Synchronized automata (1)

Marie-Pierre Béal, Dominique Perrin

May 20, 2012
Outline

- Černý conjecture (DP)
  - Synchronized automata
  - Nondeterministic automata
  - Unambiguous automata
  - Aperiodic automata
  - Rational series
- Road coloring (MPB)
We first consider deterministic automata. A synchronizing word is a word $w$ such that there are paths labeled by $w$ and all have the same terminal state. An automaton which has a synchronizing word is called synchronized. Elementary arguments give the cubic bound $\frac{1}{2}n(n-1)^2$ for the length of a synchronizing word if there is one. Computing the minimal length of a synchronizing word is NP-complete (Eppstein, 1990).
A deterministic automaton with four vertices and two edges going out of each vertex labeled $B$ or $R$.

$BRB$ is a synchronization sequence: starting anywhere, it leads to 1.
Another example

With a different coloring:

there is no synchronizing sequence: the sets $\{1, 3\}$ and $\{2, 4\}$ cannot be collapsed further:
Cubic bounds

Let $\mathcal{A}$ be a synchronized automaton with $n$ states. For each set $P$ of states there is a word $w$ of length at most $n(n - 1)/2$ such that $\text{Card}(P \cdot w) < \text{Card}(P)$. By iteration, this proves the existence of a synchronizing word of length at most $\frac{1}{2}n(n - 1)^2$.

Refinement: a sequence $P_1, P_2, \ldots$ of $k$ element subsets of an $n$ element set is called 2-renewing if each $P_i$ contains a pair $R_i$ such that $R_i$ is not contained in $P_j$ for $j < i$.

**Proposition (Frankl, 1982)**

A 2-renewing sequence has at most $\binom{n-k+2}{2}$ elements.

Consequence: existence of a synchronizing word of length at most $(n^3 - n)/6$. 
Proposition (Eppstein, 1990)

The existence of a synchronizing word of length \( \ell \) is NP-complete.

Proof: polynomial reduction from the satisfiability of boolean formulas. Given a formula \( \psi = c_1 \land c_2 \land \ldots \land c_m \) with each \( c_i \) a disjunction of \( x_1, x_2, \ldots, x_n \) or their negations, one builds an automaton \( A(\psi) \) which has a synchronizing word on length \( n \) if and only if \( \psi \) is satisfiable.

As usual a binary word of length \( n \) is interpreted as a truth assignnement for the \( n \) variables.
Conjecture (Černý, 1964)

Any complete $n$-state synchronized automaton has a synchronizing word of length at most $(n - 1)^2$.

The hypothesis that the automaton is complete is perhaps not necessary?
Figure: The automata $C_4, C_5, C_6$

The word $(ba^{n-1})^{n-2}b$ is a synchronizing word of length $(n - 1)^2$. There is no shorter one.
A nice proof (Ananisev, Gusev, Volkov)

The cycles of $D_n$ are nonnegative integer combinations of $n$ and $n - 1$.

**Lemma**

*If $k, \ell \geq 1$ are relatively prime then $k\ell - k - \ell$ is the largest integer not expressible as a nonnegative integer combination of $k, \ell$.**
Proof (Cont’d)

Let $w$ be a synchronizing word of minimal length for $C_n$. Set $w = w'b$. Then $Q \cdot w' = \{0, 1\}$. Let $v$ be obtained from $w'$ replacing each $ba$ by $c$ (there are no consecutive $b$ in $w$). Then $vc$ is synchronizing for $D_n$ and thus $D_n$ has cycles of all lengths $\geq |vc|$. This forces

$$|vc| > n(n - 1) - n - (n - 1) = n^2 - 3n + 1.$$  

We cannot have $|vc| = n^2 - 3n + 2$ because $1 \cdot vc = 2$ would imply that $vc$ minus its first letter sends 2 to 2 although there is no cycle of length $n^2 - 3n + 1$ in $D_n$. Thus $|vc| \geq n^2 - 3n + 3$ and $|w'| = |v| - (n - 2) \geq n^2 - 2n$ and finally $|w| \geq n^2 - 2n + 1$. 

The rank of a word $w$ in a deterministic automaton $A = (Q, E)$ is the size of its image $Q \cdot w$. A synchronizing word has rank 1. It has been conjectured by Pin that if an automaton admits a word of rank at most $k$, then there exists such a word of length at most $(n - k)^2$. This generalizes the Černý conjecture. Pin’s conjecture was proved to be false (Kari, 2001). The counterexample corresponds to $n = 6$ and $k = 2$. A new conjecture proposed by Volkov states that an automaton of minimal rank $k$ admits a word of rank $k$ of length at most $(n - k)^2$. 
Kari’s automaton

Shortest words of rank 2: $x = baabababaabbabaab$ and $y = baababaabaababaab$ of length 17. Shortest synchronizing word $xaababaab$ of length $24 = 5^2 - 1$. 

