Row monomial matrices and Černy conjecture, short proof

A.N. Trahtman

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

Abstract. A word \( w \) of letters on edges of underlying graph \( \Gamma \) of deterministic finite automaton (DFA) is called synchronizing if \( w \) sends all states of the automaton to a unique state. J. Černy discovered in 1964 a sequence of \( n \)-state complete DFA possessing a minimal synchronizing word of length \((n-1)^2\).

The hypothesis, well known today as the Černy conjecture, claims that \((n-1)^2\) is also precise upper bound on the length of synchronizing word for a complete DFA. The hypothesis was formulated in 1966 by Starke. The problem has motivated great and constantly growing number of investigations and generalizations.

We present the proof of the Černy-Starke conjecture: the deterministic complete \( n \)-state synchronizing automaton has synchronizing word of length at most \((n-1)^2\).

The proof used also the space of row monomial matrices (one unit and rest zeros in every row) and connection between dimension of the space and the length of words on paths of edges in underlying graph of automaton.

Introduction

The long and fascinating history of state machine synchronization and the problems around was reflected in hundreds articles. The problem of synchronization of finite automata is a natural one and various aspects of this problem have been touched in the literature. Different problems of synchronization and achievements one can find in surveys \cite{7,10} and works \cite{17,13,19}.

The synchronizing word limits the propagation of errors for a prefix code. Deterministic finite automaton is a tool that helps to recognized language in a set of DNA strings.

A problem with a long story is the estimation of the minimal length of synchronizing word. J. Černy in 1964 \cite{3} found the infinite sequence of \( n \)-state complete DFA with shortest synchronizing word of length \((n-1)^2\) for an alphabet of size two. Since then, only 27 small automata with synchronizing word of length \((n-1)^2\) of size \( n \leq 6 \) have been added to the single Černy sequence \cite{19}.

The following hypothesis is well known today as Černy’s conjecture:

* Email: avraham.trakhtman@gmail.com
Conjecture 1 The deterministic complete $n$-state synchronizing automaton over alphabet $\Sigma$ has synchronizing word in $\Sigma$ of length at most $(n-1)^2$ [10] (Starke, 1966).

The problem can be reduced to automata with a strongly connected graph [3].

We consider a class of matrices $M_u$ of mapping induced by words $u$ in the alphabet of letters on edges of the underlying graph $\Gamma$. Every edge has length one.

A sequence of distinct non-trivial spaces ordered by inclusion of matrices is used in the study.

Initially found upper bound for the minimal length of synchronizing word was big and has been consistently improved over the years by different authors. The upper bound found by Frankl in 1982 [6] is equal to $(n^3 - n)/6$. The result was reformulated in terms of synchronization in [14] and repeated independently in [8]. The cubic estimation of the bound exists since 1982.

Examples of automata such that the length of the shortest synchronizing word is greater than $(n-1)^2$ are unknown. Moreover, the examples of automata with shortest synchronizing word of length $(n-1)^2$ are infrequent. After the sequence of Černý and the example of Černý, Piricka and Rosenauerova [4] of 1971 for $|\Sigma| = 2$, the next such examples were found by Kari [9] in 2001 for $n = 6$ and $|\Sigma| = 2$ and by Roman [15] for $n = 5$ and $|\Sigma| = 3$ in 2004. The package TESTAS [20] studied all automata with strongly connected underlying graph of restricted size and alphabet from 2 to 4 found five new examples of DFA with shortest synchronizing word of length $(n-1)^2$ for $n \leq 4$.

Don and Zantema present in [5] an ingenious method of designing several new automata, a kind of "hybrids" from existing examples from [3], [4], [20] of size three and four. They proved that for $n \geq 5$ the method does not work. So there are up to isomorphism exactly 15 DFA for $n = 3$ and 12 DFA for $n = 4$ with shortest synchronizing word of length $(n-1)^2$.

The authors of [5] support the hypothesis from [20] that all automata with shortest synchronizing word of length $(n-1)^2$ are known, of course, with essential correction found by themselves for $n = 3, 4$.

We propose below an attempt to prove the Černý conjecture.

Preliminaries

We consider a complete $n$-state DFA with strongly connected underlying graph $\Gamma$ over a fixed finite alphabet $\Sigma$ of labels on edges of $\Gamma$ of an automaton $A$. The trivial cases $n \leq 2$, $|\Sigma| = 1$ and $|A\sigma| = 1$ for $\sigma \in \Sigma$ are excluded.

The restriction on strongly connected graphs is based on [3]. The states of the automaton $A$ are considered also as vertices of the graph $\Gamma$.

