Open access • Book Chapter • DOI:10.1007/978-3-319-53733-7_18

# Finding DFAs with Maximal Shortest Synchronizing Word Length — Source link

Henk Don, Hans Zantema, Hans Zantema

**Institutions:** Radboud University Nijmegen, Eindhoven University of Technology

**Published on:** 06 Mar 2017 - Language and Automata Theory and Applications

**Topics:** Synchronizing word

# Finding DFAs with maximal shortest synchronizing word length

Henk Don[2] and Hans Zantema[1,2]

[1] Department of Computer Science, TU Eindhoven, P.O. Box 513,
5600 MB Eindhoven, The Netherlands, email: h.zantema@tue.nl
[2] Radboud University Nijmegen, P.O. Box 9010,
6500 GL Nijmegen, The Netherlands email: h.don@math.ru.nl

**Abstract.** It was conjectured by Černý in 1964 that a synchronizing DFA on $n$ states always has a shortest synchronizing word of length at most $(n-1)^2$, and he gave a sequence of DFAs for which this bound is reached. In 2006 Trahtman conjectured that apart from Černý's sequence only 8 DFAs exist attaining the bound. He gave an investigation of all DFAs up to certain size for which the bound is reached, and which do not contain other synchronizing DFAs. Here we extend this analysis in two ways: we drop this latter condition, and we drop limits on alphabet size. For $n \leq 4$ we do the full analysis yielding 19 new DFAs with smallest synchronizing word length $(n-1)^2$, refuting Trahtman's conjecture. Several of these new DFAs admit more than one synchronizing word of length $(n-1)^2$, and even the synchronizing state is not unique. All these new DFAs are extensions of DFAs that were known before. For $n \geq 5$ we prove that none of the DFAs in Trahtman's analysis can be extended similarly. In particular, as a main result we prove that the Černý examples $C_n$ do not admit non-trivial extensions keeping the same smallest synchronizing word length $(n-1)^2$.

## 1 Introduction

A *deterministic finite automaton (DFA)* over a finite alphabet $\Sigma$ is called *synchronizing* if it admits a *synchronizing word*. Here a word $w \in \Sigma^*$ is called *synchronizing* (or directed, or reset) if starting in any state $q$, after processing $w$ one always ends in one particular state $q_s$. So processing $w$ acts as a reset button: no matter in which state the system is, it always moves to the particular state $q_s$. Now Černý's conjecture ( [3]) states:

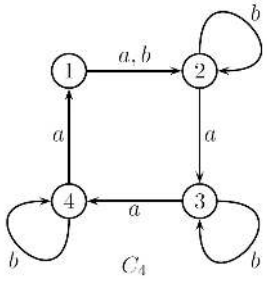
> Every synchronizing DFA on $n$ states admits a synchronizing word of length $\leq (n-1)^2$.

Surprisingly, despite of extensive effort this conjecture is still open, and even the best known upper bound is still cubic in $n$. Černý himself ( [3]) provided an upper bound of $2^n - n - 1$ for the length of the shortest synchronizing word. A substantial improvement was given by Starke [14], who was the first to give a polynomial upper bound, namely $1 + \frac{1}{2}n(n-1)(n-2)$. The best known upper

bound is $\frac{1}{6}(n^3-n)$, established by Pin in 1983 [11]. He reduced proving this upper bound to a purely combinatorial problem which was then solved by Frankl [8]. Since then for more than 30 years no progress for the general case has been made.

The conjecture has been proved for some particular classes of automata , such as circular automata, aperiodic automata and one-cluster automata with prime length cycle. For these results and some more partial answers, see [1, 2, 5–7, 10, 12, 15, 17, 18]. For a survey on synchronizing automata and Černý's conjecture, we refer to [19].

In [3] Černý already gave DFAs for which the bound of the conjecture is attained: for $n \geq 2$ the DFA $C_n$ is defined to consist of $n$ states $1, 2, \ldots, n$, and two symbols $a, b$, acting by $\delta(i, a) = i + 1$ for $i = 1, \ldots, n-1$, $\delta(n, a) = 1$, and $\delta(i, b) = i$ for $i = 2, \ldots, n$, $\delta(1, b) = 2$. For $n = 4$ this is depicted on the right.

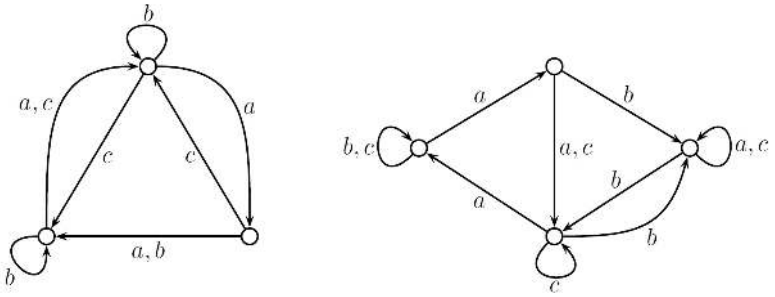



For $C_n$ the string $w = b(a^{n-1}b)^{n-2}$ of length $|w| = (n-1)^2$ satisfies $qw = 2$ for all $q \in Q$, so is synchronizing. No shorter synchronizing word exists for $C_n$ as is shown in [3], showing that the bound in Černý's conjecture is sharp.

The topic of this paper is to investigate all DFAs for which the bound is reached; these DFAs are called *critical*. A DFA for which the bound is exceeded is called *super-critical*, so Černý's conjecture states that no super-critical DFA exists. To exclude infinitely many trivial extensions, we only consider *basic* DFAs: no two distinct symbols act in the same way in the automaton, and no symbol acts as the identity. Obviously, adding the identity or copies of existing symbols has no influence on synchronization.

An extensive investigation was already done by Trahtman in [16]: by computer support and clever algorithms all critical DFAs on $n$ states and $q$ symbols were investigated for $3 \leq n \leq 7$ and $q \leq 4$, and for $n = 8, 9, 10$ and $q = 2$. Here a minimality requirement was added: examples were excluded if criticality may be kept after removing one symbol. Then up to isomorphism there are exactly 8 of them, apart from the basic Černý examples: 3 with 3 states, 3 with 4, one with 5 and one with 6. So apart from the basic Černý examples only 8 other critical DFAs were known. It was conjectured in [16] that no more exist, which is refuted in this paper by finding several more not satisfying the minimality condition, all being extensions of known examples. As one main result we prove that up to isomorphism for $n = 3$ there are exactly 15 basic critical DFAs and for $n = 4$ there are exactly 12 basic critical DFAs, 19 more than the four for $n = 3$ and the four for $n = 4$ that were known before.

Two typical examples are depicted as follows.




The left one restricted to $a, b$ is exactly $C_3$, while restricted to $a, c$ it is exactly a DFA found in [16] that we call T3-1 in Section 3. So this example is a kind of union of $C_3$ and T3-1. It has four distinct synchronizing words of the minimal length 4 described by $(b + c)aa(b + c)$, having two distinct synchronizing states.

The right one restricted to $a, b$ is the example found in [4] that we call CPR in Section 3. However, the extra non-trivial symbol $c$ does not occur in any known critical DFA on four states. It has eight distinct synchronizing words of the minimal length 9 described by $(b + c)aa(b + c)abaa(b + c)$, again having two distinct synchronizing states.

In the partial order on the 15 critical basic DFAs on three states, the four given in [16] are the minimal ones, but there is only one maximal one, being an upper bound of all. Here *maximal* means that that it does not admit an extension that is still basic and critical. In the partial order on the 12 critical basic DFAs on four states, the four given in [16] are the minimal ones, and exactly three are maximal. Two of the maximal examples are also minimal; the other is an upper bound of the two remaining minimal ones.

