there is at most one path with a given origin, end, and label. Carpi proved that an n-state unambiguous, transitive and synchronized automaton has a synchronizing pair of length $O(n^3)$ [5].

Local automata form a particular class of unambiguous automata. They appear in several contexts since the beginnings of automata theory. Informally, a non-deterministic automaton is local if the knowledge of a fixed amount of symbols

Research supported by the Academy of Finland grant 211967 and by the Turku University Foundation grant.

of the input determines the current state. Thus strongly connected local automata are unambiguous. Local deterministic automata are the definite automata of [14]. They are also called finite-memory automata in [11].

Local automata have very strong synchronization properties that make them very helpful in the construction of encoders or decoders used in constrained coding. Their synchronization properties guarantee that the decoder does not propagate errors. An automaton is (m, a)-local (m stands for memory and a for anticipation) if whenever two paths of length $m + a$ have the same label, they go through the same state at time m. The minimal length $m+a$ such that the automaton fits this condition is called the synchronization delay of the local automaton. For coding purposes, the synchronization delay is made as small as possible.

It is known since Perles, Rabin and Shamir [14] that an n-state deterministic local and complete automaton has a synchronization delay at most $n-1$. Recently, two of the authors proved [7, 8] that this result still holds in the non-deterministic case for complete local automata. In the general case, when the local automaton is not assumed to be complete, the upper bound of the synchronization delay of a local automaton is quadratic (see for instance [1]).

In this paper, we present some results of Césari on unambiguous monoids of relations. We consider automata for which all states have to be considered as both initial and final. We restrict to the case of irreducible complete synchronized automata, which includes irreducible complete local automata. The transition monoid M of such an automaton is an unambiguous monoid of relations, which is complete and synchronizing. Césari proved in[6] that the set of maximal rows of elements of M is invariant under the right action of M. He also proved that each idempotent of M is a sum of elements of the form $c \cdot r$, where r is a maximal row vector and c is a maximal column vector. This property holds even if the automaton is not local. We then show how to derive the bound of the synchronization delay of local complete automata from the stability property of the set of maximal rows and the set of maximal columns.

The paper is organized as follows. In Section 2, we recall basic definitions about automata. In Sections 3 and 4 we recall results from Césari and Boë (see [6] and [3]) about unambiguous monoids of relations that are used in the last sections.

In Section 5, we reproduce the result of Carpi [5] on the bound of the length of a synchronizing word in a complete and unambiguous automaton. In Section 6, we give our proof of the upper bound of the synchronization delay of irreducible local complete automata.

2. Unambiguous and local automata

We begin with some definitions of automata. A (finite) *automaton* is a pair $\mathcal{A} = (Q, E)$, where Q is a finite set of states, and E is a finite set of edges labeled in a finite alphabet A. Note that no initial or final states are specified. Actually, all states have to be considered as both initial and final. We say that an automaton is *irreducible* if it has a strongly connected graph.

An automaton is *deterministic* if two edges with the same origin carry different labels. An automaton is *unambiguous*[1] if two paths with the same origin, end and label are equal.

A word w is *synchronizing* for an automaton $\mathcal{A} = (Q, E)$ if for any $p, q, r, s \in Q$, $p \xrightarrow{w} q$ and $r \xrightarrow{w} s$ imply $p \xrightarrow{w} s$ and $r \xrightarrow{w} q$. An automaton is *synchronized* if it has a synchronizing word.

Let m and a be two nonnegative integers (m stands for memory, and a for anticipation). We say that an automaton is (m, a)-*local* if whenever two paths $((p_i, a_i, p_{i+1}))_{0 \le i < (m+a)}$ and $((p'_i, a_i, p'_{i+1}))_{0 \le i < (m+a)}$ of length $m+a$ have the same label, then $p_m = p'_m$. Note that a deterministic and local automaton is $(m, 0)$-local for some integer m. We say that an automaton is local if it is (m, a)-local for some integers m, a. The minimal value $m + a$ for all (m, a) such that an automaton is (m, a)-local is called its *synchronization delay*. Note that any word of length $m+a$ in an (m, a)-local automaton is synchronizing.

An irreducible automaton is local if and only if any two distinct cycles carry different labels. It is known that an irreducible local automaton is unambiguous.

Let A be a finite alphabet. An automaton is *complete* for A if any word of A^* is label of some path of the automaton.

The upper bound of the synchronization delay of a local automaton is quadratic in general (see for instance [1]). It is known from Perles, Rabin and Shamir [14] that an n-state deterministic local and complete automaton has a synchronization delay of at most $n - 1$. It is proved in [7, 8] that this result holds even if the automaton is not deterministic.

An example of a $(2, 1)$-local automaton is displayed in Figure 1.

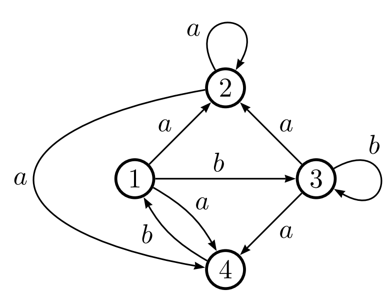

FIGURE 1. The automaton $\mathcal{A}$ is $(2, 1)$-local and complete for $A = \{a, b\}$.

