# ČERNY CONJECTURE FOR DFA ACCEPTING STAR-FREE LANGUAGES

A.N. Trahtman*

Bar-Ilan University, Dep. of Math. and St., 52900, Ramat Gan, Israel

ICALP, Workshop synchr. autom., Turku, Finland, 2004

**Abstract.** A word $w$ is called synchronizing (recurrent, reset) word of an automaton if $w$ sends all states of the automaton on a unique state. Černy had conjectured in 1964 that an $n$-state automaton with non-empty set of synchronizing words possesses a synchronizing word of length not greater than $(n-1)^2$.
We consider automata accepting star-free languages (or aperiodic, viz all subgroups of the syntactic monoid of the language are trivial). The minimal length of synchronizing word for $n$-state aperiodic automaton is not greater than $n(n-1)/2$.
Thus the Černy conjecture holds true for automata accepting star-free languages. A new wording of necessary and sufficient conditions of synchronizability are found for an arbitrary and for an aperiodic automaton.
http://www.cs.biu.ac.il/∼trakht/Testas.html

**Keywords**: synchronizing word, star-free language, deterministic finite automata.

## Introduction

We consider a deterministic finite automaton (DFA) with complete state transition graph $\Gamma$ and transition semigroup $S$ over alphabet $\Sigma$.
An important problem with a long story is the estimation of the shortest length of synchronizing word. Most known as a Černy conjecture, it was aroused independently by distinct authors. Jan Černy had found in 1964 [2] an $n$-state automaton with minimal length of synchronizing word of $(n-1)^2$. He had conjectured that it is the minimal length of synchronizing word in $\Sigma$ for any $n$-state automaton. The conjecture is valid for big list of objects [1], [3], [4], [6], [9], [8], but in general the answer is open. This simply looking conjecture is now one of the most longstanding open problem in the theory of finite automata. The best known upper bound is now equal to $(n^3-n)/6$ [5], [7], [8]. For reach and intriguing story of investigation in this area see [8], [11].
The existence of some non-trivial subgroup in the transition semigroup of the automaton is essential in many investigation of Černy conjecture. We use another approach and consider transition semigroups without non-trivial subgroups. It is

* Email: trakht@macs.biu.ac.il

wide class of automata that accept star-free languages also known as languages of star height 0. The class was involved and studied by Schŭtzenberger [12]. Star-free languages play a significant role in the formal languages theory.
We prove that the aperiodic DFA with sink state has a synchronizing word of length not greater than $n(n-1)/2$ and therefore the Černy conjecture holds true for DFA accepting star-free languages.
In the case the aperiodic DFA is also strongly connected, the obtained upper bound was essentially improved by Volkov. He reduced the estimation to $n(n+1)/6$.
A necessary and sufficient conditions of synchronizeability of an arbitrary automaton are known [2] and we present them in the following form:
An automaton with transition graph $\Gamma$ is synchronizing iff $\Gamma^2$ has sink state.
It is a base for an implemented quadric algorithm and is essential for algorithms suggested by Eppstein [4] and Ananichev. The following statement holds for automata accepting star-free languages:

An aperiodic automaton with sink state is synchronizing.