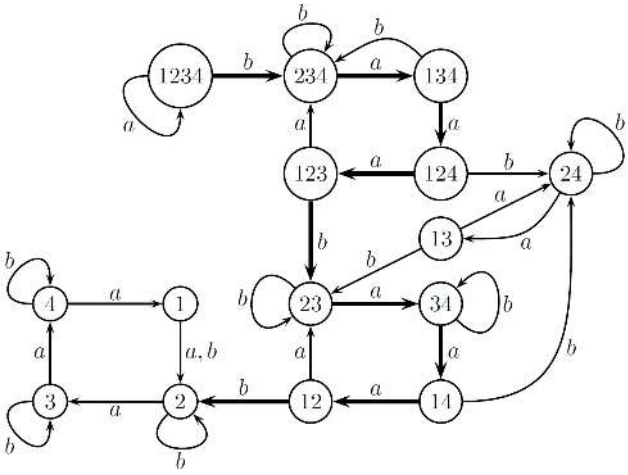
For $n \geq 5$, we wonder whether the minimal critical DFAs in Trahtman's analysis admit critical extensions just as for $n \leq 4$. The answer is negative. Apart from $C_n$ these include only two minimal critical DFAs: one with 5 and one with 6 states, for which a simple computer search applies. For $C_n$ this boils down to the main theorem stating that when adding an extra symbol to $C_n$ not acting as the identity or as one of the existing symbols, always a strictly shorter synchronizing word can be obtained. The theorem is proved by a case analysis in how this extra symbol acts on the states.

This paper is organized as follows. In Section 2 we give some preliminaries. In Section 3 we consider DFAs of at most six states. First we give a self-contained analysis of all critical DFAs on $\leq 4$ states. Next, for the known critical DFAs on five and six states we show that they do not admit critical extensions. The most substantial part is Section 4, where we prove our property for $C_n$ for arbitrary $n$: $C_n$ has no critical extension for $n \geq 5$. An extensive case analysis on how an extra non-trivial symbol $c$ acts on the $n$ states shows that this always yields a shorter synchronizing word. We conclude in Section 5.

## 2 Preliminaries

A *deterministic finite automaton (DFA)* over a finite alphabet $\Sigma$ consists of a finite set $Q$ of states and a map $\delta : Q \times \Sigma \to Q$.[3] A DFA is called *basic* if the mappings $q \mapsto \delta(a, q)$ are distinct for all $a \in \Sigma$, and are not the identity. For $w \in \Sigma^*$ and $q \in Q$ define $qw$ inductively by $q\epsilon = q$ and $qwa = \delta(qw, a)$ for $a \in \Sigma$. So $qw$ is the state where one ends when starting in $q$ and applying $\delta$-steps for the symbols in $w$ consecutively, and $qa$ is a short hand notation for $\delta(q, a)$. A word $w \in \Sigma^*$ is called *synchronizing* if a state $q_s \in Q$ exists such that $qw = q_s$ for all $q \in Q$. Stated in words: starting in any state $q$, after processing $w$ one always ends in state $q_s$. Obviously, if $w$ is a synchronizing word then so is $wu$ for any word $u$. A DFA on $n$ states is *critical* if its shortest synchronizing word has length $(n-1)^2$; it is *super-critical* if its shortest synchronizing word has length $> (n-1)^2$. A critical DFA is *minimal* if it is not the extension of another critical DFA by one or more extra symbols; it is *maximal* if it does not admit a basic critical extension.

The basic tool to analyze synchronization is by exploiting the *power set automaton*. For any DFA $(Q, \Sigma, \delta)$ its power set automaton is the DFA $(2^Q, \Sigma, \delta')$ where $\delta' : 2^Q \times \Sigma \to 2^Q$ is defined by $\delta'(V, a) = \{q \in Q \mid \exists p \in V : \delta(p, a) = q\}$. For any $V \subseteq Q, w \in \Sigma^*$ we define $Vw$ as above, using $\delta'$ instead of $\delta$. From this definition one easily proves that $Vw = \{qw \mid q \in V\}$ for any $V \subseteq Q, w \in \Sigma^*$. A set of the shape $\{q\}$ for $q \in Q$ is called a *singleton*. So a word $w$ is synchronizing if and only if $Qw$ is a singleton. Hence a DFA is synchronizing if and only if its power set automaton admits a path from $Q$ to a singleton, and the shortest length of such a path corresponds to the shortest length of a synchronizing word.



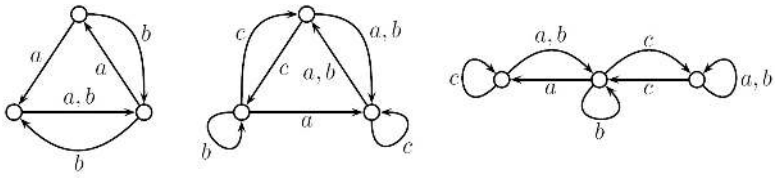

The power set automaton of $C_4$ is depicted on the right, in which indeed the unique shortest path from $Q$ to a singleton (indicated by fat arrows from 1234 to 2) has length 9.

[3] For synchronization the initial state and the set of final states in the standard definition may be ignored.

## 3 Small DFAs

### 3.1 Three states

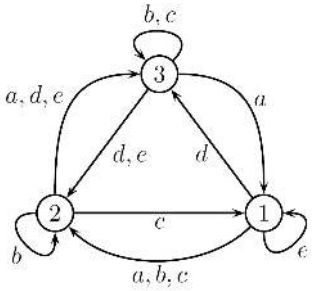
First we give the minimal critical DFAs as presented in [16] on three states, apart from $C_3$:




T3-1 T3-2 T3-3

We call them T3-1, T3-2 and T3-3, as they were found by Trahtman. They all have a unique synchronizing word of length 4, being *baab*, *acba*, *bacb*, respectively.

They can be combined to a single DFA A3 on five symbols $a, b, c, d, e$, depicted as follows.



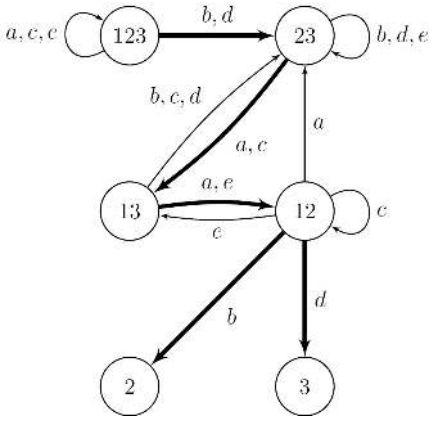

A3

Observe that A3 restricted to $a, b$ coincides with $C_3$, A3 restricted to $a, d$ coincides with T3-1, A3 restricted to $c, d, e$ coincides with T3-2 and A3 restricted to $b, c, e$ coincides with T3-3, so exactly the four minimal critical automata on three states from [16]. On the other hand, as all minimal basic critical DFAs on three states are contained in A3, A3 is the only maximal basic critical DFA on three states. It admits 16 synchronizing words of length 4, expressed by the regular expression $(b + d)(a + c)(a + e)(b + d)$, where state 2 is the synchronizing state if the word ends in $b$ and state 3 if the word ends in $d$. This follows from the analysis of the power set automaton of A3 as depicted below (we stopped when a singleton was reached).

Here the shortest paths from 123 to a singleton are indicated by fat arrows.

The relationship between A3 and critical DFAs is given in the following theorem.