We associate to any unambiguous automaton $\mathcal{A} = (Q, E)$, a deterministic automaton $\det(\mathcal{A}) = (\mathcal{P}(Q), F)$, where (P, a, R) is an edge of $\det(\mathcal{A})$ if and only if

[1] or has no diamond.

$R = \{q \in Q \mid \exists\, p \in P \text{ with } (p, a, q) \in E\}$. We keep only the states accessible from the states $\{q\}$, where $q \in Q$. Similarly, we associate a codeterministic automaton $\mathrm{codet}(\mathcal{A})$ to $\mathcal{A}$. It is defined as the reverse automaton of the deterministic automaton associated to the reverse automaton of $\mathcal{A}$.

The deterministic and codeterministic automaton associated to the automaton of Figure 1 are shown in Figures 2 and 3 respectively.

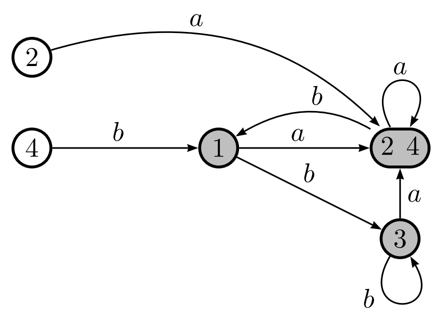

FIGURE 2. The deterministic automaton $\det(\mathcal{A})$ associated to $\mathcal{A}$. The states in gray are maximal for the inclusion. They form a final strongly connected component of the graph of $\det(\mathcal{A})$.

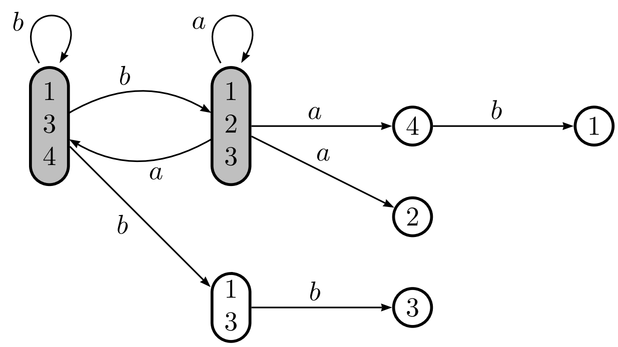

FIGURE 3. The codeterministic automaton $\mathrm{codet}(\mathcal{A})$ associated to $\mathcal{A}$. The states in gray are maximal for the inclusion. They form a final strongly connected component of the graph of $\mathrm{codet}(\mathcal{A})$.

3. Unambiguous monoids of relations

Unambiguous monoids of relations have been studied by Césari [6], Boë [3], and Carpi [4], [5], see also Berstel, Perrin and Reutenauer [2]. The transition monoids of unambiguous automata are unambiguous monoids of relations.

Let Q be a finite set. A *relation* over Q is a subset of $Q \times Q$. The product mn of two relations m and n on Q is the relation defined by $(p,q) \in mn$ if and only if there exists $r \in Q$ such that $(p,r) \in m$ and $(r,q) \in n$. The product is said to be *unambiguous* if for every $p,q \in Q$, there exists at most one r such that $(p,r) \in m$ and $(r,q) \in n$. A relation can be considered as a $Q \times Q$ matrix with $0-1$ coefficients. The product of relations corresponds to the product of matrices with Boolean coefficients. If it is unambiguous, it corresponds also the to the product of matrices over $\mathbb{N}$.

An *unambiguous monoid of relations* M on a set Q is a submonoid of the monoid of all relations on Q such that for every $m,n \in M$ the product mn is unambiguous.

An unambiguous monoid of relations over Q is *complete* if it does not contain the null matrix.

An unambiguous monoid of relations over Q is *transitive* if for any $p,q \in Q$, there is an element $m \in M$ such that $(p,q) \in m$.

An unambiguous monoid of relations over Q is *synchronizing* if it contains a matrix of rank 1.

A complete and transitive unambiguous monoid of relations over a finite set Q is obtained from a transitive and complete unambiguous automaton. Let $\mathcal{A} = (Q,E)$ be a transitive and complete unambiguous automaton over the alphabet A. To each word $w \in A^*$ we associate a relation $\mu(w)$ on Q defined by $(p,q) \in \mu(w)$ if there is a path from p to q labelled by w. Then the monoid $\mu(A^*)$ is an unambiguous monoid of relations if and only if the automaton $\mathcal{A}$ is unambiguous. It is transitive when $\mathcal{A}$ is irreducible. It is complete when $\mathcal{A}$ is complete. The monoid M is called the *transition monoid* of the automaton $\mathcal{A}$.

A Q-vector is a vector indexed by Q. If v is a Q-zero-one vector, then it is the characteristic vector of a set. We say that v contains $q \in Q$ if $v_q = 1$. If $q \in Q$ and when there is no confusion, we sometimes denote by q the row characteristic vector of q, and by $\bar{q}$ the column characteristic vector of q. If v,w are two Q-zero-one vectors, we say that v contains w if $v_q = 1$ whenever $w_q = 1$.

