

Sofic shifts and synchronizing words I

MARIE-PIERRE BÉAL

Université Paris-Est
Laboratoire d'informatique Gaspard-Monge

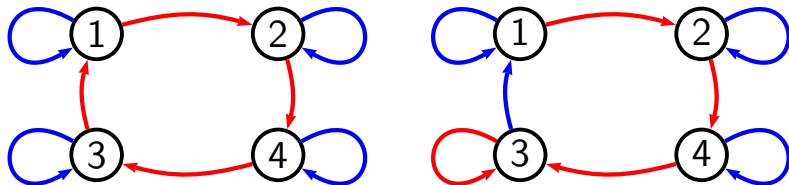


Jorcad September, 2008

Content of the talk

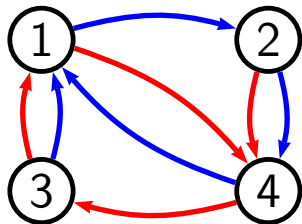
- Synchronizing words
- The Road Coloring Theorem
- The Černý Conjecture
- Application to Huffman compression
- Application to Symbolic Dynamics

Synchronizing words



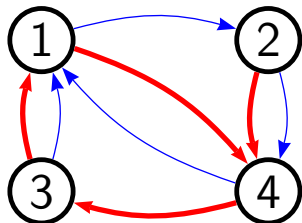
Two irreducible deterministic complete automata. Only one on the right is **synchronized**: the word **RRR** is a synchronizing word (magic word, homing sequence, Rome word).

Another synchronized automaton



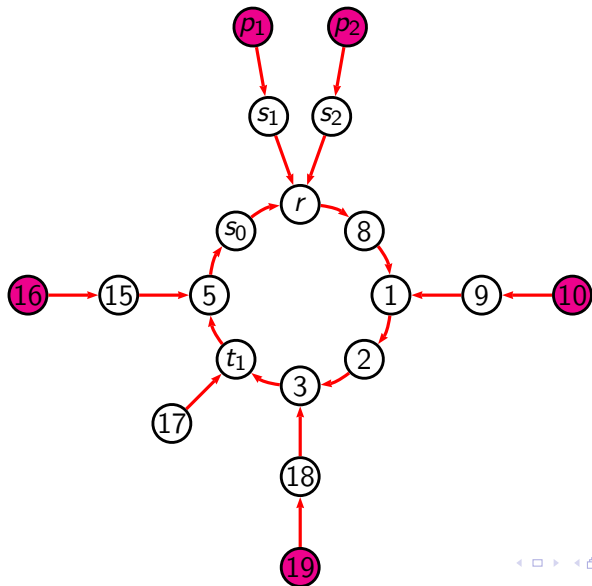
The word *BRB* focuses to state 1.

One-cluster automata



The notion of cluster: this automaton has a **single red cycle** and trees attached to it.

Another one-cluster automaton



The Road Coloring problem (Adler, Goodwin, Weiss, 1977)

A graph is **road colorable** if there is a coloring of its edges such that

- 1 the edges going out of a vertex have distinct colors;
- 2 there is a synchronizing word.

aperiodic graph: the gcd of the cycle lengths is 1.

Theorem (A. Trahtman 2007)

Any irreducible directed graph which is aperiodic and has constant out-degree is road colorable.

admissible: aperiodic + constant out-degree

- **lossless source coding**: the Huffman decoder can be chosen synchronized, hence resistant against errors
- **communication protocols**: test sequences to check whether a protocol conforms to its specification
- **symbolic dynamics**: for two aperiodic shifts of finite type X, Y with the same entropy, there exists a factor map $\varphi : X \rightarrow Y$ which is almost one-to-one. (invertible on "typical" sequences, *i.e.* on bi-infinite sequences which contain a synchronizing word infinitely often to the left).

- Statement of the problem (Adler et al., 1977).
- Graphs with a cycle of prime length (O'Brien, 1981).
- Colorings and eigenvectors (Friedman, 1990)
- Synchronized finite prefix codes (Perrin and Schützenberger, 1992)
- Eulerian graphs (Kari, 2001)
- Solution (Trahtman, 2007)

Trahtman's solution: stable pairs

Definition

A pair of states (p, q) in a colored graph is **synchronizable** if there is a color sequence u such that $p \cdot u = q \cdot v$.