**Theorem 1.** *No super-critical DFAs on three states exist, and a basic DFA on three states is critical if and only if up to isomorphism it is one of the 15 automata that can be obtained from A3 by removing zero or more symbols and keeping at least one of the sets* $\{a, b\}, \{a, d\}, \{b, c, e\}, \{c, d, e\}$ *of symbols.*



*Proof.* Let 1,2,3 be the three states. The automaton has a shortest synchronizing word of length $\geq 4$ if and only if the shortest path from $\{1,2,3\}$ to a singleton in the power set automaton has length $\geq 4$. There is a step from $\{1,2,3\}$ to a smaller set. Since the length of the shortest path is $\geq 4$, this smaller set is not a singleton, so it is a pair; without loss of generality we may assume this is $\{2,3\}$. Let $b$ be the first symbol of a shortest synchronizing word, so $\{1,2,3\} \xrightarrow{b} \{2,3\}$. Since the shortest path from $\{2,3\}$ to a singleton consists of at least three steps, it meets the other two pairs and consists of exactly three steps, yielding shortest synchronizing word length 4. May be after swapping 2 and 3 we may assume this shortest path is $\{1,2,3\} \xrightarrow{b} \{2,3\} \to \{1,3\} \to \{1,2\} \to$ singleton. As it is the shortest path, we conclude that for every symbol $a$ we have

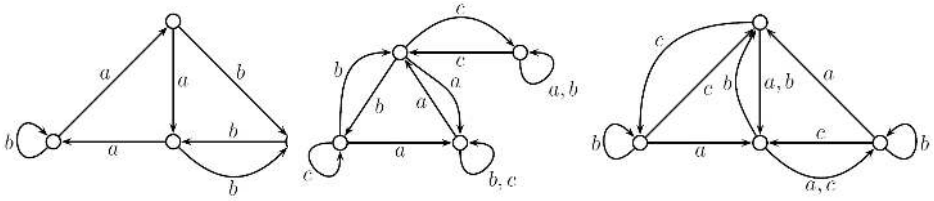
- either $\{1,2,3\} \xrightarrow{a} \{1,2,3\}$ or $\{1,2,3\} \xrightarrow{a} \{2,3\}$,
- either $\{2,3\} \xrightarrow{a} \{2,3\}$ or $\{2,3\} \xrightarrow{a} \{1,3\}$, and
- not $\{1,3\} \xrightarrow{a}$ singleton.

A small program investigates that among the $3^3 = 27$ possible symbol actions in a DFA on three states exactly 6 satisfy these properties: exactly the symbols $a, b, c, d, e$ in A3 and the identity. So for all DFAs being a sub-automaton of A3 it holds that if it is synchronizing, then the shortest synchronizing word length is 4. Restricting A3 to either $\{a, b\}$, $\{a, d\}$, $\{b, c, e\}$ or $\{c, d, e\}$ yields one of the known synchronizing DFAs, so every extension is synchronizing too. Conversely, it is easily checked that all of these restrictions are minimal: all symbols are required for synchronization. This concludes the proof. □

As a consequence of Theorem 1 apart from the four minimal critical DFAs that were known on three states, we obtain 11 more that are not minimal.

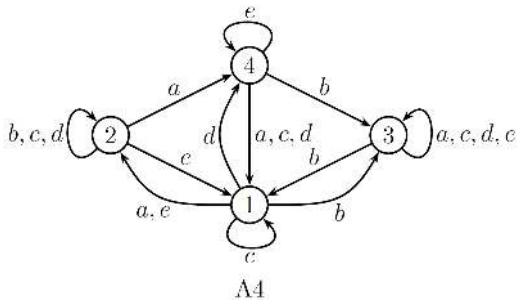
### 3.2 Four states

First we give the minimal critical DFAs as presented in [16] on four states, apart from $C_4$. The first one is CPR, found by Černý, Piricka and Rosenauerova, [4], and has unique synchronizing word of length 9, being *baababaab*. The next two we call T4-1 and T4-2, as they were found by Trahtman. The DFA T4-1 has a unique synchronizing word of length 9, being *abcacabca*; for T4-2 there are 4 synchronizing words of length 9 represented by $acb(a+c)a(a+b)cba$.



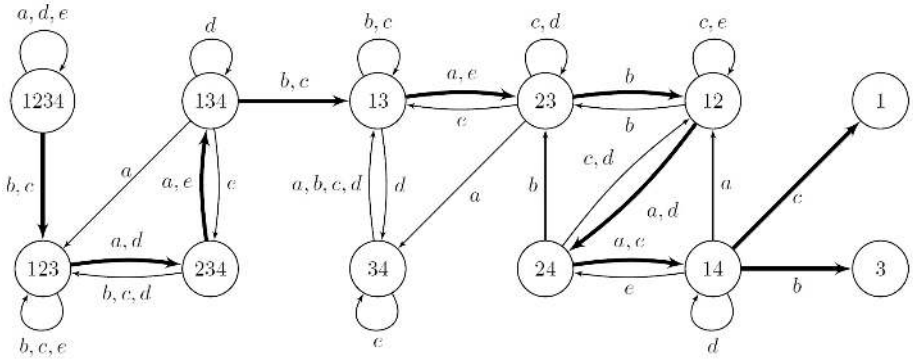

CPR T4-1 T4-2

In order to investigate all critical DFAs with four states, we introduce the DFA A4 on five symbols $a, b, c, d, e$, depicted as follows.



A4

Observe that A4 restricted to $a, b$ coincides with CPR and A4 restricted to $b, d, e$ coincides with T4-1, so together with $C_4$ and T4-2 exactly the four automata with four states from [16], being the minimal ones. On the other hand, $C_4$, T4-2 and A4 are the only maximal basic critical DFAs on four states. We will prove this in Theorem 2. The DFA A4 admits 256 synchronizing words of length 9, expressed by the regular expression $(b + c)(a + d)(a + e)(b + c)(a + e)b(a + d)(a + e)(b + c)$, where the synchronizing state is 1 or 3, depending on the last symbol. This follows from the analysis of the power set automaton of A4 that looks as follows:




Here the shortest paths from 1234 to a singleton are indicated by fat arrows.

The relationship between A4 and critical DFAs is given in the following theorem.

**Theorem 2.** *No super-critical DFAs on four states exist, and a basic DFA on four states is critical if and only if up to isomorphism it is $C_4$, T4-2, or one of the 10 automata that can be obtained from A4 by removing zero or more symbols and keeping at least one of the sets $\{a, b\}, \{b, d, e\}$ of symbols.*

*Proof.* Let 1,2,3,4 be the four states. We have to prove that the shortest path in the power set automaton from $\{1, 2, 3, 4\}$ to a singleton never has length $> 9$ (this would be super-critical), and that length 9 only occurs in the cases indicated by the theorem. So assume this length is $\geq 9$. Since there is a step from $\{1, 2, 3, 4\}$ to a smaller set, and since the length is $\geq 9$, it is not to a pair since there are only 6 distinct pairs. So after one step only one element is removed from $\{1, 2, 3, 4\}$, say, 4. By the same argument also the next step in a shortest path to a singleton is not to a pair; by possibly renaming we may assume it is to $\{2, 3, 4\}$, so a shortest path is of the shape $\{1, 2, 3, 4\} \stackrel{a_1}{\to} \{1, 2, 3\} \stackrel{a_2}{\to} \{2, 3, 4\} \to^{\geq 7}$ singleton.

Note that every shortest path can be renamed to this form, but not necessarily all simultaneously. So we assume at least one shortest path to have this form.

The approach is to generate all solutions by a computer program. In order to reduce the search space, first we prove that for every symbol $a$ we have

1. if $\{1,2,3,4\} \xrightarrow{a} V$ then $\{1,2,3\} \subseteq V$;
2. if $\{1,2,3\} \xrightarrow{a} V$ then either $V = \{1,2,3\}$ or $V = \{2,3,4\}$;
3. if $\{2,3,4\} \xrightarrow{a} V$ then $|V| = 3$.