Proposition 1 (Césari). *Let Q be a finite set and M a complete and transitive unambiguous monoid of relations over Q. Let r be a non null row vector. The following assertions are equivalent.*

(i) *r is a maximal row*[2].
(ii) $0 \notin rM$.
(iii) *r is a row of a matrix $m \in M$ such that m has a minimal number of non null rows.*

Proof. (i) $\Rightarrow$ (iii).
Let r be a maximal row. It is non null. Let us assume that r is the row of index q of a matrix $m \in M$. Let m' be a matrix of M that has a minimal number of non null rows. Let us assume that $m'_{ps} = 1$ for some $p,s \in Q$. Since M is transitive,

[2]the maximality is for the inclusion and among all rows of all matrices of M.

there is a matrix n such that $n_{sq} = 1$. Hence $(m'n)_{pq} = 1$. Since m' has a minimal number of non null rows, $m'n$ and $m'nm$ also. From $(m'n)_{pq} = 1$ and since r is the row of index q of a matrix m, we get that the row of index p of $(m'nm)$ is greater than or equal to r. Since r is maximal, it is equal to r.

(iii) $\Rightarrow$ (ii).
Let r be a non null row of index q of $m \in M$ such that m has a minimal number of non null rows. Let us denote the set of indices of the null rows of m by S. Hence $q \notin S$. If there is $n \in M$ such that $rn = 0$, then the rows of indices in $S \cup \{q\}$ of mn are null. Hence mn has more null rows than m, which is a contradiction.

(ii) $\Rightarrow$ (i).
Let us assume that r is not a maximal row. There is a maximal row r' of a matrix in M with $r < r'$. Let $q \in Q$ such that $r_q = 0$ and $r'_q = 1$. Let m be a matrix of M that has a maximal row ℓ of index p. Since M is transitive, there is a matrix n with $n_{qp} = 1$. Hence the row of index q of nm is greater than or equal to ℓ. By maximality of ℓ, it is equal to ℓ. Then from $r'_q = 1$ we get that $r'nm$ is a row of an element of M greater than or equal to ℓ. Again by maximality of ℓ, it is equal to ℓ. We get $\ell = r'nm \geq r'nm - rnm = (r' - r)nm \geq \ell$. As a consequence, $rnm = 0$. □

The previous proposition yields the following result from Césari, which says that the set of maximal rows is invariant under the right action of the monoid.

We denote by $\mathcal{R}$ (resp. $\mathcal{C}$) the set of maximal rows (resp. maximal columns) of an unambiguous monoid of relations.

Proposition 2 (Césari, Boë). *Let Q be a finite set and M a complete and transitive unambiguous monoid of relations over Q. For any maximal row r and any element m of M, rm is a maximal row. Let $\mathcal{R}$ be the set of maximal rows of M. The monoid M acts transitively on $\mathcal{R}$ on the right.*

Proof. Let r be a maximal row and $m \in M$. We show that rm is a maximal row. By Proposition 1, if rm is not maximal, then there is $n \in M$ such that $rmn = 0$. In this case, r is not maximal either, again by Proposition 1.

Let $r, r' \in \mathcal{R}$. Let us assume that r is the row of index p (resp. p') of a matrix $m \in M$ (resp. m'). Hence $pm = r$ and $p'm' = r'$. Since M is transitive, there is a matrix n such that $(rn)_{p'} = 1$. Hence $rnm' = r'$ since r' is maximal. It follows that M acts transitively on the set $\mathcal{R}$ on the right ($\mathcal{R}M = \mathcal{R}$). □

Propositions similar to Propositions 1 and 2 hold for maximal columns. The monoid M acts transitively on $\mathcal{C}$ on the left.

We add a further property of maximal rows which has been remarked by Friedman [9].

Let $\mathcal{A} = (Q, E)$ be an irreducible and complete unambiguous automaton over the alphabet A. For $p, q \in Q$, let X_{pq} be the number of edges from p to q in $\mathcal{A}$. Let X be the $Q \times Q$-matrix with elements X_{pq}.

Let $k = \operatorname{card} A$. It is known that X has a simple eigenvalue equal to k and a positive corresponding column eigenvector w, unique up to a positive multiplicative scalar. Indeed, since $\mathcal{A}$ is complete and unambiguous, the matrix X has the eigenvalue $\log k$ (see for instance [12, p. 269]). Since $\mathcal{A}$ is irreducible, X is irreducible and the eigenvalue $\log k$ is simple.

For each state $q \in Q$, we call w_q the *weight* of q. The weight of a subset of Q is the sum of the weights of its elements.