## Preliminaries

Let us consider a deterministic finite automaton with state transition graph $\Gamma$ and transition semigroup $S$ over alphabet $\Sigma$.
A maximal strongly connected component of a directed graph will be denoted for brevity as **SCC**.
A semigroup without non-trivial subgroups is called *aperiodic*. A DFA with syntactic aperiodic semigroup is called *aperiodic* too.
If there exists a path from the state $\mathbf{p}$ to the state $\mathbf{q}$ and the transitions of the path are consecutively labelled by $\sigma_1, ..., \sigma_k$ then for $s = \sigma_1 ... \sigma_k$ let us write $\mathbf{q} = \mathbf{p}s$.
The state $\mathbf{q}$ is called *sink* state if for every state $\mathbf{p}$ there exists a word $s$ such that $\mathbf{p}s = \mathbf{q}$.
The binary relation $\beta$ is called *stable* if for any pair of states $\mathbf{q}, \mathbf{p}$ and any $\sigma \in \Sigma$ from $\mathbf{q}\ \beta\ \mathbf{p}$ follows $\mathbf{q}\sigma\ \beta\ \mathbf{p}\sigma$.
The graph $\Gamma$ is called *complete* if for every vertex $\mathbf{p}$ and every $\sigma \in \Sigma$ the vertex $\mathbf{p}\sigma$ exists.
$|s|$ - the length of the word $s$ in alphabet $\Sigma$.
$|P|$ - the size of the set of states of the automaton (of vertices of the transition graph) $P$.
Let $Ps$ denote the mapping of the graph (of the automaton) $P$ by help of $s \in \Sigma^*$.
A word $s \in \Sigma^+$ $(\in S)$ is called *synchronizing* word of an automaton with transition graph $\Gamma$ if $|\Gamma s| = 1$.
The direct product $\Gamma^2$ of two copies of the transition graph $\Gamma$ over an alphabet $\Sigma$ consists of pairs $(\mathbf{p}, \mathbf{q})$ and edges $(\mathbf{p}, \mathbf{q}) \to (\mathbf{p}\sigma, \mathbf{q}\sigma)$ labelled by $\sigma$. Here $\mathbf{p}, \mathbf{q} \in \Gamma$, $\sigma \in \Sigma$ [13].
If $\rho$ is a congruence of a DFA $A$, we denote by $[\mathbf{q}]_\rho$ the $\rho$-class containing the state $\mathbf{q}$ of $A$. The *quotient* $A/\rho$ is the automaton with the set of states $[\mathbf{q}]_\rho$ and

the transition function defined by the rule $[\mathbf{q}]_\rho \sigma = [\mathbf{q}\sigma]_\rho$ for any $\sigma \in \Sigma$.
We shall use also the following concepts:
An $SCC$ $M$ from $\Gamma^2$ will be called *almost minimal* if for every its state $(\mathbf{p}, \mathbf{q})$ $\in M$ and for every $\sigma \in \Sigma$ such that $\mathbf{p}\sigma \neq \mathbf{q}\sigma$ there exists a word $s$ such that $(\mathbf{p}\sigma, \mathbf{q}\sigma)s = (\mathbf{p}, \mathbf{q})$.
Let $\Gamma(M)$ be the set of coordinates of states from almost minimal $SCC$ $M$.
Let us define a relation $\succ_M$ where $M$ is an almost minimal $SCC$. Suppose $\mathbf{p} \succ_M \mathbf{q}$ if $(\mathbf{p}, \mathbf{q}) \in M$ and let $\succ_M$ be the transitive closure of this relation. Let $\succeq_M$ be the reflexive closure and $\rho_M$ be equivalent closure of the relation $\succ_M$. A cycle defined by the transitive relation $\succ_M$ let as call *t-cycle*.
So $\mathbf{r} \succ_M \mathbf{q}$ if there exists a sequence of states $\mathbf{r} = \mathbf{p}_1, ..., \mathbf{p}_n = \mathbf{q}$ such that $(\mathbf{p}_i, \mathbf{p}_{i+1}) \in M$ for $i = 1, ..., n-1$ and almost minimal $SCC$ $M$. In the case $\mathbf{r} = \mathbf{q}$ we have $t$-cycle with at least two states.

## 1 The graph $\Gamma^2$

**Lemma 11** *The relation $\succeq_M$ is stable. The equivalent closure $\rho_M$ of the relation $\succ_M$ is a congruence.*