We start by property 3. Assume that a symbol $a$ exists such that $\{2,3,4\} \xrightarrow{a} \{p_1, q_1\}$. As the DFA is synchronizing, there is a path from $\{p_1, q_1\}$ to a singleton in the power set automaton. Take a shortest such path. As the shortest path from $\{2,3,4\}$ to a singleton consists of at least 7 steps, the shortest path from $\{p_1, q_1\}$ to a singleton consists of at least 6 steps. As there are only 6 distinct unordered pairs, it can not be longer than 6 steps, and this shortest path is of the shape

$$\{p_1, q_1\} \to \{p_2, q_2\} \to \{p_3, q_3\} \to \{p_4, q_4\} \to \{p_5, q_5\} \to \{p_6, q_6\} \to \text{singleton},$$

in which $\{p_i, q_i\}$ are the six distinct unordered pairs, together with the starting part $\{1,2,3,4\} \xrightarrow{a_1} \{1,2,3\} \xrightarrow{a_2} \{2,3,4\} \xrightarrow{a} \{p_1, q_1\}$ yielding a shortest synchronizing sequence of length 9. From this pattern we will derive a contradiction. In principle this could be done by hand using a lot of case analysis. In order to avoid this, and as in circumstances like these computers may be more reliable than humans, we chose to do this by computer support. We built a formula on eight unknown functions $a_i : \{1,2,3,4\} \to \{1,2,3,4\}$ for $i = 1$ to 8, and 12 values $p_i, q_i \in \{1,2,3,4\}$ for $i = 1$ to 6, stating that

- $a_1(\{1,2,3,4\}) = \{1,2,3\}$, $a_2(\{1,2,3\}) = \{2,3,4\}$, $a_3(\{2,3,4\}) = \{p_1, q_1\}$,
- $a_{i+3}(\{p_i, q_i\}) = \{p_{i+1}, q_{i+1}\}$ for $i = 1,2,3,4,5$,
- $p_i \neq q_i$ for $i = 1,2,3,4,5,6$,
- $\{p_i, q_i\} \neq \{p_j, q_j\}$ for $1 \leq i < j \leq 6$,
- $a_i(p_j) \neq a_i(q_j)$ for all $i = 1, \ldots, 8$, $j = 1,2,3,4,5$, and
- $a_k(\{p_i, q_i\}) \neq \{p_j, q_j\}$ for $i + 1 < j \leq 6$ and $k = 1, \ldots, 8$.

The last two requirements may be stated since otherwise a shorter path to a singleton can be obtained. Next we applied the SMT solver Yices ( [20]) to this formula, that stated that this formula is unsatisfiable in a fraction of a second. So this yields the required contradiction, proving property 3. To check that we generated the correct formulas, we also applied the SMT solver on variants of the formula in which minor parts of the formula were removed, and the obtained satisfying assignments yielded solutions of the modified problem that could be checked by hand.

In fact we proved that in the power set automaton no pair can be reached from $\{1,2,3,4\}$ in less than four steps.

In order to prove properties 1 and 2, observe that from $a_1(\{1,2,3,4\}) = \{1,2,3\}$ we conclude that $x \neq y \in \{1,2,3,4\}$ exist such that $a_1(x) = a_1(y)$. Since $a_1(\{2,3,4\})$ consists of three elements by property 3, we have $\{x, y\} \not\subseteq \{2,3,4\}$,

so $1 \in \{x, y\}$. If $a_1(\{1,2,3\})$ consists a two elements this gives rise to a shorter synchronizing sequence, so $\{x, y\} \not\subseteq \{1,2,3\}$, hence $4 \in \{x, y\}$. So $a_1(1) = a_1(4)$.

For property 1 assume that $\{1,2,3\} \not\subseteq V$ for $V = a(\{1,2,3,4\})$. Since $|V| = 3$, the set $V$ is one of the three sets $\{2,3,4\}, \{1,2,4\}, \{1,3,4\}$. If it is $\{2,3,4\}$, we have a shorter synchronizing sequence; otherwise $\{1,4\} \subseteq V$, by which $|a_1(V)| = 2$, and this pair $a_1(V)$ can be reached in two steps $a, a_1$ from $\{1,2,3,4\}$. Both cases yield a contradiction, proving property 1.

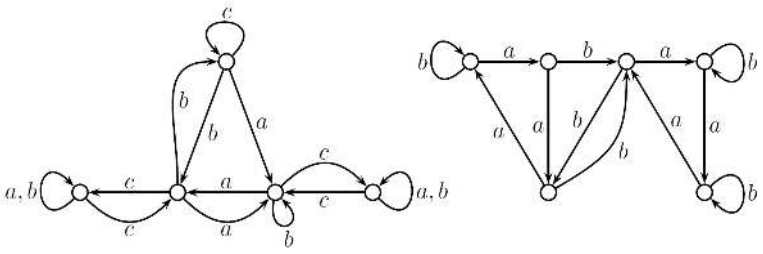
The proof of property 2 is similar: if $V = a(\{1,2,3\})$ is not $\{1,2,3\}$ or $\{2,3,4\}$, then by $|V| = 3$ we have $V$ is $\{1,2,4\}$ or $\{1,3,4\}$, so $|a_1(V)| = 2$ by $a_1(1) = a_1(4)$, and this pair $a_1(V)$ can be reached in three steps $a_1, a, a_1$ from $\{1,2,3,4\}$, contradicting property 3 stating that no shortest path from $\{1,2,3,4\}$ to a singleton reaches a pair in three steps.

It turns out that among the 256 functions from $\{1,2,3,4\}$ to itself exactly 18 satisfy properties 1, 2 and 3. This includes the identity that has to be excluded since that is not allowed in a basic DFA, leaving 17 functions. So we may restrict to basic DFAs on four states such that all symbols act as one of these 17 functions. There are $2^{17} = 131072$ of them. A simple computer program generates all these, and computes for all of them the power automaton and checks whether it admits a path from $\{1,2,3,4\}$ to a singleton, and if so, gives the length of the shortest such path. If this length is $\geq 9$, the DFA is reported. This full computation is executed in less than 10 seconds. As expected, no shortest path length longer than 9 is obtained, proving that no super-critical DFA on 4 states exists. Exactly 24 basic DFAs are obtained with shortest path length 9. These 24 automata exactly coincide with the 12 automata indicated in the theorem; each occurring twice up to swapping 2 and 3. □

As a consequence of Theorem 2 apart from the four minimal critical DFAs that were known on four states, we obtain 8 more that are not minimal.

### 3.3 Five and six states

In Section 3 we saw that for $n = 3, 4$ a critical DFA may have a basic critical extension. We now claim that for $n \geq 5$ this does not occur any more for the known critical DFAs. In Section 4 we will prove that this holds for $C_n$ for all $n \geq 5$. By Trahtman's investigation the only two more critical DFAs to consider are one on five states from Roman [13] and one on six states from Kari [9], depicted as follows.



For Roman's DFA the shortest synchronizing word $abcacacbcaacabca$ is unique; for Kari's DFA there are two shortest synchronizing words, described by $baababababaabbaba(baab + abaa)babaab$.

For $n = 5, 6$ we wrote a program that takes a DFA on $n$ states and computes for all $n^n$ ways to add a fresh symbol, the shortest path length in the power set automaton from the full set to a singleton. For both candidates it turns out that the only extensions keeping this shortest path length to be $(n-1)^2$ is by adding either a copy of one of the existing symbols, or a symbol that acts as the identity. This proves our claim.

To check our results, we also applied this approach to $n = 3, 4$: all 19 new critical DFAs from Theorem 1 and Theorem 2 were obtained as extensions of the earlier known DFAs.

## 4 Extending $C_n$

In this section we show that for all $n \geq 5$ the DFA $C_n$ is maximal: it cannot be extended to a basic critical DFA. The main result of this section is the following:

**Theorem 3.** *Let $n \geq 5$ and let $C_n^c$ be a basic extension of $C_n$ by a symbol $c$. Then $C_n^c$ admits a synchronizing word of length strictly less than $(n-1)^2$.*

Recall that *basic* means that $c$ is not equal to $a$ or $b$ and that $c$ is not the identity function on $Q$. This section is organized as follows: first we collect some properties of $C_n$ and its unique shortest synchronizing word. Then we consider the cases $|Qc| = n$, $|Qc| = n-1$ and $|Qc| \leq n-2$ separately.