Proposition 3. *Let M be a complete and transitive unambiguous monoid of relations over Q and let $\mathcal{R}$ be the set of maximal rows of M. All elements of $\mathcal{R}$ have the same weight.*

Proof. For each row $r \in \mathcal{R}$, we denote by $w(r)$ the weight of r. We denote by $r \cdot a$ the maximal row $r\mu(a)$. Since $Xw = kw$, we get

$$(Xw)_q = kw_q,$$

$$\sum_{q \in r} (Xw)_q = kw(r),$$

$$\sum_{a \in A} w(r \cdot a) = kw(r).$$

Hence this implies that the weights of the maximal rows are equal. Indeed, if r has maximal weight, then each $r \cdot a$ has the same maximal weight. The result follows from the fact that the action of A is transitive on $\mathcal{R}$. □

Example 1. Consider again the automaton of Figure 1. We have $k = 2$ and

$$X = \begin{bmatrix} 0 & 1 & 1 & 1 \\ 0 & 1 & 0 & 1 \\ 0 & 1 & 1 & 1 \\ 1 & 0 & 0 & 0 \end{bmatrix}, \quad w = \begin{bmatrix} 2 \\ 1 \\ 2 \\ 1 \end{bmatrix}.$$

All maximal rows, corresponding to sets $\{1\}, \{2, 4\}, \{3\}$ (see Figure 2), have the same weight 2.

A similar result holds for maximal columns.

4. Boxes

The notion of box is related to the notion of maximal rows and maximal columns in an unambiguous monoid of relations. It was introduced and studied by Boë in [3].

Let Q be a finite set. We say that a set S of Q-zero-one vectors *covers* Q if the union of the coordinates 1 of all vectors of S is Q.

A pair $(\mathcal{R}, \mathcal{C})$ is called a *box*[3] of Q if and only if

[3] Actually the definition of a box given by Boë in [3] differs slightly from the definition given in this paper. In [3], Boë defines a strongbox as a pair $(\mathcal{R}, \mathcal{C})$ satisfying conditions (1),(2), and (3),

1. $\mathcal{R}$ is a set of Q-zero-one row vectors covering Q.
2. $\mathcal{C}$ is a set of Q-zero-one column vectors covering Q.
3. for any vectors $r \in \mathcal{R}$, $c \in \mathcal{C}$, $r \cdot c = 1$.

Note that when $(\mathcal{R}, \mathcal{C})$ is a box, any vector of $\mathcal{R}$ (resp. $\mathcal{C}$) is maximal with respect to the inclusion in $\mathcal{R}$ (resp. in $\mathcal{C}$). Indeed, let us assume that $v, w \in \mathcal{R}$ and v contains w. Let q such that $w_q = 0$ and $v_q = 1$. Let $c \in \mathcal{C}$ containing q. Since $v \cdot c = w \cdot c = 1$, we have $(v - w) \cdot c = 0$, which is a contradiction with c contains q.

The following proposition was obtained by Césari [6].

Proposition 4 (Césari). *Let M be a transitive and complete unambiguous monoid of relations over a finite set Q. Let $\mathcal{R}$ be its set of maximal rows, and $\mathcal{C}$ be its set of maximal columns. Any element m of M of minimal rank k can be written*

$$m = \sum_{i=1}^{k} c_i r_i,$$

with $r_i \in \mathcal{R}, c_i \in \mathcal{C}$.

Proof. In the case m is an idempotent of minimal rank, the decomposition of m in the form $\sum_{i=1}^{k} c_i r_i$ is quite simple. Let $r_1, \ldots, r_n$ (resp. $c_1, \ldots, c_n$) be the rows (resp. the columns) of m. By Proposition 1, $r_i \in \mathcal{R}$ and $c_j \in \mathcal{C}$. Let I be the intersection of the set of indices of the non null rows of m and of the set of indices of the non null columns of m. We have $m = m^2 = \sum_{i=1}^{|I|} c_i r_i$, $r_i \in \mathcal{R}, c_i \in \mathcal{C}$, and $|I| = k$.

We now consider the case where m is not supposed to be idempotent. Since M is finite, there is an idempotent e that has the same rank k as m, and such that $em = me = m$. Let C be the $|Q| \times k$ matrix whose column of index i is c_i, and let R be the $k \times |Q|$ matrix whose row of index j is r_j. We have

$$m = eme = CRmCR.$$

The matrix RmC is a nonnegative integral invertible matrix. Moreover, its coefficients are 0 or 1 since the rows r_i and the columns c_i are non-null. Hence RmC is a permutation matrix, and $m = \sum_{i=1}^{k} c_i r_{\pi(i)}$, where π is a permutation of $\{1, 2, \ldots, k\}$.

□

Proposition 5. *Let M be a transitive and complete unambiguous monoid of relations over a finite set Q. Let $\mathcal{R}$ be the set of maximal rows of M, and $\mathcal{C}$ be the set of maximal columns of M. Then $(\mathcal{R}, \mathcal{C})$ is a box if and only if M is synchronizing.*

Proof. Since M is transitive, for any element q of Q, there is a row r such that $r_q = 1$. Hence there is a maximal row ℓ such that $\ell_q = 1$. It follows that $\mathcal{R}$ covers Q. Similarly, $\mathcal{C}$ covers Q.

---

such that moreover $\mathcal{R}$ is the set of sections of $\mathcal{C}$, and $\mathcal{C}$ is the set of sections $\mathcal{R}$. A section of $\mathcal{C}$ is a part intersecting each set of $\mathcal{C}$ in exactly one element. Hence our definition of a box is weaker than Boë's definition of a strongbox.