Proof.Suppose $\mathbf{u}$ $\rho_M$ $\mathbf{v}$. Then there exist a sequence of states $\mathbf{u} = \mathbf{p}_1, ..., \mathbf{p}_n = \mathbf{v}$ such that for every integer $i < n$ at least one of the states $(\mathbf{p}_{i+1}, \mathbf{p}_i)$, $(\mathbf{p}_i, \mathbf{p}_{i+1})$ belongs to the almost minimal $SCC$ $M$. Therefore in the sequence of states $\mathbf{r}s = \mathbf{p}_1 s, ..., \mathbf{p}_n s = \mathbf{q}s$ for any two distinct neighbors $\mathbf{p}_i s, \mathbf{p}_{i+1} s$ the state $(\mathbf{p}_i s, \mathbf{p}_{i+1} s)$ or dual belongs to $M$. Consequently, $\mathbf{r}s$ $\rho_M$ $\mathbf{q}s$.
If for every integer $i < m$ the state $(\mathbf{p}_i, \mathbf{p}_{i+1})$ belongs to $M$ then $(\mathbf{p}_i s, \mathbf{p}_{i+1} s) \in M$ or $\mathbf{p}_i s = \mathbf{p}_{i+1} s$ and therefore $\mathbf{p}_i s \succeq_M \mathbf{p}_{i+1} s$. Consequently, $\mathbf{u}s \succeq_M \mathbf{v}s$.

From the definition of the relation $\succ_M$ follows

**Proposition 12** *If $\mathbf{r} \succ_M \mathbf{q}$ and $\mathbf{r}s \notin \Gamma(M)$ for some word $s$, then $\mathbf{r}s = \mathbf{q}s$.*

Let us present a following new wording of some result from [2]:

**Lemma 13** *An automaton $A$ with transition graph $\Gamma$ is synchronizing if and only if $\Gamma^2$ has a sink state.*

Proof. Let $s$ be synchronizing word of $A$. Then $\Gamma^2 s$ is a sink state of $\Gamma^2$.
Conversely, the coordinates of a sink state of $\Gamma^2$ obviously are equal and let $(\mathbf{t}, \mathbf{t})$ be a sink state. For any state $(\mathbf{p}, \mathbf{q})$ from $\Gamma^2$ there exists a word $s$ such that $(\mathbf{p}, \mathbf{q})s = (\mathbf{t}, \mathbf{t})$. We have $\mathbf{p}s = \mathbf{q}s = \mathbf{t}$ for any pair of states $\mathbf{p}, \mathbf{q}$ from $\Gamma$. Some product of such words $s$ for distinct pairs of states from $\Gamma$ presents a synchronizing word of the graph $\Gamma$.

**Lemma 14** *The sets of synchronizing words of the graphs $\Gamma$ and $\Gamma^2$ coincide.*

Proof. Let $s$ be a synchronizing word of the graph $\Gamma$. Then for every state $\mathbf{p}$ and fixed state $\mathbf{q}$ from $\Gamma$ holds $\mathbf{p}s = \mathbf{q}$. Therefore for every state $(\mathbf{p}, \mathbf{r})$ from $\Gamma^2$

holds $(\mathbf{p},\mathbf{r})s = (\mathbf{q},\mathbf{q})$. Thus $s$ is a synchronizing word of the graph $\Gamma^2$.
Now let $t$ be a synchronizing word of the graph $\Gamma^2$. Then for every state $(\mathbf{p},\mathbf{r})$ and fixed state $(\mathbf{q},\mathbf{v})$ from $\Gamma^2$ holds $(\mathbf{p},\mathbf{r})t = (\mathbf{q},\mathbf{v})$. Therefore $\mathbf{p}t = \mathbf{q}$ for arbitrary $\mathbf{p}$ from $\Gamma$ and $\mathbf{r}t = \mathbf{v}$ for arbitrary $\mathbf{r}$ from $\Gamma$. Consequently, $\mathbf{v} = \mathbf{q}$ and $t$ is a synchronizing word of the graph $\Gamma$.

## 2 Aperiodic automata

Let us recall that the syntactic semigroup of star-free language is finite and aperiodic [12] and the semigroup satisfies identity $x^n = x^{n+1}$ for some suitable $n$. Therefore for any state $\mathbf{p} \in \Gamma$, any $s \in S$ and for some suitable $k$ holds $\mathbf{p}s^k = \mathbf{p}s^{k+1}$.