### 4.1 Properties of $C_n$

Recall that $C_n$ is defined by $n$ states $1, 2, \ldots, n$, and two symbols $a, b$, acting by $qa = q+1$ for $q = 1, \ldots, n-1$, $na = 1$, and $qb = q$ for $q = 2, \ldots, n$, $1b = 2$. It is well known that $w_n = b(a^{n-1}b)^{n-2}$ of length $|w_n| = (n-1)^2$ is its shortest synchronizing word. It is synchronizing since

$$Qb = \{2, 3, \ldots, n\} \tag{1}$$

$$\{2, 3, \ldots, k\}\, a^{n-1}b = \{2, 3, \ldots, k-1\}, \quad 3 \leq k \leq n. \tag{2}$$

The first part of this word defines the path

$$Q \xrightarrow{b} Q \setminus \{1\} \xrightarrow{a} Q \setminus \{2\} \xrightarrow{a} \ldots \xrightarrow{a} Q \setminus \{n\}. \tag{3}$$

We now extend the alphabet of the automaton by a non-trivial new symbol $c$. Non-trivial means that the transitions defined by $c$ are not all equal to the transitions of $a$ or the transitions of $b$ and furthermore that $c$ is not the identity function. We will distinguish three cases:

1. $|Qc| = n$, i.e. $c$ is a permutation.

2. $|Qc| = n - 1$, i.e. $c$ has deficiency 1.
3. $|Qc| \leq n - 2$, i.e. $c$ has deficiency 2.

We will show that in all these cases a shorter synchronizing word exists. The general pattern in the arguments is as follows. The shortest synchronizing word $w_n$ corresponds to a path from $Q$ to a singleton in the power automaton of $C_n$. Take two sets $S, S' \subseteq Q$ on this path which are visited in this order. Let $d$ be the distance from $S$ to $S'$, i.e.

$$d := \min\left\{|w| : Sw = S', w \in \{a, b\}^\star\right\}.$$

Now construct a word $w \in \{a, b, c\}^\star$ in the automaton $C_n^c$ for which $Sw = S'$ and $|w| < d$. Then $C_n^c$ admits a synchronizing word of length at most $|w_n| - d + |w| < (n-1)^2$.

### 4.2 Construction of a shorter synchronizing word

If $c$ defines a permutation on $Q$, we may assume that $c$ satisfies:

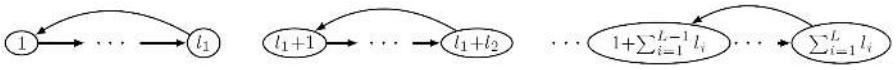
$$qc \leq q + 1 \text{ for all } q \in Q. \tag{4}$$

Indeed, if $qc = q + k$ for some $q \in Q$ and $k \geq 2$, then $(Q \setminus \{q\})c = Q \setminus \{q + k\}$, which in view of (3) would imply existence of a synchronizing word shorter than $(n-1)^2$. The following lemma describes the structure of $c$.

**Lemma 1.** *If $|Q| = n \geq 1$ and $c$ is a permutation on $Q$ satisfying (4), then there exist numbers $L$ (number of $c$-loops) and $1 \leq l_1, \ldots, l_L \leq n$ (lengths of $c$-loops) with $\sum_{i=1}^{L} l_i = n$ such that*

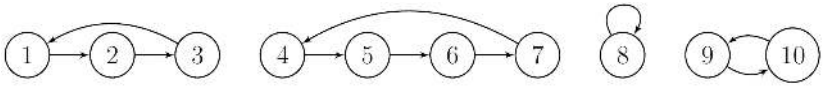
$$qc = \begin{cases} q - l_i + 1 & \text{if } q = l_1 + \ldots + l_i \text{ for some } 1 \leq i \leq L \\ q + 1 & \text{otherwise} \end{cases} \tag{5}$$

An illustration of the statement is given below.




*Proof.* We give a proof by induction. For $n = 1$, $1 \xrightarrow{c} 1$, so $L = 1$ and $l_1 = 1$. Now suppose the statement is true for all $n \leq N$ and consider the case $|Q| = N + 1$. If $1 \xrightarrow{c} 1$, then $c$ defines a permutation on $Q \setminus \{1\}$. Applying the induction hypothesis on $Q \setminus \{1\}$ gives the result. If $1 \xrightarrow{c} 2 \xrightarrow{c} \ldots \xrightarrow{c} k$ for some $k \geq 2$, then either $kc = k + 1$ or $kc = 1$. In both cases there is a number $l_1 \geq 1$ such that $1 \xrightarrow{c} \ldots \xrightarrow{c} l_1 \xrightarrow{c} 1$. Apply the induction hypothesis on the remaining $n - l_1$ states. □

Note that $L = 1$ and $L = n$ are the trivial cases, because then $c = a$ or $c$ is the identity. Before we give a general argument, we first give an example.



*Example 1.* Consider the automaton $C_{10}^c = \{Q, \Sigma, \delta\}$ with $Q = \{1, \ldots, 10\}$ and $\Sigma = \{a, b, c\}$. The actions of the symbols $a$ and $b$ are from the definition of $C_n$ and $c$ is the permutation shown above. Here we have four loops ($L = 4$) with lengths $l_1 = 3, l_2 = 4, l_3 = 1$ and $l_4 = 2$. We will show how to use the $c$-loop of length four to create a shorter synchronizing word. Consider the set $S = \{2, \ldots, 9\}$. We start by a greedy approach to reach a set of size 7:

$$Sa^3b = (\{1, 2\} \cup \{5, \ldots, 10\})\, b = \{2\} \cup \{5, \ldots, 10\}\,.$$

As a next step, we shift everything by using the symbol $a$ until the isolated state $\{2\}$ ends up in the $c$-loop of length four:

$$(\{2\} \cup \{5, \ldots, 10\})\, a^3 = \{1, 2, 3\} \cup \{5\} \cup \{8, 9, 10\}$$

Since $\{1, 2, 3\}$ and $\{8, 9, 10\}$ are (unions of) full $c$-loops, they are invariant under $c$. Therefore, we can move the isolated state $\{5\}$ to the desired position:

$$(\{1, 2, 3\} \cup \{5\} \cup \{8, 9, 10\})\, c^3 = \{1, 2, 3, 4\} \cup \{8, 9, 10\}$$

Finally, we shift again by a power of $a$ and apply $b$ to get rid of one more state:

$$(\{1, 2, 3, 4\} \cup \{8, 9, 10\})\, a^3b = \{1, \ldots, 7\}\, b = \{2, \ldots, 7\} := S'.$$

We conclude that the word $w = a^3ba^3c^3a^3b$ has the property that $Sw = S'$. In $C_{10}$ both $S$ and $S'$ are on the shortest path from $Q$ to $\{2\}$ and by (2) the distance between them is equal to $2n = 20$. The word $w$ has length $|w| = 14$, so in $C_{10}^c$ there exists a synchronizing word of length at most $(10 - 1)^2 - 6 = 75$. Note that there might be even shorter synchronizing words, but for our main goal it is sufficient to have some synchronizing word shorter than 81.

The idea of this example works in more generality if there is a $c$-loop of length at least 3, as is proved in the next lemma. If the longest loop has length 2, then basically we can do the same thing, but we need at least three $c$-loops to isolate a state.

**Lemma 2.** *Let $n \geq 5$ and let $C_n^c$ be an extension of the automaton $C_n$ by a symbol $c$ as given in Lemma 1. If $2 \leq L \leq n-1$, then $C_n^c$ admits a synchronizing word of length strictly less than $(n-1)^2$.*

*Proof.* We distinguish the following three cases:

- $L \geq 2$ and $l_k \geq 3$ for some $k$.
- $L \geq 3$ and $l_k = 2$ for some $k \leq L - 1$.
- $L \geq 3$ and $l_L = 2$.