Marie-Pierre Béal, Dominique Perrin

Synchronized automata (1)
A funny conjecture

**Conjecture (Trahtman, 2006)**

The set of $n$-state complete deterministic automata with minimal synchronizing word of length $(n - 1)^2$ consists of the sequence of Černý automata and eight automata of size at most 6.
A word $w$ is \textit{synchronizing} for a nondeterministic automaton if there is at least one path labeled by $w$ and, for any states $p, q, r, s$,

$$p \xrightarrow{w} q, \ r \xrightarrow{w} s \ \text{implies} \ \ p \xrightarrow{w} s \ \text{and} \ r \xrightarrow{w} q.$$ 

The automaton is synchronized if there is a synchronizing word. This is consistent with the previous definition. Indeed, if the automaton is deterministic, the condition implies that $q = s$. 
Associate to a word $w$ the relation $\varphi_A(w) = \{(p, q) | p \xrightarrow{w} q\}$. The **rank** of a relation $m$ on a set $Q$ is the minimal cardinality of a set $R$ such that $m = uv$ with $u$ a $Q \times R$ relation and $v$ an $R \times Q$ relation.

A word $w$ is synchronizing for a nondeterministic automaton $A$ if and only if the relation $\varphi_A(w)$ has rank one. Indeed, if $w$ is synchronizing, then $\varphi_A(w) = uv$ where $u$ is the column $Q$-vector $v$ is the row $Q$-vector defined as follows. One has $u_p = 1$ if and only if there is a path labeled $w$ starting from $p$ and $v_q = 1$ if and only if there is a path labeled $w$ ending in $q$. 
Example

The automaton below is synchronized. Indeed, the word $ab$ is such that

$$\varphi_A(ab) = \begin{bmatrix} 0 & 0 & 0 \\ 0 & 0 & 1 \\ 0 & 0 & 1 \end{bmatrix} = \begin{bmatrix} 0 \\ 1 \\ 1 \end{bmatrix} \begin{bmatrix} 0 & 0 & 1 \end{bmatrix}$$

and thus is synchronizing.

![Diagram of a synchronized automaton](image)

**Figure:** A synchronized nondeterministic automaton.
A nondeterministic automaton is unambiguous if, for any pair \( p, q \) of states, and any word \( w \), there is at most one path labeled by \( w \) going from \( p \) to \( q \). A deterministic automaton is unambiguous. The converse is not true.

One may check that an automaton is unambiguous by computing its square. The square of \( \mathcal{A} \) is the automaton on \( Q \times Q \) with edges \((p, q) \xrightarrow{a} (r, s)\) if \( p \xrightarrow{a} r \) and \( q \xrightarrow{a} s \) are edges of \( \mathcal{A} \). The automaton \( \mathcal{A} \) is unambiguous if and only if there is no path in its square of the form \((p, p) \xrightarrow{u} (r, s) \xrightarrow{v} (q, q)\) with \( r \neq s \).
Example

Let $\mathcal{A}$ be the automaton represented on the left.

![Diagram of automata]

**Figure:** An unambiguous automaton and part of its square.

This automaton is unambiguous as one may check by computing the square of the automaton $\mathcal{A}$ represented on the right (with only the states accessible from the states $(p, p)$ represented).
Let $f : A^* \rightarrow A^*$ be a morphism with $f(a) \in aA^*$. Let $x = f^\omega(a)$. Let

$$l = \{|f(x_0 \cdots x_n)| \mid n \geq 0\}$$

The morphism $f$ is $n$-recognizable if for $i \in l$, $x_{i-n} \cdots x_{i+n} = x_{j-n} \cdots x_{j+n}$ implies $j \in l$.

Example: The Fibonacci morphism $f(a) = ab$, $f(b) = a$ is 2-recognizable.

Martin’s theorem: any primitive morphism such that $x$ is not periodic is recognizable.
Using the fact that

(i) for any code $X$ there is an unambiguous automaton $A = (Q, 1, 1)$ recognizing $X^*$

(ii) The fixpoint of a primitive morphism is uniformly recurrent

one obtains

**Proposition**

Assume that $f$ is injective and primitive and let $A = (Q, 1, 1)$ be an unambiguous automaton recognizing $f(A)^*$. If there is a synchronizing word $w \in F(x)$, then

(i) There is an $n \geq 1$ such that any word of $F$ of length $n$ is synchronizing.