**Lemma 21** *Let $A$ be an aperiodic automaton with $n$ states. Then the existence of sink state in $A$ is equivalent to the existence of synchronizing word.*

Proof. It is clear that the existence of synchronizing word implies the existence of sink state.
So let us consider a $DFA$ with at less one sink state. For any state $\mathbf{p}$ and any sink state $\mathbf{p_0}$ there exists an element $s$ from transition semigroup $S$ such that $\mathbf{p}s = \mathbf{p_0}$. $S$ is aperiodic, whence for some integer $m$ holds $s^{m-1} \neq s^m = s^{m+1}$. Therefore $\mathbf{p}s^m = \mathbf{p_0}s^m$, whence $s^m$ is a synchronizing word for these two states and the state $\mathbf{p_0}s^m$ is a sink state too. We repeat the process reducing the number of states on each step. The product of all such powers $s^m$ maps all states of the automaton on some sink state. So we obtain in this way a synchronizing word.

Let us go to the key lemma of the proof.

**Lemma 22** *Let a DFA with the transition graph $\Gamma$ be aperiodic.*
*Then the graph $\Gamma$ has no $t$-cycle and the quasi-order $\succeq_M$ is a partial order.*

Proof. Suppose the states $\mathbf{p}_1 \succ_M \mathbf{p}_2, .., \mathbf{p}_{m-1} \succ_M \mathbf{p}_m = \mathbf{p}_1$ form $t$-cycle of the minimal size $m$ for some almost minimal $SCC$ $M$.
Let us establish that $m > 2$. Indeed, $\mathbf{p}_1 \neq \mathbf{p}_2$ by the definition of the relation $\succ_M$, whence $m > 1$. If $m = 2$ then two states $(\mathbf{p}_1,\mathbf{p}_2)$ and $(\mathbf{p}_2,\mathbf{p}_1)$ belong to common $SCC$. For some element $u$ from transition semigroup $S$, we have $(\mathbf{p}_1,\mathbf{p}_2)u = (\mathbf{p}_2,\mathbf{p}_1)$. Therefore $\mathbf{p}_1u = \mathbf{p}_2$, $\mathbf{p}_2u = \mathbf{p}_1$, whence $\mathbf{p}_1u^2 = \mathbf{p}_1 \neq \mathbf{p}_1u$. It implies $\mathbf{p}_1u^{2k} = \mathbf{p}_1 \neq \mathbf{p}_1u = \mathbf{p}_1u^{2k+1}$ for any integer $k$. However, semigroup $S$ is finite and aperiodic and therefore for some $k$ holds $u^{2k} = u^{2k+1}$, whence $\mathbf{p}_1u^{2k} = \mathbf{p}_1u^{2k+1}$. Contradiction.
Thus we can assume that $m > 2$ and suppose that the states $\mathbf{p}_1, \mathbf{p}_2, \mathbf{p}_3$ are distinct. For some element $s \in S$ and for the states $\mathbf{p}_1$, $\mathbf{p}_2$, $\mathbf{p}_3$ from considered $t$-cycle holds $(\mathbf{p}_1,\mathbf{p}_2)s = (\mathbf{p}_2,\mathbf{p}_3)$. We have

$$\mathbf{p}_2 = \mathbf{p}_1s,\ \mathbf{p}_3 = \mathbf{p}_1s^2$$

For any word $v \in S$ and any state $(\mathbf{p}_i,\mathbf{p}_{i+1})$ from $M$ by Lemma 11 $\mathbf{p}_iv \succeq_M \mathbf{p}_{i+1}v$. Therefore for any word $v \in S$ the non one-element sequence of states $\mathbf{p}_1v, ..., \mathbf{p}_mv$ forms $t$-cycle of minimal size $m$. It is also true for $v = s^i$ for any

integer $i$.
The states $\mathbf{p}_1, \mathbf{p}_1 s, \mathbf{p}_1 s^2$ are distinct. Let us notice that in aperiodic finite semigroup for some $l$ holds $s^l \neq s^{l+1} = s^{l+2}$. Therefore there exists such maximal integer $k \leq l$ such that $\mathbf{p}_1 s^k \neq \mathbf{p}_1 s^{k+1} = \mathbf{p}_1 s^{k+2}$ and in the $t$-cycle $\mathbf{p}_1 s^k$, $\mathbf{p}_2 s^k = \mathbf{p}_1 s^{k+1}$, $\mathbf{p}_3 s^k = \mathbf{p}_1 s^{k+2}$,..., $\mathbf{p}_m s^k$ holds $\mathbf{p}_1 s^k \neq \mathbf{p}_2 s^k = \mathbf{p}_3 s^k$. So the cardinality of the obtained $t$-cycle is greater than one and less than $m$. Contradiction.