If there exists a path in an automaton from the state $p$ to the state $q$ and the edges of the path are consecutively labelled by $\sigma_1, \ldots, \sigma_k$, then for $s = \sigma_1 \ldots \sigma_k \in \Sigma^+$ let us write $q = ps$. 
Let \( P \) be the set of states \( q = px \) for all \( p \) from the subset \( P \) of states and \( x \in \Sigma^+ \). Let \( Ax \) denote the set \( P \) for the set \( P \) of all states of the automaton.

A word \( s \in \Sigma^+ \) is called synchronizing (reset, magic, recurrent, homing) word of an automaton \( A \) with underlying graph \( \Gamma \) if \( |As| = 1 \). The word \( s \) below denotes minimal synchronizing word.

We consider row monomial matrices (one unit and rest of zeros in every row) \( n \times n \)-matrices and edges on underlying graph of length one.

The states of the automaton are enumerated, the state \( q \) has number one.

An automaton (and its underlying graph) possessing a synchronizing word is called synchronizing.

We connect a mapping of the set of states of the automaton made by a word \( u \) of row monomial \( n \times n \)-matrix \( M_u \) such that for an element \( m_{i,j} \in M_u \) takes place

\[
m_{i,j} = \begin{cases} 
1, & q_i = q_j; \\
0, & \text{otherwise}. 
\end{cases}
\]

Any mapping of the set of states of the automaton \( A \) can be presented by some word \( u \) and by a corresponding matrix \( M_u \). For instance,

\[
M_u = \begin{pmatrix}
0 & 0 & 1 & \ldots & 0 \\
1 & 0 & 0 & \ldots & 0 \\
0 & 0 & 0 & \ldots & 1 \\
1 & \ldots & \ldots & \ldots & 0 \\
0 & 1 & 0 & \ldots & 0 \\
1 & 0 & 0 & \ldots & 0
\end{pmatrix}
\]

Let us call the matrix \( M_u \) of the mapping induced by the word \( u \), for brevity, the matrix of word \( u \).

\[
M_u M_a = M_{au}. \tag{2}
\]

The set of nonzero columns of \( M_u \) (set of second indexes of its elements) of \( M_u \) is denoted as \( R(u) \) of size \( |R(u)| \).

For linear algebra terminology and definitions, see [11], [12].

1 Some properties of row monomial matrices

Remark 1 The invertible matrix \( M_a \) does not change the number of units of every column of \( M_u \) in its image of the product \( M_a M_u \).

Every unit in the product \( M_a M_u \) is the product of two units, first unit from nonzero column of \( M_u \) and second unit from a row with one unit of \( M_a \).

Remark 2 The columns of the matrix \( M_u M_a \) are obtained by permutation of columns \( M_u \). Some columns can be merged (units of columns are moved along row to a common column) with \( |R(ua)| < |R(u)| \).

The rows of the matrix \( M_a M_u \) are obtained by permutation of rows of the matrix \( M_u \). Some of these rows may disappear and replaced by another rows of \( M_u \).

Lemma 1 The number of nonzero columns \( |R(b)| \) is equal to the rank of \( M_b \).

\[
|R(ua)| \leq |R(u)| \text{ and } R(au) \subseteq R(u).
\]
For invertible matrix $M_a$ we have $R(au) = R(u)$ and $|R(ua)| = |R(u)|$. Nonzero columns of $M_{aa}$ have units also in $M_a$.

Proof. The matrix $M_b$ has submatrix with nonzero determinant having only one unit in every row and in every nonzero column. Therefore $|R(b)|$ is equal to the rank of $M_b$.

The matrix $M_a$ in the product $M_u M_a$ shifts column of $M_u$ to columns of $M_u M_a$ without changing the column itself by Remark 2 or merging some columns of $M_u$. In view of possible merged columns, $|R(ua)| \leq |R(u)|$.

Some rows of $M_u$ can be replaced in $M_u M_a$ by another row and therefore some rows from $M_a$ may be changed, but zero columns of $M_u$ remain in $M_u M_a$ (Remark 1).

Hence $R(au) \subseteq R(u)$ and $|R(ua)| \leq |R(u)|$.

For invertible matrix $M_a$ we have $R(au) = R(u)$ and $|R(ua)| = |R(u)|$. Nonzero columns of $M_{aa}$ have units also in $M_a$ in view of $R(ua) \subseteq R(u)$.

Corollary 1 All matrices of prefixes of synchronizing matrix $s$ also have at least one unit in nonzero column of $s$.