(ii) The morphism $f$ is $n$-recognizable.
Set \( f(a) = ab \) and \( f(b) = a \). Then \( x = abaababa \cdots \) is the Fibonacci word. The automaton \( A \) is represented below.

![Automaton Diagram]

The word \( a \) is synchronizing. Thus \( aa, ab \) and \( ba \) are synchronizing and \( f \) is 2-recognizable.
Set $f(a) = ab$ and $f(b) = ba$. Then $x = abbabaab \cdots$ is the Thue-Morse word. The automaton $A$ is represented below.

```
2 --- a --- 1
   \
   \ b
 --- a --- 3
```

The words $aa$ and $bb$ are synchronizing. Thus $f$ is 5-recognizable.
A nondeterministic automaton is strongly connected if, for any word $w$, there exists a pair $p, q$ of states such that $p \xrightarrow{w} q$ is a path.

The following result gives a cubic upper bound for the length of synchronizing words.

**Theorem (Carpi, 1988)**

Let $A$ be an $n$-state strongly connected unambiguous automaton. If $A$ is synchronized, it has a synchronizing word of length at most $(n^2 - n + 2)(n - 1)/2$. 
We introduce some terminology concerning monoids of relations on a set. For a relation $m$ on a set $Q$, we denote indifferently $(p, q) \in m$ or $m_{pq} = 1$, considering the relation either as a subset of $Q \times Q$ or as a boolean $Q \times Q$ matrix. We denote by $m_p^*$ the row of index $p$ of the matrix $m$, which is the characteristic vector of the set $\{ q \in Q \mid (p, q) \in m \}$.

A monoid of relations $M$ on a set $Q$ is unambiguous if for any $m, n \in M$ and $p, q \in Q$ there exists an at most one element $r$ of $Q$ such that $(p, r) \in m$ and $(r, q) \in n$. The monoid is transitive if for any $p, q \in M$ there is an $m \in M$ such that $(p, q) \in m$.

An automaton is unambiguous if and only if its transition monoid is unambiguous. The automaton is strongly connected if and only if its transition monoid is transitive.
We use the following lemmas.

A row of an element of $M$ which is maximal among the rows of the elements of $M$ is called a maximal row.

**Lemma**

Let $M$ be a transitive unambiguous monoid of relations not containing zero. If $v$ is a maximal row, then $vm$ is a maximal row for any $m \in M$.

**Lemma**

Let $M$ be a transitive unambiguous monoid of relations not containing zero. For two elements $m, m'$ of $M$, if $m \leq m'$ then $m = m'$.
The following lemma is the key of Carpi’s theorem.

**Lemma**

For a state \( p \in Q \) and a word \( u \in A^* \), if \( \varphi(u)_p^* \) is not a maximal row, 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 of Carpi’s theorem.

By the last Lemma and its symmetric form, there exist 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 \(A^* \times Q\) such that, with 
\(x_i = \varphi (u_1 \cdots u_1)_{p_i}^*\) and \(y_j = \varphi (v_1 \cdots v_j)_{q_i}^*\),

(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) \(x_s\) is a maximal row and \(y_t\) is a maximal column.

Let \(u = u_s \cdots u_1\) and \(v = v_1 \cdots v_t\). 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 
\(x_s y_t = 1\).
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 $w = vz\nu$ is such that $y_t x_s \leq \wp(w)$. By the second Lemma, this implies $\wp(w) = y_t x_s$, whence the conclusion.
We illustrate the proof on the automaton of the previous example. We start with $u_1 = v_1 = \varepsilon$ and $p_1 = q_1 = 1$. The row $\varphi(\varepsilon)_1 = [1 \ 0 \ 0]$ is not maximal but $\varphi(\varepsilon)_1 < \varphi(a)_3 = [1 \ 1 \ 0]$. Thus we choose $u_2 = a$ and $p_2 = 3$. Symmetrically, the column $\varphi(\varepsilon)_1^t = [1 \ 0 \ 0]^t$ is not maximal but $\varphi(\varepsilon)_1^t < \varphi(b)_3^t = [1 \ 1 \ 0]^t$. Thus we choose $v_2 = b$ and $q_2 = 3$. Then $ab$ is a synchronizing word.
Question: is Černý conjecture true for strongly connected unambiguous automata?