**Corollary 23** *Let $M$ be almost minimal SCC of aperiodic DFA with transition graph $\Gamma$. Then the relation $\succ_M$ is anti-reflexive and any state $\mathbf{p}$ from $\Gamma(M)$ belongs to $\rho_M$ - class of size at least two.*

**Lemma 24** *Let $M$ be almost minimal SCC of aperiodic DFA with transition graph $\Gamma$. Let the state $\mathbf{q}$ be a minimal element of the partial order $\succeq_M$ and suppose that for some word $s$ the state $\mathbf{q}s$ is either a maximal element of the order $\succeq_M$ or $\mathbf{q}s \notin \Gamma(M)$.*
*Then for any state $\mathbf{t}$ such that $\mathbf{t} \succeq_M \mathbf{q}$ holds $\mathbf{t}s = \mathbf{q}s$. The word $s$ unites all predecessors of $\mathbf{q}$.*

Proof. By Lemma 11, $\mathbf{t}s \succ_M \mathbf{q}s$ or $\mathbf{q}s = \mathbf{t}s$. The case $\mathbf{t}s \succ_M \mathbf{q}s$ is excluded because the state $\mathbf{q}s$ is a maximal state. In the case $\mathbf{r} \notin \Gamma(M)$ also $\mathbf{t}s = \mathbf{q}s$ by Proposition 12.
Thus the word $s$ is a common synchronizing word for all states $\mathbf{t}$ such that $\mathbf{t} \succ_M \mathbf{q}$.

## 3 Černy conjuncture

**Lemma 31** *Let $r$ be the number of $\rho_M$ - classes of almost minimal SCC $M$ and let $R$ be a $\rho_M$ - class.*
*Then $|Rs| = 1$ for some word $s \in \Sigma^*$ of length not greater than $(n-r)(n-1)/2$.*

Proof. Suppose $|R| > 1$. Let $Max$ be the set of all maximal and $Min$ be the set of all minimal states from $R$ according to the order $\succ_M$. The sets $Max$ and $Min$ are not empty in view of Lemma 22. $|Max| \cap |Min| = \emptyset$ because the relation $\succ_M$ in virtue of Corollary 23 is anti-reflexive. Without loss of generality, let us assume that $|Max| > |Min|$. Then $|Min| \leq (n-r)/2$. For any word $s$, $Rs$ is a class of congruence $\rho_M$ (Lemma 11). The number of all minimal states of $Rs$ is not greater than in $R$ because the relation $\succeq_M$ is stable (Lemma 11).
For any minimal state $\mathbf{q}s \in Rs$ there exists a word $w$ of length not greater $n-1$ that maps $\mathbf{q}s$ together with all its predecessors in $Max$ (Lemma 24). On this way, the number of minimal states can be reduced to zero by help of at most $|Min|$ words of length $n-1$ or less. The $\rho_M$-class without minimal elements is one-element, $|Min| \leq (n-r)/2$, whence for a word $u$ of length not greater than $(n-1)(n-r)/2$ holds $|Ru| = 1$.

**Theorem 32** *The strongly connected complete transition graph $\Gamma$ of aperiodic automaton with $n$ states has synchronizing word of length less than $(n-1)n/2$.*

Proof. All states of an automaton with strongly connected transition graph $\Gamma$ are sink states. Therefore the automaton is synchronizing (Lemma 21).
There exists at least one almost minimal $SCC$ $M$ in $\Gamma^2$ because the number of $SCC$ is finite and the set of $SCC$ is partially ordered. Suppose $|M| = m$. Let us consider the congruence $\rho_M$ (Lemma 11) and the quotient $\Gamma/\rho_M$.
Any synchronizing word of $\Gamma$ synchronizes also $\Gamma/\rho_M$. Therefore the graph $\Gamma$ has a synchronizing word $uv$ where $u$ is a synchronizing word of $\Gamma/\rho_M$ and $v$ is a synchronizing word of the pre-image $R$ of $\Gamma/\rho_M$ $u$. By Corollary 23, $\rho_M$ is not trivial, therefore $|\Gamma/\rho_M| = r < n$ and we can use induction, assuming $|u| \leq (r-1)r/2$. By Lemma 31, the word $v$ of length not greater than $(n-r)(n-1)/2$ synchronizes $R$. Therefore

$$|uv| \leq r(r-1)/2 + (n-r)(n-1)/2 < n(n-1)/2.$$

Let us go to the general case.