Let $r \in \mathcal{R}$. Let us assume that r is the row of index p of a matrix $m \in M$. Hence $pm = r$. Let $c \in \mathcal{C}$. There is $q \in Q$ and a matrix $n \in M$ with $c = n\bar{q}$. This implies that $r \cdot c = pmn\bar{q} = (mn)_{pq}$. Hence $r \cdot c \leq 1$ since M is an unambiguous monoid of relations.

If M is synchronizing, let m be an element of rank 1. By Proposition 4, $m = z \cdot \ell$, where $z \in \mathcal{C}$ and $\ell \in \mathcal{R}$. Let $r \in \mathcal{R}$ and $c \in \mathcal{C}$, there are matrices $n, s \in M$ with $r = \ell n$ and $c = sz$. Hence $c \cdot r = (sz) \cdot (\ell n) = s(z \cdot \ell)n = smn$ belongs to M. As a consequence, since M is complete,

$$c \cdot (r \cdot c) \cdot r \neq 0.$$

This implies that $r \cdot c \neq 0$, *i.e.* $r \cdot c = 1$, and thus $(\mathcal{R}, \mathcal{C})$ is a box.

If M is not synchronizing, let m be an element of minimal rank k, with $k \geq 2$. By Proposition 4, $m = \sum_{i=1}^{k} c_i r_i$ with $c_i \in \mathcal{C}$, $r_i \in \mathcal{R}$. We have

$$m^2 = (\sum_{i=1}^{k} c_i r_i)(\sum_{i=1}^{k} c_i r_i) = \sum_{i=1}^{k}\sum_{j=1}^{k} c_i (r_i \cdot c_j) r_j.$$

Since m^2 has rank k, the matrix $(r_i \cdot c_j)_{1 \leq i,j \leq n}$ is a permutation matrix. As $k \geq 2$, there are indices i, j such that $r_i \cdot c_j = 0$. This implies that $(\mathcal{R}, \mathcal{C})$ is not a box. □

One side of Proposition 5 was proved by Boë in [3]. Both sides were obtained by Nasu [13] in the particular case of the transition monoid of an unambiguous and complete automaton which has a De Bruijn graph. In this case, the in-degree and the out-degree of each state are constant. For this kind of automata, it was proved by L. R. Welch (see [10, Theorem 14.4]) that all maximal rows contain the same number of states $n(\mathcal{R})$ of the automaton, and all maximal columns contain the same number $n(\mathcal{C})$ of states of the automaton. Moreover, a reformulation of a result from G. A. Hedlund in [10, Theorem 14.9, Theorem 15.1] states that, in a complete automaton on a d-letter alphabet, if k is the minimal rank of the elements of M, and N is the number of states of the De Bruin graph, then one has

$$n(\mathcal{R})n(\mathcal{C})k = N/d.$$

Finally Nasu proved in [13, Theorem 3] that $n(\mathcal{R})n(\mathcal{C}) = N/d$ if and only if $(\mathcal{R}, \mathcal{C})$ is a box, which is the statement of Proposition 5 for transition monoids of unambiguous automaton with a De Bruin graph.

The transition monoid of the automaton of Figure 1 has a box $(\mathcal{R}, \mathcal{C})$ shown in Figure 4.

Proposition 6 (Boë). *Let $(\mathcal{R}, \mathcal{C})$ be a box of a finite set Q. There is a transitive complete and synchronizing unambiguous monoid of relations M such that $\mathcal{R}$ is the set of maximal rows of M, and $\mathcal{C}$ is the set of maximal columns of M.*

Proof. Let $(\mathcal{R}, \mathcal{C})$ be a box of Q. We define M as

$$M = \mathcal{C} \cdot \mathcal{R} = \{c \cdot r \mid r \in \mathcal{R}, c \in \mathcal{C}\}.$$

Let $m_1 = c_1 r_1 \in M$, $m_2 = c_2 r_2 \in M$. We have $m_1 m_2 = c_1 r_1 c_2 r_2 = c_1 r_2$ since $(\mathcal{R}, \mathcal{C})$ is a box. Hence M is a monoid of unambiguous relations. It is complete

$\mathcal{R} \setminus \mathcal{C}$	123	134
1	1	1
24	2	4
3	3	3

Figure 4. The transition monoid M of the automaton of Figure 1 is a monoid of unambiguous relations. It is complete, transitive and synchronizing. Let $\mathcal{R}$ (resp. $\mathcal{C}$) be the set of maximal rows (resp. columns) of M. The set $\mathcal{R}$ is the set of characteristic vectors of the states of the strongly connected component of the deterministic automaton $\det(\mathcal{A})$ (resp. the codeterministic automaton $\mathrm{codet}(\mathcal{A})$) associated to $\mathcal{A}$. These states are represented in gray in Figures 2 and 3. The box $(\mathcal{R}, \mathcal{C})$ is displayed here. The unique state corresponding to a row r and a column c is the intersection of r and c.

by definition. It is synchronizing since every element of M has rank 1. Since $\mathcal{R}$ and $\mathcal{C}$ covers Q, for any elements $p, q \in Q$, there are $r \in \mathcal{R}$, $c \in \mathcal{C}$ such that $r_q = 1$ and $c_p = 1$. Hence $(cr)_{pq} = 1$ and M is transitive. Finally, the elements of $\mathcal{R}$ are maximal rows of M (and the elements of $\mathcal{C}$ are maximal columns of M). Indeed, let $r \in \mathcal{R}$. If there is an element $m_1 = c_1 r_1 \in M$ such that $rm_1 = 0$, then $0 = rm_1 = rc_1r_1 = m_1$, which is a contradiction. □

The transitive complete and synchronizing unambiguous monoid of relations constructed in Proposition 6 for the box $(\mathcal{R}, \mathcal{C})$ of Figure 5 is the transition monoid of the automaton of Figure 6.