Definition

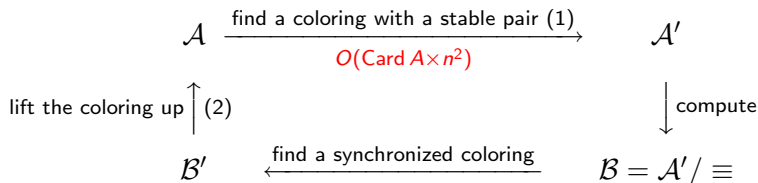
A pair of states (p, q) in a colored graph is **stable** if for any color sequence v , the pair $(p \cdot v, q \cdot v)$ is synchronizable.

Property

p is equivalent to q if (p, q) is a stable pair, is a congruence relation.

The congruence **generated by a stable pair** (s, t) is the least congruence \equiv such that s and t belong to a same class.

Trahtman's solution: a cubic algorithm

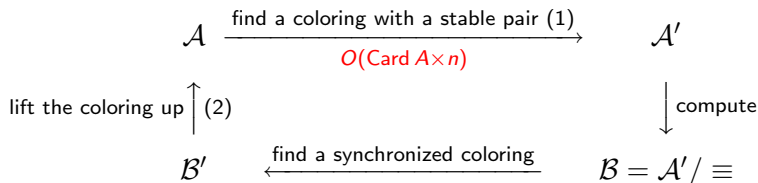


FINDCOLORING(aperiodic colored graph \mathcal{A})

- 1 $\mathcal{A}_0 \leftarrow \mathcal{A}$
- 2 **while** \mathcal{A} has at least two states
- 3 **do** $\mathcal{A}, (s, t) \leftarrow \text{FINDSTABLEPAIR}(\mathcal{A})$
- 4 lift the coloring of \mathcal{A} up to \mathcal{A}_0
- 5 $\mathcal{A} \leftarrow$ the quotient of \mathcal{A} by the congruence generated par (s, t)
- 6 **return** \mathcal{A}_0

(2) was known from Kari et al. 2001

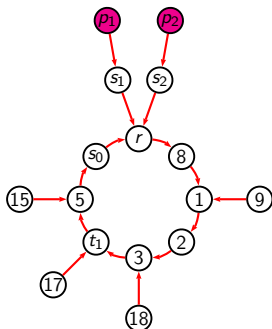
A quadratic algorithm for road coloring



Theorem (B. Perrin 2008)

A synchronized coloring of an n -state admissible graph is computable in time $O(\text{Card } \mathcal{A} \times n^2)$.

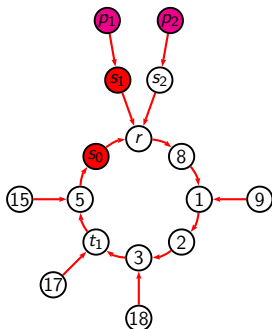
Finding a coloring which has a stable pair



Lemma (Trahtman)

If all maximal states in an irreducible colored graph \mathcal{A} belong to the same tree, then \mathcal{A} has a stable pair.

Finding a coloring which has a stable pair



Lemma (Trahtman)

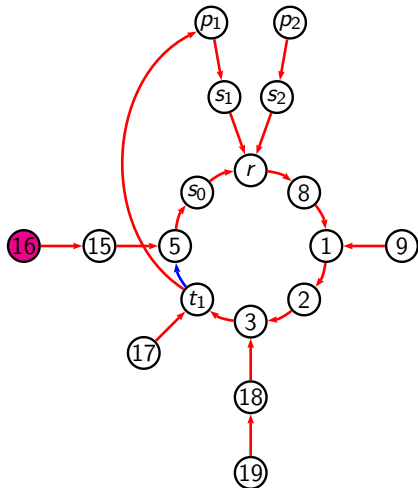
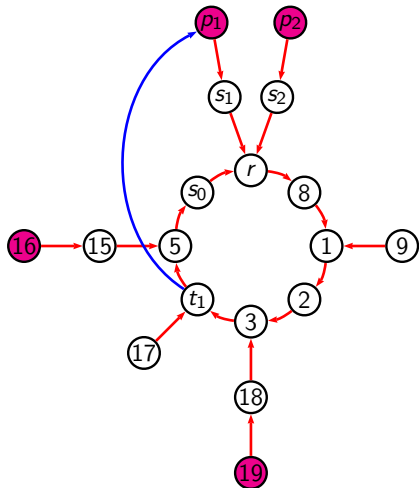
If all maximal states in an irreducible colored graph \mathcal{A} belong to the same tree, then \mathcal{A} has a stable pair.