Note that for all $n \geq 5$ and all possible non-trivial choices of $c$, the extended automaton $C_n^+$ satisfies at least one of these cases.

*Case 1: $L \geq 2$ and $l_k \geq 3$ for some $k$.* Take $k$ such that $l_k \geq 3$ and write $\Lambda^- = \sum_{i=1}^{k-1} l_i$, $\Lambda^+ = \sum_{i=k+1}^{L} l_i$, for the sum of the loop lengths before the $k$th loop and after the $k$th loop respectively. These sums can be zero if $k = 1$ or $k = L$. Define $\Lambda = \Lambda^- + \Lambda^+ = n - l_k \leq n - 3$. Since $L \geq 2$, we have $\Lambda \geq 1$. Take

$$S = \{2, 3, \ldots, n - l_k + 3\}, \quad S' = \{2, 3, \ldots, n - l_k + 1\}.$$

and define the word

$$w = a^{l_k - 1} b a^{\Lambda^-} c^{l_k - 1} a^{\Lambda^+} b. \tag{6}$$

We will show that $Sw = S'$. Write $S = S_1 \cup S_2$ with

$$\begin{aligned} S_1 &= \{2, \ldots, n - l_k + 1\} = \{2, \ldots, 1 + \Lambda\}, \\ S_2 &= \{n - l_k + 2, n - l_k + 3\} = \{2 + \Lambda, 3 + \Lambda\}. \end{aligned}$$

Then

$$\begin{aligned} S_1 w &= \{2, \ldots, 1 + \Lambda\} a^{l_k - 1} b a^{\Lambda^-} c^{l_k - 1} a^{\Lambda^+} b \\ &= \{l_k + 1, \ldots, n\} b a^{\Lambda^-} c^{l_k - 1} a^{\Lambda^+} b \\ &= \{l_k + 1, \ldots, n\} a^{\Lambda^-} c^{l_k - 1} a^{\Lambda^+} b \\ &= \left(\{1, \ldots, \Lambda^-\} \cup \{\Lambda^- + l_k + 1, \ldots, n\}\right) c^{l_k - 1} a^{\Lambda^+} b \\ &= \left(\{1, \ldots, \Lambda^-\} \cup \{\Lambda^- + l_k + 1, \ldots, n\}\right) a^{\Lambda^+} b \\ &= \{1, \ldots, \Lambda\} b \\ &= \begin{cases} \{2\} = \{1 + \Lambda\} & \text{if } \Lambda = 1 \\ \{2, \ldots, \Lambda\} & \text{if } \Lambda \geq 2, \end{cases} \end{aligned} \tag{7}$$

where sets of the form $\{x, \ldots, y\}$ with $x > y$ should be interpreted as being empty. This occurs if $\Lambda^- = 0$ or $\Lambda^+ = 0$. Furthermore

$$\begin{aligned} S_2 w &= \{2 + \Lambda, 3 + \Lambda\} a^{l_k - 1} b a^{\Lambda^-} c^{l_k - 1} a^{\Lambda^+} b \\ &= \{1, 2\} b a^{\Lambda^-} c^{l_k - 1} a^{\Lambda^+} b = \{2\} a^{\Lambda^-} c^{l_k - 1} a^{\Lambda^+} b \\ &= \left\{2 + \Lambda^-\right\} c^{l_k - 1} a^{\Lambda^+} b = \left\{1 + \Lambda^-\right\} a^{\Lambda^+} b \\ &= \{1 + \Lambda\} b = \{1 + \Lambda\}. \end{aligned} \tag{8}$$

It follows that the word $w$ has the property

$$Sw = (S_1 \cup S_2)w = S_1 w \cup S_2 w = \{2, \ldots, \Lambda + 1\} = S'.$$

and its length is $|w| = l_k - 1 + 1 + \Lambda^- + l_k - 1 + \Lambda^+ + 1 = 2l_k + \Lambda = l_k + n < 2n$. In the automaton $C_n$ the sets $S$ and $S'$ are both on the shortest path from $Q$ to a singleton and the shortest path is defined by $S(a^{n-1}b)^2 = S'$. Since $|(a^{n-1}b)^2| = 2n > |w|$, the statement of the lemma follows.

The above proof fails in case $l_k \leq 2$, since then $n - l_k + 3 > n$. However, the proofs for the other cases use pretty much the same ideas.

*Case 2:* $L \geq 3$ *and* $l_k = 2$ *for some* $k \leq L-1$. Take $k$ such that $l_k = 2$ and write

$$\Lambda^- = \sum_{i=1}^{k-1} l_i, \quad \Lambda^+ = \sum_{i=k+2}^{L} l_i,$$

for the sum of the loop lengths before the $k$th loop and after the $(k+1)$th loop respectively. These sums can be zero if $k = 1$ or $k = L-1$. Define $\Lambda = \Lambda^- + \Lambda^+ = n - l_k - l_{k+1} \leq n-3$. From the assumption $L \geq 3$ it follows that $\Lambda \geq 1$. Take

$$S = \{2, 3, \ldots, \Lambda + 3\}, \quad S' = \{2, 3, \ldots, \Lambda + 1\}.$$

and define the word

$$w = a^{l_k + l_{k+1} - 1} b a^{\Lambda^-} c a^{\Lambda^+} b.$$

By a similar argument as in Case 1 it follows that $Sw = S'$: Let $S_1 = \{2, \ldots, \Lambda + 1\}$, then

$$\begin{aligned}
S_1 w &= \{2, \ldots, \Lambda + 1\}\, a^{l_k + l_{k+1} - 1} b a^{\Lambda^-} c a^{\Lambda^+} b \\
&= \{l_k + l_{k+1} + 1, \ldots, n\}\, b a^{\Lambda^-} c a^{\Lambda^+} b \\
&= \{l_k + l_{k+1} + 1, \ldots, n\}\, a^{\Lambda^-} c a^{\Lambda^+} b \\
&= \left(\{1, \ldots, \Lambda^-\} \cup \{\Lambda^- + l_k + l_{k+1} + 1, \ldots, n\}\right) c a^{\Lambda^+} b \qquad (9) \\
&= \left(\{1, \ldots, \Lambda^-\} \cup \{\Lambda^- + l_k + l_{k+1} + 1, \ldots, n\}\right) a^{\Lambda^+} b \\
&= \{1, \ldots, \Lambda\}\, b = \begin{cases} \{2\} = \{1 + \Lambda\} & \text{if } \Lambda = 1 \\ \{2, \ldots, \Lambda\} & \text{if } \Lambda \geq 2, \end{cases}
\end{aligned}$$

Completely analogous to Case 1, we have

$$\{\Lambda + 2, \Lambda + 3\}\, w = \{1 + \Lambda\}.$$

Therefore,

$$Sw = \{2, \ldots, \Lambda + 1\}\, w \cup \{\Lambda + 2, \Lambda + 3\}\, w = \{2, \ldots, \Lambda + 1\} = S'$$

Since $w$ has length $n + 2 < 2n$, the statement of the lemma follows.

*Case 3:* $L \geq 3$ *and* $l_L = 2$. Define

$$S = \{2, \ldots, n\}, \qquad w = a^2 b a^{n-3} c a b. \qquad (10)$$

Then

$$\begin{aligned}
Sw &= (\{1, 2\} \cup \{4, \ldots, n\}) b a^{n-3} c a b = (\{2\} \cup \{4, \ldots, n\}) a^{n-3} c a b \\
&= (\{n-1\} \cup \{1, \ldots, n-3\}) c a b = (\{n\} \cup \{1, \ldots, n-3\}) a b \qquad (11) \\
&= \{1, \ldots, n-2\}\, b = \{2, \ldots, n-2\}.
\end{aligned}$$

Since $|w| = n + 3 < 2n$, the result follows. □

### 4.3 The additional symbol has deficiency 1

In this section we assume that the additional symbol $c$ satisfies $|Qc| = n-1$. We will prove that the extended automaton $C_n^c$ admits a synchronizing word of length strictly less than $(n-1)^2$ for every non-trivial choice of $c$. The first step (Lemma's 3, 4, 5 and Corollary 1) is to show that the only candidates to preserve the shortest synchronizing word length have a loop structure similar to the permutations in Lemma 1. In Lemma 6 we couple such candidates $c$ to a permutation $\tilde{c}$, which leads to the conclusion that the automaton with $c$ synchronizes at least as fast as the automaton with $\tilde{c}$.