$\mathcal{R} \setminus \mathcal{C}$	12	13	34
14	1	1	4
23	2	3	3

Figure 5. A box (R, C).

5. A bound on the length of a synchronizing word

The following result is due to Carpi [5]. We give here a slightly simplified proof.

Proposition 7 (Carpi). *If $\mathcal{A}$ is an n-state irreducible synchronized unambiguous automaton, it has a synchronizing word of length at most $(n^2 - n + 2)(n - 1)/2$.*

We prove several preliminary results.

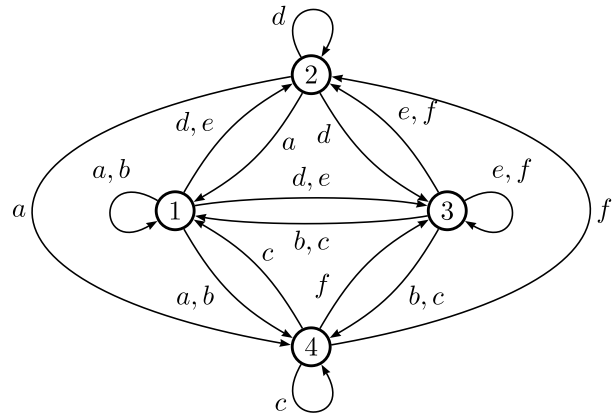

FIGURE 6. The automaton displayed here is irreducible, $(1,1)$-local and complete for $A = \{a,b,c,d,e,f\}$. Each letter corresponds to a relation defined by the product of a row in $\mathcal{R}$ and a column in $\mathcal{C}$ of Figure 5. For instance $a = (1100) \cdot (1001)^t$. The transitive complete and synchronizing monoid of unambiguous relations constructed in Proposition 6 is the transition monoid of this automaton.

For two relations m, m' on a set Q, we denote $m \le m'$ if for all $p, q \in Q$ one has $m_{pq} \le m'_{pq}$. For a relation m on a set Q and a state $p \in Q$, we denote by m_{p*} the row of index p of m, and by m_{*p} the column of index p of m.

Proposition 8. *Let M be a transitive complete unambiguous monoid of relations. Let $m, m' \in M$. If $m \le m'$ then $m = m'$.*

Proof. Suppose that $m'_{pq} = 1$ for some $p, q \in Q$. Since M is complete and transitive, there exists a row $\ell \in \mathcal{R}$ such that $\ell_p = 1$. Let us assume that $\ell = n_{r*}$ for some $n \in M$. Then $nm \le nm'$ and $(nm)_{r*} = n_{r*}m$ is a maximal row by Proposition 2. Thus $(nm)_{r*} = (nm')_{r*}$. This forces $m_{pq} = 1$ since $m \le m'$. □

Consider an unambiguous complete automaton $\mathcal{A} = (Q, E)$. Set $n = \mathrm{card}(Q)$. We denote by φ the morphism from A^* into the transition monoid of $\mathcal{A}$. Let $M = \varphi(A^*)$. We still denote by $\mathcal{R}$ (resp. $\mathcal{C}$) the set of rows (resp. columns) of the matrices in M which are maximal for inclusion.

Proposition 9. *Let $p \in Q$ and $u \in A^*$. If $\varphi(u)_{p*}$ is not in $\mathcal{R}$ there is a state q and a word v of length at most $n(n-1)/2$ such that $\varphi(u)_{p*} < \varphi(vu)_{q*}$.*

Proof. Since $\mathcal{A}$ is complete and transitive, there exists a row $r \in \mathcal{R}$ such that $r_p = 1$. There is at least a state p' distinct of p such that $r_{p'} = 1$ and $\varphi(u)_{p'*} \neq 0$ since otherwise $r\varphi(u) \notin \mathcal{R}$. Hence there is a state $q \in Q$ and a word v of length at most $n(n-1)/2$ such that $q \xrightarrow{v} p$ and $q \xrightarrow{v} p'$. Then $\varphi(u)_{p*} < \varphi(vu)_{q*}$. □

The symmetric result holds for columns.

Proposition 10. *There exists* $w \in A^*$ *with* $|w| \leq n(n-1)^2/2 + (n-1)$ *such that* $cr \leq \varphi(w)$ *for some* $r \in \mathcal{R}$ *and* $c \in \mathcal{C}$.