Corollary 2 The invertible matrix $M_a$ keeps the number of units of any column of $M_u$ in corresponding column of the product $M_u M_a$.

$|R(u)|$ we call also rank of word $u$.

1.1 Dimension of the space of row monomial matrices.

Lemma 2 The set $V$ of all $n \times k$-matrices having precisely one unit in any row and zeroes everywhere else (row monomial) (or $n \times n$-matrices with zeros in fixed $n - k$ columns for $k \leq n$) has at most $n(k - 1) + 1$ linear independent matrices.

Proof. Let us consider distinct $n \times k$-matrices of word with at most only one nonzero cell outside the last nonzero column $k$.

Let us begin from the matrices $V_{i,j}$ with unit in $(i, j)$ cell $(j < k)$ and units in $(m, k)$ cells for all $m$ except $i$. The remaining cells contain zeros. So we have $n - 1$ units in the $k$-th column and only one unit in remaining $k - 1$ columns of the matrix $V_{i,j}$. Let the matrix $K$ have units in the $k$-th column and zeros in the other columns. There are $n(k - 1)$ matrices $V_{i,j}$. Together with $K$ they belong to the set $V$. So we have $n(k - 1) + 1$ matrices. For instance,

$$V_{i,1} = \begin{pmatrix} 1 & 0 & 0 & \ldots & 0 \\ 0 & 0 & 1 & \ldots & 0 \\ 0 & 0 & 1 & \ldots & 0 \\ \ldots & \ldots & \ldots & \ldots & \ldots \\ 0 & 0 & 0 & \ldots & 1 \\ 0 & 0 & 0 & \ldots & 1 \end{pmatrix} \quad V_{3,2} = \begin{pmatrix} 0 & 0 & 0 & \ldots & 1 \\ 0 & 0 & 0 & \ldots & 1 \\ 0 & 0 & 0 & \ldots & 1 \\ 0 & 1 & 0 & \ldots & 0 \\ \ldots & \ldots & \ldots & \ldots & \ldots \end{pmatrix} \quad K = \begin{pmatrix} 0 & 0 & 0 & \ldots & 1 \\ 0 & 0 & 0 & \ldots & 1 \\ 0 & 0 & 0 & \ldots & 1 \\ 0 & 0 & 0 & \ldots & 1 \\ 0 & 0 & 0 & \ldots & 1 \end{pmatrix}$$

The first step is to prove that the matrices $V_{i,j}$ and $K$ generate the space with the set $V$. For arbitrary matrix $T$ of word from $V$ for every $t_{i,j} \neq 0$ and
$j < k$, let us consider the matrices $V_{i,j}$ with unit in the cell $(i,j)$ and the sum of them $Z = \sum V_{i,j}$. So the first $k-1$ columns of $T$ and $Z$ coincide.

Hence in the first $k-1$ columns of the matrix $Z$ there is at most only one unit in any row. Therefore in the cell of $k$-th column of $Z$ one can find only value of $m$ or $m-1$. The value of $m$ appears if there are only zeros in other cells of the considered row. Therefore $\sum V_{i,j} - (m-1)K = T$. Thus every matrix from the set $V$ is a span of $(k-1)n + 1$ matrices from $V$.

It remains now to prove that the set of matrices $V_{i,j}$ and $K$ is a set of linear independent matrices. If one excludes a certain matrix $V_{i,j}$ from the set of these matrices, then it is impossible to obtain a nonzero value in the cell $(i,j)$ and therefore to obtain the matrix $V_{i,j}$. So the set of matrices $V_{i,j}$ is linear independent.

Every non-trivial linear combination of the matrices $V_{i,j}$ equal to a matrix of word has at least one nonzero element in the first $k-1$ columns. Therefore, the matrix $K$ could not be obtained as a linear combination of the matrices $V_{i,j}$. Consequently the set of matrices $V_{i,j}$ and $K$ forms a basis of the set $V$.

**Corollary 3** The set of all row monomial $n \times (n-1)$-matrices of words has at most $(n-1)^2$ linear independent matrices.

The set of row monomial $n \times n$-matrices has at most $n(n-1) + 1$ linear independent matrices.

There are at most $n+1$ linear independent row monomial matrices of words in the set of matrices with 2 nonzero columns. and at most $n$ row monomial linear independent matrices in the set of matrices with one common nonzero column.

**Corollary 4** There exists a sequence of length at most $n(n-1) + 1$ of distinct not empty subspaces ordered by inclusion of matrices for $n$-state automaton.

### 2 Linear independent matrices $M_u$

**Lemma 3** Let the space $W$ be generated by linear independent $n \times n$-matrices $M_u$ of words $u$ with $|R(u)| > 1$ of synchronizing complete strongly connected DFA.