**Lemma 3.** *Let $n \geq 5$ and let $C_n^c$ be an extension of the automaton $C_n$ by a symbol $c$ for which $|Qc| = n-1$. If the shortest synchronizing word for $C_n^c$ has length $(n-1)^2$, then $Qc = Q \setminus \{1\}$ and $c$ defines a permutation on $Q \setminus \{1\}$.*

*Proof.* If $Qc = Q \setminus \{q\}$ with $q \neq 1$, then $w = ca^{n-q}b(a^{n-1}b)^{n-3}$ is synchronizing and $w$ has length

$$|w| = 1 + n - q + 1 + n(n-3) = (n-1)^2 - q + 1 < (n-1)^2.$$

If $Qc = Q \setminus \{1\}$ and $|Qc^2| \leq n-2$, then one of the following two is true:

- $Qc^2 = Q \setminus \{1, 2\}$.
  In this case $w = c^2a^{n-2}b(a^{n-1}b)^{n-4}$ is synchronizing and has length
  $$|w| = 2 + n - 2 + 1 + n(n-4) = (n-1)^2 - n < (n-1)^2.$$
- $Qc^2 \subset Q \setminus \{q\}$ for some $q \geq 3$.
  In this case $w = c^2a^{n-q}b(a^{n-1}b)^{n-3}$ is synchronizing and $w$ has length
  $$|w| = 2 + n - q + 1 + n(n-3) = (n-1)^2 - q + 2 < (n-1)^2.$$

Therefore, we may assume that $Qc = Q \setminus \{1\}$ and $|Qc^2| = n-1$. This means that $(Q \setminus \{1\})c = Q \setminus \{1\}$, so $c$ defines a permutation on $Q \setminus \{1\}$. □

The next lemma shows that $c$ can be assumed to satisfy $qc \leq q+1$ for all $q$.

**Lemma 4.** *Let $n \geq 5$ and let $C_n^c$ be an extension of the automaton $C_n$ by a symbol $c$ for which $|Qc| = n-1$. If $qc = q + k$ for some $q \in Q$ and $k \geq 2$, then $C_n^c$ admits a synchronizing word of length strictly less than $(n-1)^2$.*

*Proof.* If $qc = q + k$ for some $q \in Q$ and $k \geq 2$, then either $1c \neq 2$ or $qc = q + k$ for some $q \geq 2$ and $k \geq 2$. We distinguish these two cases:

- $1c \neq 2$. In this case there exists a singleton $\tilde{q} := 2c^{-1}$, so
  $$(Q \setminus \{\tilde{q}\})c = Q \setminus \{1, 2\}.$$
  The sets $Q \setminus \{\tilde{q}\}$ and $Q \setminus \{1, 2\}$ are both on the shortest path in $C_n$, where
  $$(Q \setminus \{\tilde{q}\})a^{n-\tilde{q}}ba = Q \setminus \{1, 2\}.$$
  Since $a^{n-\tilde{q}}ba \geq 2$, the shortest synchronizing word in $C_n^c$ has length at most $(n-1)^2 - 1$.

– $1c = 2$ and there exist $q \geq 2$ and $k \geq 2$ such that $qc = q + k$. In this case

$$(Q \setminus \{q\})c = Q \setminus \{1, q + k\} \subseteq Q \setminus \{q + k\},$$

which means that there is synchronizing word of length $(n-1)^2 - k + 1$ in $C_n^c$, see (3).

□

**Lemma 5.** *Suppose $|Q| = n \geq 2$ and $c$ is such that*

$$Qc = Q \setminus \{1\}, \quad (Q \setminus \{1\})\, c = Q \setminus \{1\} \quad \text{and} \quad qc \leq q + 1 \text{ for all } q. \tag{12}$$

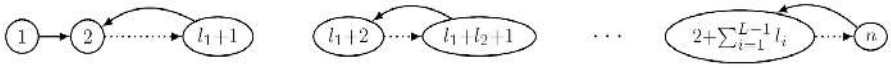
*Then there exist numbers $L$ (number of c-loops) and $1 \leq l_1, \ldots, l_L \leq n-1$ (lengths of c-loops) with $\sum_{i=1}^{L} l_i = n - 1$ such that*

$$qc = \begin{cases} q - l_i + 1 \text{ if } q = l_1 + \ldots + l_i + 1 \text{ for some } 1 \leq i \leq L \\ q + 1 \qquad \text{otherwise} \end{cases} \tag{13}$$

*Proof.* Similar to the proof of Lemma 1. □

**Corollary 1.** *Let $n \geq 5$ and let $C_n^c$ be an extension of the automaton $C_n$ by a symbol $c$ for which $|Qc| = n - 1$. If the shortest synchronizing word for $C_n^c$ has length $(n-1)^2$, then $c$ has the structure described in Lemma 5.*

An illustration of the statement is given below. The structure of $c$ if $|Qc| = n - 1$. Dotted arrows represent chains of transitions of the form $qc = q + 1$.




Finally, in the next lemma, we handle symbols $c$ having the structure described in Lemma 5. If all loops of $c$ have length 1, then $qc = qb$ for all $q \in Q$. Therefore the case $L = n - 1$ is excluded.

**Lemma 6.** *Let $n \geq 5$ and let $C_n^c$ be an extension of the automaton $C_n$ by a symbol $c$ as given in Lemma 5. If $1 \leq L \leq n-2$, then $C_n^c$ admits a synchronizing word of length strictly less than $(n-1)^2$.*

*Proof.* We distinguish two cases: $2 \leq l_1 \leq n - 1$ and $l_1 = 1$.
**Case 1:** $2 \leq l_1 \leq n - 1$. In this case

$$Q \setminus \{l_1\} \xrightarrow{c} Q \setminus \{1, l_1 + 1\} \xrightarrow{c} Q \setminus \{1, 2\}.$$

In $C_n$ the shortest path between these sets is given by

$$Q \setminus \{l_1\} \xrightarrow{a^{n-l_1}} Q \setminus \{n\} \xrightarrow{b} Q \setminus \{1, n\} \xrightarrow{a} Q \setminus \{1, 2\},$$

which has length $n - l_1 + 2 \geq 3$. Therefore, $C_n^c$ has a synchronizing word of length at most $(n-1)^2 - 1$.

*Case 2:* $l_1 = 1$. This means that $2c = 2$. We define a permutation $\tilde{c}$ on $Q$ by

$$q\tilde{c} = \begin{cases} 1 & \text{if } q = 1, \\ qc & \text{if } q \neq 1. \end{cases}$$

The permutation $\tilde{c}$ has $\tilde{L} := L + 1 \geq 2$ loops and $L$ of them coincide with the loops of $c$. Since $c$ has a loop of length at least 2, so does $\tilde{c}$. The loop lengths of $\tilde{c}$ are given by $\tilde{l}_1 = 1$ and $\tilde{l}_k = l_{k-1}$ for $2 \leq k \leq \tilde{L}$

By Lemma 2 we already know that there exists a synchronizing word $\tilde{w} \in \{a, b, \tilde{c}\}^*$ with $|\tilde{w}| < (n-1)^2$. Define $w \in \{a, b, c\}^*$ as the word that is obtained from $\tilde{w}$ by replacing all instances of $\tilde{c}$ by $c$. Clearly this operation preserves the word length. We will show that the word $w$ is a synchronizing word for $C_n^c$.