Proof. By Proposition 9 and its symmetric, there exists pairs (p_1, u_1), (p_2, u_2), $\ldots, (p_s, u_s)$ in $Q \times A^*$ and $(v_1, q_1), (v_2, q_2), \ldots, (v_t, q_t)$ in $A^* \times Q$, with $x_i = \varphi(u_i \cdots u_1)_{p_i *}$ and $y_j = \varphi(v_1 \cdots v_j)_{* q_j}$, such that:

(i) $u_1 = v_1 = 1$ and $p_1 = q_1$.
(ii) for $2 \leq i \leq s$, the word u_i has length at most $n(n-1)/2$ and $x_i > x_{i-1}$.
(iii) for $2 \leq j \leq t$, the word v_j has length at most $n(n-1)/2$ and $y_j > y_{j-1}$.
(iv) $x_s \in \mathcal{R}$ and $y_t \in \mathcal{C}$.

Let $u = u_s \ldots u_1$ and $v = v_1 \ldots v_t$. We have $|u| \leq (s-1)n(n-1)/2$ and $|v| \leq (t-1)n(n-1)/2$. Thus $|uv| \leq (s+t-2)n(n-1)/2$. Since $\mathcal{A}$ is unambiguous, we have that $x_s y_t = 1$. Thus $s + t \leq \sum_{q \in Q} (x_s)_q + \sum_{q \in Q} (y_t)_q \leq n+1$. Let finally $z \in A^*$ be such that $q_t \xrightarrow{z} p_s$ with $|z| \leq n-1$. Then $w = vzu$ satisfies the conditions of the statement. □

Proof of Proposition 7. By Proposition 10 there is a word w of length at most $n(n-1)^2/2 + (n-1) = (n^2 - n + 2)(n-1)/2$ such that $cr \leq \varphi(w)$ for some $r \in \mathcal{R}$ and $c \in \mathcal{C}$. Let $m \in M$ be a relation of rank 1 such that c is a column of m. Then $m = cr'$ for some $r' \in \mathcal{R}$. Let $m' \in M$ be such that $m' = c'r$ for some $c' \in \mathcal{C}$. Then $mm' = cr$ and thus cr is in M. By Proposition 8, we have $\varphi(w) = cr$ and thus $\varphi(w)$ has rank 1. ■

6. Synchronization delay of a local automaton

We now come to the result of Czeizler and Kari [7] giving a linear upper bound of the synchronization delay of an irreducible local automaton. We give a proof using unambiguous monoids of relations.

Proposition 11 (Czeizler and Kari). *An* n*-state irreducible local complete automaton has a synchronization delay at most* $n-1$.

Proof. Let $\mathcal{A} = (Q, E)$ be an irreducible local and complete automaton. Let M be the transition monoid of $\mathcal{A}$. Let $\mathcal{R}$ (resp. $\mathcal{C}$) be the set of maximal rows (resp. columns) of M. It is the set of characteristic vectors of the states of the strongly connected component of the deterministic automaton $\det(\mathcal{A})$ (resp. the codeterministic automaton $\mathrm{codet}(\mathcal{A})$) associated to $\mathcal{A}$.

We define the following vectorial spaces (for instance over $\mathbb{R}$) D_i for $i \geq 0$.

$$D_0 = \langle \{r - r' \mid r, r' \in \mathcal{R}\} \rangle ,$$
$$D_i = \left\langle \{ru - r'u \mid r, r' \in \mathcal{R}, u \in A^i\} \right\rangle ,$$

where $\langle S \rangle$ denotes the vectorial space over $\mathbb{R}$ generated by the vectors in a set S.

Since $\mathcal{R}M \subseteq \mathcal{R}$ by Proposition 2, $D_1 \subseteq D_0$. It follows that, for any integer $i \geq 0$, $D_{i+1} \subseteq D_i$. Let us assume that $\mathcal{A}$ is (m, a)-local. Then $\det(\mathcal{A})$ is deterministic and $(m, 0)$-local. Hence all Ru with $u \in A^m$ are equal to the same characteristic vector of some state $q \in Q$. As a consequence, for any $u \in A^m$, $ru - r'u = 0$ for any $r, r' \in R$. We get

$$D_0 \supseteq D_1 \supseteq \cdots \supseteq D_i \supseteq D_{i+1} \supseteq \cdots \supseteq D_m = \{0\}.$$

If $D_i = D_{i+1}$ for an integer i, then $D_i = D_{i+1} = \cdots = D_m = \{0\}$. This implies that $m \leq \dim D_0 = \dim \langle \mathcal{R} - \mathcal{R} \rangle$, where $\mathcal{R} - \mathcal{R}$ is the set of vectors $r - r'$ for $r, r' \in \mathcal{R}$. Similarly, $a \leq \dim \langle \mathcal{C} - \mathcal{C} \rangle$, where $\mathcal{C} - \mathcal{C}$ is the set of vectors $c - c'$ for $c, c' \in \mathcal{C}$. We have $\langle \mathcal{R} - \mathcal{R} \rangle \subsetneq \langle \mathcal{R} \rangle$. The inclusion is strict since for any $r, r' \in R$ and some $c \in \mathcal{C}$, $(r - r') \cdot c = 0$. Similarly $\langle \mathcal{C} - \mathcal{C} \rangle \subsetneq \langle \mathcal{C} \rangle$.