Then some matrix $M_{u,\beta} \notin W$ for generator $M_u$ of $W$ and some letter $\beta$.

**Proof.** Assume the contrary: for every generator $M_u$ of $W$ and every letter $\beta$ the matrix $M_{u,\beta}$ in $W$.

So the space $W$ has the same set of generators. Consequently for every generator $M_u$ and every word $t$ the product $M_{u,t} \in W$.

For every matrix $M_v \in W$ $M_v = \sum \tau_i M_{u,i}$

with generators $M_{u,i}$ in $W$. Then $M_{v,\beta} = (\sum \tau_i M_{u,i}) M_{\beta} = \sum \tau_i M_{u,i} M_{\beta}$ for matrices $M_{u,i} M_{\beta} = M_{u,i,\beta}$ un $W$.

Therefore also $M_{v,\beta} = M_{v,\beta} M_{\beta} = M_{v} \sum \tau_i M_{u,i,\beta}$ belongs to $W$. 

By induction for every word $t$ the matrix $M_{u,t}$ belongs to $W$ and the matrix $M_{v,t} \in W$.

However in synchronizing complete strongly connected DFA for some generator $M_u$ and some $t$ $|R(u,t)| = 1$ outside $W$.

**Corollary 5** Let the sequence of spaces $W_j$ of matrices be ordered by inclusion with increasing $j$. The basis of $W_0$ contains the matrix $M_0$ with $|R(u)| = n$ (n units in one row, say, q).

By Lemma 3 for the space $W_{j+1}$ $W_{j+1} \supset W_j$. The space $W_j$ is can be extended by matrix $M_{u\beta}$ of a letter $\beta$ and some $M_u$ from the basis of $W_j$ by Lemma 3.

Matrices from basis $W_j$ are included in basis $W_{j+1}$ as far as possible.

Hence by induction in the sequence of all matrices of kind $M_{u\beta}$ all matrices are linear independent.

### 3 The sequence of spaces $W_j$ ordered by inclusion

We study matrices from $W_j$. Moreover, $W_{j+1}$ includes basis $W_j$. Matrices of generators of $W_j$ are linear independent (Corollary 5). All generators were obtained by adding a letter from right to generator of the former $W_{j-1}$.

The rank $|R(u)|$ plays some role in the study.

By Corollary 4 there exists a sequence of length at most $n(n - 1) + 1$ of distinct non-empty subspaces ordered by inclusion of row monomial matrices for $n$-state automaton.

### 4 Theorems

**Theorem 1** The length of synchronizing word in complete deterministic finite $n$-state synchronizing automaton with strongly connected underlying graph is not more than $(n - 1)^2$.

Proof. The space $W_0$ has matrix $M_0$ with rank $n$.

By Lemma 3 for some matrix of letter $\beta_0$ the matrix $M_0\beta_0 = M_1$ outside $W_0$. The basis of $W_1$ contains $M_0$ with rank $n$ and the matrix $M_1 = M_0\beta_0$ with $|R(u)| \leq n$ in view of Remark 2. So the set of generators of the space $W_1$ has two generators $M_0$ and $M_1$.

The edge with letter $\beta_0$ in underlying graph of automaton corresponds extension $M_0$ by letter $\beta_0$ into matrix $M_1$. The edges of letters $\beta_k$ form some kind of tree in view of Corollary 5.

By Lemma 3 some matrix $M_{u\beta} \not\in W_j$ for generator $M_u$ of $W_j$ and some letter $\beta$. So we can extend $W_j$ by matrix $M_{u\beta}$ in basis of $W_{j+1}$ $W_{j+1} \supset W_j$.

The fixed set of generators $M_{u\beta}$ of the set of spaces $W_j$ corresponds the set of spaces.

The fixed set of generators of $W_i$ can be extended by help of matrix of letter $\beta_i$, due to lemma 3.
Rank of every next generator may decrease with growth of $i$ according to Remark 2. Some sequences of letters $\beta$ corresponds paths in underlying graph of strongly connected synchronizing automaton.

The word $u$ of matrix $M_i$ with $|R(u)| = 1$ corresponds synchronizing word because the matrix maps all $n$ units of the matrix $M_0$ to one nonzero column.

Therefore we consider also matrices $M_i$ with rank $|R(u)| > 1$, because in the case $|R(u)| = 1$ the word of matrix defines a synchronizing word.

By Corollary 4 there exists a sequence of length at most $n(n-1)+1$ of distinct non empty subspaces $W_i$ ordered by inclusion of matrices for $n$-state automaton.