Finding a coloring with all maximal states in a same tree

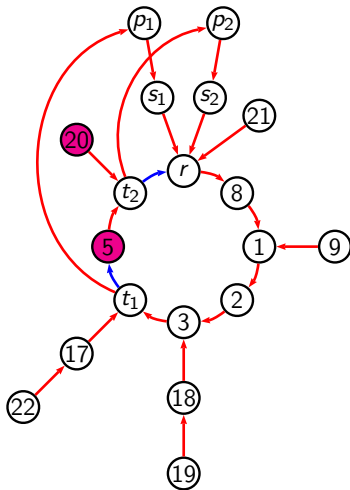
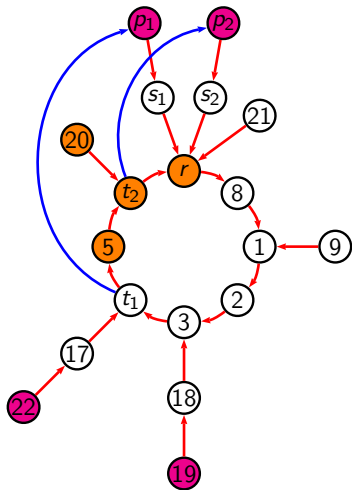
Sequence of flips

- do a sequence of flips of edges (at each time, one flips two edges going out of some state);
- the sequence of flips depends of the position of the roots of the maximal trees on the cycle and of the blue edges.

A simple case: flip the outgoing edges of t_1



A harder case: flip the outgoing edges of t_1 and of t_2



The height of the tree \mathcal{T} , in orange is equal to the maximal level.

Questions

- What is the size of a shortest synchronizing sequence after running the algorithm? (there is a trivial cubic upper bound).
- Does the algorithm also gives a way to find a shortest synchronizing word? (No)
- Are there algorithms to find "short" synchronizing words?

Conjecture (Černý 1964)

Each synchronized n -state automaton graph possesses a synchronizing word of length at most $(n - 1)^2$.

True in some particular cases:

- circular automata with a prime-length red cycle (Pin, 1980)
- circular automata (Dubuc, 1998)
- eulerian automata (Kari, 2001)
- aperiodic automata (Trahtman, 2002)
- circular automata and eulerian automata, using rational series (B, unpublished 2003)
- automata preserving a chain of partial orders (Volkov, 2007)
- strongly transitive automata (using rational series also) (Carpi, 2008)
- ...

Černý's Conjecture for one-cluster automata

Proposition (B. Perrin 2008)

Let \mathcal{A} be a synchronized n -state deterministic, complete and irreducible automaton over an alphabet A . If \mathcal{A} is one-cluster, then it has a synchronizing word of length at most $2(n - 1)(n - 2)$.

Proof

Proof: If u is a word, M_u is the transition matrix of the action of u on the states.

$$\begin{aligned}(M_u)_{pq} &= 1 \text{ if } p \cdot u = q, \\ (M_u)_{pq} &= 0 \text{ otherwise.}\end{aligned}$$

Let C be the red cycle of length m and P be a subset of set of states with $P \cap C \neq \emptyset$. We note ℓ the maximal level of the states. A word u is said to be **(P, C) -augmenting** if

$$\text{Card}((Pu^{-1}) \cap C) > \text{Card}(P) \cap C.$$

Let \mathbf{p} and \mathbf{c} be the characteristic row vectors of P and c . Then u is a (P, C) -augmenting word if and only if

$$\mathbf{c}M_u\mathbf{p}^t > \mathbf{c} \cdot \mathbf{p}^t.$$

We denote by S, T the rational series (of rank at most m)

$$\langle S, u \rangle = \mathbf{c}(M_u - I)\mathbf{p}^t \quad \text{and} \quad \langle T, u \rangle = \langle S, ua^\ell \rangle.$$

Let us assume that $C \not\subseteq P$. We show that there is (P, C) -augmenting word of size at most $2(m - 1) + \ell$.