**Theorem 33** *Let $A$ be an aperiodic $DFA$ with $n$ states. Then the existence of sink state in $A$ is equivalent to the existence of synchronizing word of length not greater than $n(n-1)/2$.*

Proof. It is clear, that the existence of synchronizing word implies the existence of sink state.
So let us consider a $DFA$ with at less one sink state. By lemma 21, the automaton is synchronizing. We can assume in view of theorem 32 that $\Gamma$ is not strongly connected.
Let $\Gamma_i$ be an $SCC$ of $\Gamma$ of cardinality $r_i < n$ without sink state $(i = 1, 2, ..., k)$ and let $C$ be $SCC$ of $\Gamma$ of cardinality $r < n$ with sink state. By [10] (Theorem 6.1), there exists a word $s_i$ of length at most $r_i(r_i-1)/2$ such that $\Gamma_i s_i \cap \Gamma_i$ is empty. Some product of such words $s_i$ of length not greater than $\sum r_i(r_i-1)/2$ maps $\Gamma$ on the $SCC$ $C$. By Theorem 32, $C$ has synchronizing word $s$ of length not greater than $r(r-1)/2$. We have $r + \sum r_i = n$. Therefore the word $s_1...s_k s$ synchronizes $\Gamma$ and $|s_1...s_k s| \leq \sum r_i(r_i-1)/2 + r(r-1)/2 \leq (n-1)n/2$.

**Corollary 34** *The Černy conjecture holds for aperiodic $DFA$.*

## Acknowledgments

I am very grateful to the M.V. Volkov for helpful and detailed comments that proved highly useful in improving the presentation and style of the paper.

## References

1. D. S. Ananichev, M.V. Volkov, Collapsing words vs. synchronizing words. Springer, Lect. Notes in Comp. Sci. 2295(2002), 166-174.
2. J. Černy, Poznamka k homogenym eksperimentom s konechnymi automatami, Math.-Fyz. Čas., 14(1964) 208-215.
3. L.Dubuc, Sur le automates circulaires et la conjecture de Černy, RAIRO Inform. Theor. Appl., no 1-3, 32(1998) 21-34.

4. D. Eppstein, Reset sequences for monotonic automata. SIAM J. Comput., **19**(1990) 500-510.
5. P. Frankl, An extremal problem for two families of sets, Eur. J. Comb., 3(1982) 125-127.
6. J. Kari, Synchronizing finite automata on Eulerian digraphs. Springer, Lect. Notes in Comp. Sci., 2136( 2001), 432-438.
7. A.. Kljachko, I.K. Rystsov, M.A. Spivak, An extremely combinatorial problem connected with the bound on the length of a recurrent word in an automata. Kybernetika. 2(1987) 16-25.
8. J.E. Pin, Sur un cas particulier de la conjecture de Černy, Springer, Lect. Notes Comp. Sci., 62(1978) 345-352.
9. I.K. Rystsov, Almost optimal bound on recurrent word length for regular automata. Cybernetics and System An. 31, 5(1995) 669-674.
10. I.K. Rystsov, Reset words for commutative and solvable automata. Theoret. Comput. Sci. 172(1997) 273-279.
11. A. Salomaa, Generation of constants and synchronization of finite automata, J. of Univers. Comput. Sci., 8(2) (2002), 332-347.
12. M.P. Schŭtzenberger, On finite monoids having only trivial subgroups. Inf. control, 8(1965) 190-194.
13. A.N. Trahtman, Algorithms finding the order of local testability of deterministic finite automaton and estimation of the order, Theoret. Comput. Sci., 235(2000), 183-204.