The sequence of subspaces $W_i$ has a sequence of at most $n(n-1)$ linear independent matrices $M_i = M_{i-1}\beta_{i-1}$. The matrix $M_0$ does not belong to the set of such $M_i$ for $i > 0$.

There are in general $n$ linear independent matrices $M_i$ with $|R(u)| = 1$ (with only one nonzero column) and one matrix $M_0$ with $|R(u)| = n$. Hence the number of linear independent matrices with $n > |R(u)| \geq 2$ is at most $n(n-1)+1-n-1 = n(n-2)$, no more than $n(n-2)$ such matrices correspond the edges of letters $\beta$ for rank from $n-1$ to 2.

So the matrix with $|R(u_i)| = 1$ for $i \leq n(n-2)+1 = (n-1)^2$ is possible only for path of length at most $(n-2)^2$ in underlying graph of arbitrary complete strongly connected DFA and $(n-2)^2$ is upper bound for length of synchronizing word.

**Theorem 2** The length of synchronizing word in complete deterministic finite $n$-state synchronizing automaton is not more than $(n-1)^2$.

Follows from Theorem 1 because the restriction for strongly connected graphs can be omitted due to [3].

5 Conclusion

The minimal length of synchronizing word for arbitrary $n$-state complete deterministic finite automaton is restricted by $(n-1)^2$.

The following conjecture is a matter of interest: all automata with shortest synchronizing word of length $(n-1)^2$ are already known. [20], [5].

Acknowledgments

I would like to express my gratitude to François Gonze, Dominique Perrin, Marie Béal, Akihiro Munemasa, Wit Forys, Benjamin Weiss, Mikhail Volkov, Mikhail Berlinkov and Evgeny Kleiman for fruitful and essential remarks throughout the study.
References

1. R.L. Adler, B. Weiss, *Similarity of automorphisms of the torus*, Memoirs of the Amer. Math. Soc., 98, Providence, RI, 1970.
2. M.-P. Béal, *A note on Černý Conjecture and rational series*, technical report, Inst. Gaspard Monge, Univ. de Marne-la-Vallee, 2003.
3. J. Černý, *Posnámka k homogenným eksperimentom s konečnými automatami*, Math.-Fyz. Čas., 14(1964), 208-215.
4. J. Černý, A. Pitricka, B. Rosenauerova, *On directable automata*. Kybernetika 7(1971), 289-298.
5. H. Don, H. Zantema, *Finding DFAs with maximal shortest synchronizing word length*. LATA 2017, LNCS, v.10168, 249-260. arXiv:1609.06853, 2016
6. P. Frankl, *An extremal problem for two families of sets*. Eur. J. Comb., 3(1982), 125-127.
7. H. Jurgensen, *Synchronization*. Inf. and Comp. 206(2008), 9-10, 1033-1044.
8. A.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.
9. J. Kari, *A counter example to a conjecture concerning synchronizing word in finite automata*, EATCS Bulletin, 73(2001), 146-147.
10. J. Kari, M. V. Volkov, *Černý’s conjecture and the road coloring problem*. Handbook of Automata, 2013.
11. P. Lancaster, *Theory of matrices*, Acad. Press, 1969.
12. A. I. Malcev, *Foundations of linear algebra*. San Francisco, Freeman, 1963. (Nauka, 1970, in Russian.)
13. A. Mateescu and A. Salomaa, *Many-Valued Truth Functions, Černý’s conjecture and road coloring*. Bulletin EATCS, 68 (1999), 134-148.
14. J.E. Pin, *On two combinatorial problems arising from automata theory*. Annals of Discrete Math., 17(1983), 535-548.
15. A. Roman, *Experiments on Synchronizing Automata*. Schedae Informaticae, Versita, Warsaw, 19(2010), 35-51.
16. P. H. Starke, *Eine Bemerkung über homogene Experimente*. Elektronische Informationverarbeitung und Kybernetik, 2(1966), 257-259.
17. A.N. Trahtman, *Synchronizing Road Coloring*, Fifth Ifip Int. Conf., TCS-WCC, 2008, 43-53.
18. A.N. Trahtman, *The algebra of row monomial matrices* [http://arXiv:2022.110.06839] [cs.FL].
19. A.N. Trahtman, *Some new Features of Row Monomial Matrices for the Study of DFA*. Recent Advances in Math. Res. and CS. V 9, 2022, 126-137.
20. A.N. Trahtman, *Notable trends concerning the synchronization of graphs and automata*, CTW06, El. Notes in Discrete Math., 25(2006), 173-175.