From $(r - r') \cdot c = 0$ for any $r, r' \in \mathcal{R}$, $c \in \mathcal{C}$, we get that $\dim \langle \mathcal{R} - \mathcal{R} \rangle + \dim \langle \mathcal{C} \rangle \leq \operatorname{card} Q$. Finally, we obtain

$$\begin{aligned} m + a &\leq \dim \langle \mathcal{R} - \mathcal{R} \rangle + \dim \langle \mathcal{C} - \mathcal{C} \rangle \\ &\leq \dim \langle \mathcal{R} - \mathcal{R} \rangle + \dim \langle \mathcal{C} \rangle - 1 \\ &\leq \operatorname{card} Q - 1, \end{aligned}$$

which ends the proof. □

Example 2. Let us consider the local automaton $\mathcal{A}$ of Figure 1. The maximal rows of the transition monoid of $\mathcal{A}$ are the row characteristic vectors of the maximal states $\{1\}, \{3\}, \{2, 4\}$ of the deterministic automaton $\det(\mathcal{A})$ associated to $\mathcal{A}$. Hence the vectorial space $\langle \mathcal{R} \rangle$ has dimension 3. As a consequence the automaton is (m, a)-local with $m \leq 2$.

The maximal columns of the transition monoid of $\mathcal{A}$ are the column characteristic vectors of the maximal states $\{1, 2, 3\}, \{1, 3, 4\}$ of the codeterministic automaton $\operatorname{codet}(\mathcal{A})$ associated to $\mathcal{A}$. Hence the vectorial space $\langle \mathcal{C} \rangle$ has dimension 2. As a consequence, $a \leq 1$.

References

[1] M.-P. Béal and J. Senellart, *On the bound of the synchronization delay of a local automaton*, Theoret. Comput. Sci., 205 (1998), pp. 297–306.

[2] J. Berstel, D. Perrin, and C. Reutenauer, *Codes and Automata*, Cambridge University Press, 2007. To appear, see `http://www-igm.univ-mlv.fr/~berstel/LivreCodes/newlivreCodes.pdf`.

[3] J. M. Boë, *Les boîtes*, Theoret. Comput. Sci., 81 (1991), pp. 17–34.

[4] A. Carpi, *On unambiguous reductions of monoids of unambiguous relations*, Theoret. Comput. Sci., 51 (1987), pp. 215–220.

[5] ——, *On synchronizing unambiguous automata*, Theoret. Comput. Sci., 60 (1988), pp. 285–296.

[6] Y. Cesari, *Sur l'application du théorème de Suschkewitsch à l'étude des codes rationnels complets*, in Automata, Languages and Programming (Second Colloq., Univ. Saarbrücken, Saarbrücken, 1974), Springer, Berlin, 1974, pp. 342–350. Lecture Notes in Comput. Sci., Vol. 14.

[7] E. Czeizler and J. Kari, *A tight linear bound on the neighborhood of inverse cellular automata*, in Proceedings of ICALP 2005, Automata, Languages and Programming, vol. 3580 of Lecture Notes in Computer Science, Springer-Verlag, 2005, pp. 410–420.

[8] E. Czeizler and J. Kari, *A tight linear bound on the synchronization delay of bijective automata*, Theoret. Comput. Sci., 380 (2007), pp. 23–36.

[9] J. Friedman, *On the road coloring problem*, Proc. Amer. Math. Soc., 110 (1990), pp. 1133–1135.

[10] G. A. Hedlund, *Endomorphisms and automorphisms of the shift dynamical system*, Math. Systems Theory, 3 (1969), pp. 320–375.

[11] Z. Kohavi, *Switching and Finite Automata Theory*, McGraw-Hill Book Co., New York, 1970. McGraw-Hill Computer Science Series.

[12] D. A. Lind and B. H. Marcus, *An Introduction to Symbolic Dynamics and Coding*, Cambridge University Press, Cambridge, 1995.

[13] M. Nasu, *Local maps inducing surjective global maps of one-dimensional tessellation automata*, Math. Systems Theory, 11 (1977), pp. 327–351.

[14] M. Perles, M. O. Rabin, and E. Shamir, *The theory of definite automata*, IEEE Trans. Electronic Computers, EC-12 (1963), pp. 233–243.

[15] J. Černý, *Poznámka k homogénnym experimentom s konečnými automatmi*, Mat. fyz. čas., SAV 14 (1964), pp. 208–215.

Marie-Pierre Béal
Université Paris-Est
Institut Gaspard-Monge
77454 Marne-la-Vallée Cedex 2
France
e-mail: `beal@univ-mlv.fr`

Eugen Czeizler
Department of Mathematics
FIN-20014 University of Turku
Finland
and
Turku Center for Computer Science
FIN-20520 Turku
Finland
e-mail: `euczei@utu.fi`

Jarkko Kari
Department of Mathematics
FIN-20014 University of Turku
Finland
and
Turku Center for Computer Science
FIN-20520 Turku
Finland
e-mail: `jkari@utu.fi`

Dominique Perrin
Université Paris-Est
Institut Gaspard-Monge
77454 Marne-la-Vallée Cedex 2
France
e-mail: perrin@univ-mlv.fr