We have

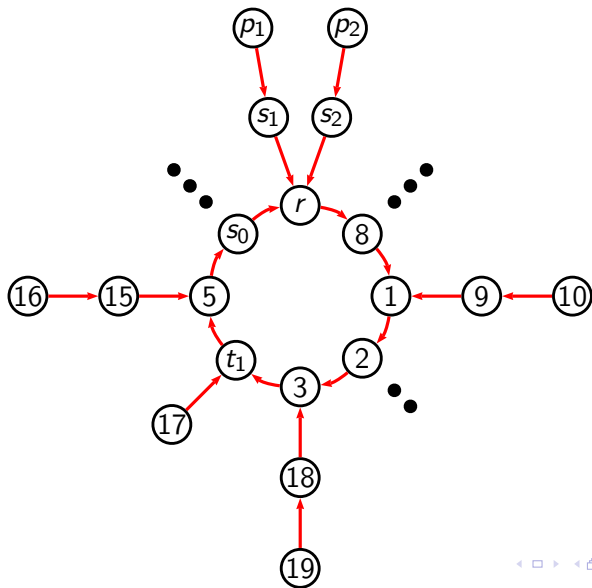
- T is non null;
- If there is u of length at most $m - 1$ such that $\langle T, u \rangle > 0$, then ua^ℓ is a (P, C) -augmenting;
- Otherwise, there is a word u of length at most $m - 1$ such that $\langle T, u \rangle < 0$;
- $\sum_{i=0}^{m-1} \langle T, ua^i \rangle = 0$.

Indeed:

$$\begin{aligned}
 \sum_{i=0}^{m-1} \langle T, ua^i \rangle &= \sum_{i=0}^{m-1} \langle S, ua^\ell a^i \rangle, \\
 &= \sum_{i=0}^{m-1} \mathbf{c}(M_{ua^\ell a^i} - I)\mathbf{p}^t, \\
 &= \sum_{i=0}^{m-1} \mathbf{c}M_{ua^\ell}M_{a^i}\mathbf{p}^t - \sum_{i=0}^{m-1} \mathbf{c}\mathbf{p}^t, \\
 &= \mathbf{c}M_{ua^\ell} \left(\sum_{i=0}^{m-1} M_{a^i} \right) \mathbf{p}^t - m\mathbf{c}\mathbf{p}^t, \\
 &= 0.
 \end{aligned}$$

Since $\mathbf{c}M_{ua^\ell} \left(\sum_{i=0}^{m-1} M_{a^i} \right) = \text{Card}(C)\mathbf{c} = m\mathbf{c}$.

Picture



As a consequence, there is a (P, C) -augmenting word u of size at most $m - 1 + m + \ell - 1$.

One can assume $m \leq n - \ell$, $1 \leq \ell \leq n - 2$.

There is a synchronizing word of length at most

$$\begin{aligned} & \ell + (m - 1)(m - 1 + m + \ell - 1) \\ \leq & n - 2 + (n - 1 - \ell)(n - 1 - \ell + n - 1) \\ \leq & n - 2 + (n - 2)(2n - 3) \\ \leq & 2(n - 2)(n - 1). \end{aligned}$$

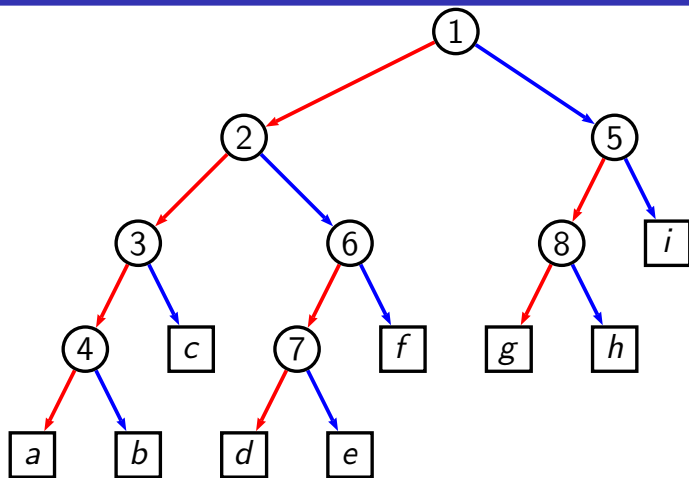
Back to the road coloring (more complicated)

Theorem (B. Perrin 2008)

*One can compute a synchronized **one-cluster** coloring of an n -state admissible graph in time $O(\text{Card } A \times n^2)$.*

Hence the result has a homing sequence of length at most $2(n-1)(n-2)$.