The key observation is that the permutation $\tilde{c}$ has the following property for $S \subseteq Q$:

$$\text{If } 1 \notin S \text{ or } 2 \in S, \text{ then } Sc^k \subseteq S\tilde{c}^k \text{ for all } k \geq 1. \tag{14}$$

We consider the same cases as in the proof of Lemma 2:

- $\tilde{L} \geq 2$ and $\tilde{l}_k \geq 3$ for some $k$. In this case

  $$\tilde{w} = a^{\tilde{l}_k - 1} b a^{\Lambda^-} \tilde{c}^{\tilde{l}_k - 1} a^{\Lambda^+} b$$

  is synchronizing (compare to (6)), where

  $$\Lambda^- = \sum_{i=1}^{k-1} \tilde{l}_i \geq \tilde{l}_1 + \tilde{l}_2 = 2, \qquad \Lambda^+ = \sum_{i=k+1}^{\tilde{L}} \tilde{l}_i.$$

  Here we used that $\tilde{l}_1 = \tilde{l}_2 = 1$ and $k \geq 3$ since $1\tilde{c} = 1$ and $2\tilde{c} = 2$. Let

  $$T_1 = \left\{1, \ldots, \Lambda^-\right\} \cup \left\{\Lambda^- + l_k + 1, \ldots, n\right\} \quad \text{and} \quad T_2 = \left\{2 + \Lambda^-\right\},$$

  and observe that $2 \in T_1$ and $1 \notin T_2$. By property (14), we obtain

  $$T_1 c^{\tilde{l}_k - 1} \subseteq T_1 \tilde{c}^{\tilde{l}_k - 1}, \quad T_2 c^{\tilde{l}_k - 1} \subseteq T_2 \tilde{c}^{\tilde{l}_k - 1}.$$

  Comparing with the argument in the proof of Lemma 2, in particular (7) and (8), we conclude that $Qw \subseteq Q\tilde{w}$ and $w$ is synchronizing.
- $\tilde{L} \geq 3$ and $\tilde{l}_k = 2$ for some $k \leq \tilde{L} - 1$. Here an analogous argument as in the previous case gives the result.
- $\tilde{L} \geq 3$ and $\tilde{l}_{\tilde{L}} = 2$. Let

  $$\tilde{w} = a^2 b a^{n-3} \tilde{c} a b,$$

  analogous to (10). Since $n \geq 5$, we have

  $$2 \in \{n-1\} \cup \{1, \ldots, n-3\}.$$

  Applying property (14) again, we obtain

  $$(\{n-1\} \cup \{1, \ldots, n-3\})\, c \subseteq (\{n-1\} \cup \{1, \ldots, n-3\})\, \tilde{c}.$$

  By comparing with (11), it follows that $Qw \subseteq Q\tilde{w}$ and therefore $w$ synchronizes.

□

### 4.4 The additional symbol has deficiency at least 2

**Lemma 7.** *Let $n \geq 5$ and let $C_n^c$ be an extension of the automaton $C_n$ by a symbol $c$ such that $|Qc| \leq n-2$. Then $C_n^c$ admits a synchronizing word of length strictly less than $(n-1)^2$.*

*Proof.* There exists $q \geq 2$ such that $Qc \subset Q \setminus \{q\}$, which implies the result. □

*Proof of Theorem 3.* Combining all results of the preceding sections completes the proof. □

## 5 Conclusions and further research

We investigated critical DFAs in two main ways: exploiting computer support we did a full investigation for $n = 3, 4$, and for $n \geq 5$ in classical mathematical style we proved that $C_n$ does not admit non-trivial critical extensions. Further we showed that neither of two more known critical DFAs on 5 and 6 states admit non-trivial critical extensions. If Trahtman's investigation gives all minimal critical DFAs (which is a weaker form of his conjecture), our results give a full characterization of all critical DFAs. In contrast to what Trahtman expected, several minimal critical DFAs on 3 and 4 states can be combined and/or extended to critical DFAs. For all of these the minimal synchronizing word is not unique, and sometimes the synchronizing state is not unique.

Despite of extensive effort, Černý's conjecture is still open after more than half a century. Being a strengthening of this long standing open problem, a full characterization of all critical DFAs may not be tractable. More feasible challenges may include

- a full investigation for $n = 5, 6$,
- proving or disproving that every non-minimal basic critical DFA admits multiple shortest synchronizing words,
- giving an upper bound on the number of symbols in a minimal critical DFA (all known examples have at most three).

## References


1. J. Almeida, S. Margolis, B. Steinberg, M. Volkov. Representation theory of finite semigroups, semigroup radicals and formal language theory. *Transactions of the American Mathematical Society*, 361, 1429–1461, 2009.
2. M.-P. Béal, M.V. Berlinkov, D. Perrin. A quadratic upper bound on the size of a synchronizing word in one-cluster automata. *Int. J. Foundations Comput. Sci.* 22, 277–288, 2011.
3. J. Černý. Poznámka k homogénnym experimentom s konečnými automatmi. *Matematicko-fyzikálny časopis, Slovensk. Akad. Vied,* Vol. 14, No. 3, 208–216, 1964.
4. J. Černý, A. Piricka, and B. Rosenauerova. On directable automata. *Kybernetika*, pages 289–298, 1971.

5. H. Don. The Černý conjecture and 1-contracting automata. *Electronic Journal of Combinatorics* 23 (3), 2016.
6. L. Dubuc. Sur les automates circulaires et la conjecture de Černý. *RAIRO Inform. Theor. Appl.* 32, 21–34, 1998.
7. D. Eppstein. Reset sequences for monotonic automata. *SIAM Journal on Computing*, 19, 500–510, 1990.
8. P. Frankl. An extremal problem for two families of sets. *European Journal of Combinatorics*, 3, 125–127, 1982.
9. J. Kari. A counterexample to a conjecture concerning synchronizing word in finite automata. *EATCS Bulletin*, 73:146–147, 2001.
10. J. Kari. Synchronizing finite automata on Eulerian digraphs. *Theoretical Computer Science*, 295 (1-3), 223–232, 2003.
11. J.-E. Pin. On two combinatorial problems arising from automata theory. *Annals of Discrete Mathematics*, 17, 535–548, 1983.
12. J.-E. Pin. Sur un cas particulier de la conjecture de Černý. *Automata, languages and programming (Fifth Internat. Colloq., Udine, 1978), Lecture Notes in Comput. Sci.*, 62, Springer, Berlin-New York, 345–352, 1978.
13. A. Roman. A note on Černý conjecture for automata with 3-letter alphabet. *Journal of Automata, Languages and Combinatorics*, 13(2), 2008.
14. P.H. Starke. Eine Bemerkung über homogene Experimente. *Elektronische Informationverarbeitung und Kybernetik*, 2, 257–259, 1966.
15. B. Steinberg. The Černý conjecture for one-cluster automata with prime length cycle. *Theoretical Computer Science*, 412 (39), 5487–5491, 2011.
16. A. N. Trahtman. An efficient algorithm finds noticeable trends and examples concerning the Černý conjecture. In Rastislav Královič and Paweł Urzyczyn, editors, *Mathematical Foundations of Computer Science 2006: 31st International Symposium, MFCS 2006*, pages 789–800. Springer Berlin Heidelberg, 2006.
17. A. Trahtman. The Černý conjecture for aperiodic automata. *Discrete Mathematics and Theoretical Computer Science*, Vol. 9, No.2, 3–10, 2007.
18. M.V. Volkov. Synchronizing automata preserving a chain of partial orders. *Theoretical Computer Science*, 410 (37), 3513–3519, 2009.
19. M.V. Volkov. Synchronizing automata and the Černý conjecture. *LATA (2008) 11–27, Lecture Notes in Comput. Sci.*, 5196, Springer, Berlin, 2008.
20. Yices homepage. `http://yices.csl.sri.com/`.