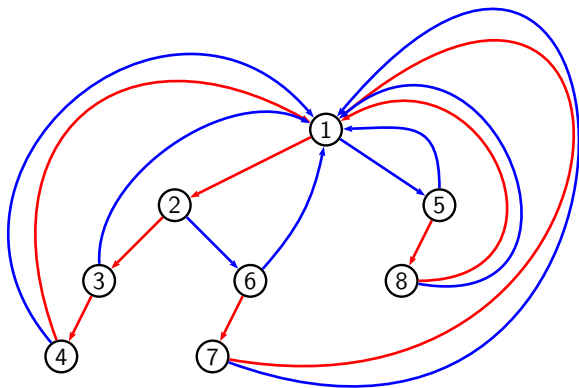
Application to Huffman compression



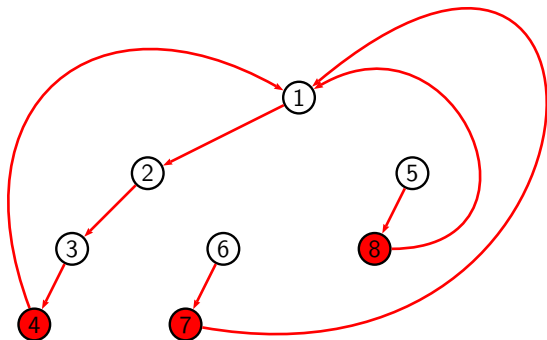
$$X = (\text{RR}, \text{RB}, \text{B})^2$$

$a : 1/16, b : 1/16, c : 1/8, d : 1/16, e : 1/16, f : 1/8, g : 1/8, h :$
 $1/8, i : 1/4$

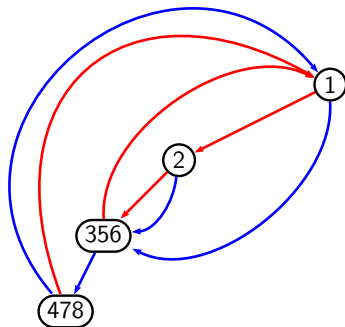
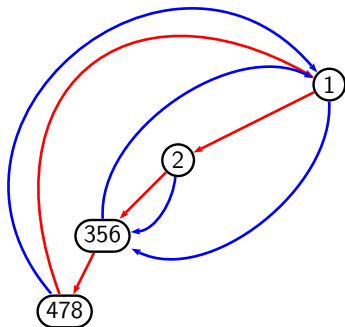
Application to Huffman compression



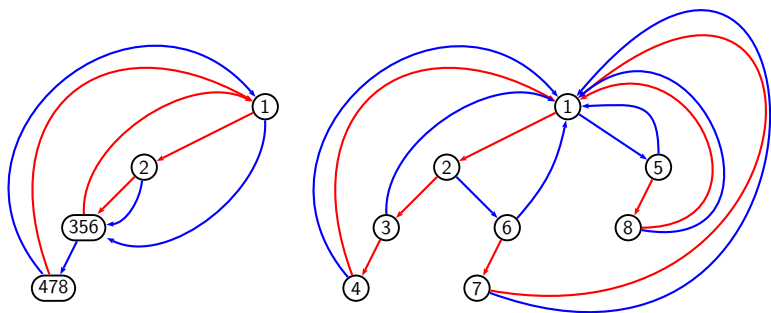
Application to Huffman compression



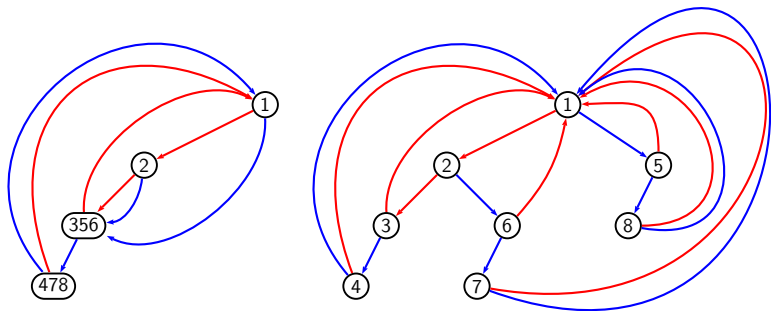
Quotient + flip at 356



Lifting up the flips



Lifting up the flips



$a \leftrightarrow \text{RRBR}$

$b \leftrightarrow \text{RRBB}$

$c \leftrightarrow \text{RRR}$

$d \leftrightarrow \text{RBR}$

...

The sequence **RBR** is a homing sequence.

The Road Coloring problem for periodic graphs

The **period** of a graph is the gcd of the lengths of the cycles.

The **minimal rank** of a colored graph (Q, E) is the minimal cardinality of the set $Q \cdot u$ for all colored sequences u .

A synchronized colored graph has minimal rank 1.

Theorem (B. Perrin 2008)

A coloring of an n -state graph (with constant out-degree) whose minimal rank is the period of the graph is computable in time $O(\text{Card } A \times n^2)$.