International Journal of Foundations of Computer Science
Vol. 23, No. 1 (2012) 115–131

DOI: 10.1142/S0129054112400084

THE MAGIC NUMBER PROBLEM FOR SUBREGULAR LANGUAGE FAMILIES

MARKUS HOLZER*, SEBASTIAN JAKOBI† and MARTIN KUTRIB‡

Institut für Informatik, Universität Giessen
Arndtstrasse 2, 35392 Giessen, Germany
**holzer@informatik.uni-giessen.de*
†jakobi@informatik.uni-giessen.de
‡kutrib@informatik.uni-giessen.de

Received 22 November 2010
Accepted 2 March 2011
Communicated by Ian McQuillan

We investigate the magic number problem, that is, the question whether there exists a minimal n-state nondeterministic finite automaton (NFA) whose equivalent minimal deterministic finite automaton (DFA) has α states, for all n and α satisfying $n \leq \alpha \leq 2^n$. A number α not satisfying this condition is called a *magic number* (for n). It was shown that no magic numbers exist for general regular languages, whereas trivial and non-trivial magic numbers for unary regular languages were identified. We obtain similar results for automata accepting subregular languages like, for example, star-free languages, prefix-, suffix-, and infix-closed languages, and prefix-, suffix-, and infix-free languages, showing that there are only trivial magic numbers, when they exist. For finite languages we obtain some partial results showing that certain numbers are non-magic.

Keywords: Descriptional complexity; finite automata; magic numbers; subregular languages.

1. Introduction

Nondeterministic finite automata (NFAs) are probably best known for being equivalent to right-linear context-free grammars and, thus, for capturing the lowest level of the Chomsky-hierarchy, the family of regular languages. It is well known that NFAs can offer exponential saving in space compared with deterministic finite automata (DFAs), that is, given some n-state NFA one can always construct a language equivalent DFA with at most 2^n states [27]. This so-called *powerset construction* turned out to be optimal, in general. That is, the bound on the number of states is tight in the sense that for an arbitrary n there is always some n-state NFA which cannot be simulated by any DFA with less than 2^n states [21, 25, 26]. On the other hand, there are cases where nondeterminism does not help for the succinct representation of a language compared to DFAs. These two milestones from the early days of automata theory form part of an extensive list of equally striking problems of

NFA related problems, and are a basis of descriptional complexity. Moreover, they initiated the study of the power of resources and features given to finite automata. For recent surveys on descriptional complexity issues of regular languages we refer to, for example, [10, 11, 12].

Nearly a decade ago a very fundamental question on the well known powerset construction was raised in [13]: does there always exists a minimal n-state NFA whose equivalent minimal DFA has α states, for all n and α with $n \leq \alpha \leq 2^n$? A number α not satisfying this condition is called a *magic number* for n. The answer to this simple question turned out not to be so easy. For NFAs over a two-letter alphabet it was shown that $\alpha = 2^n - 2^k$ or $2^n - 2^k - 1$, for $0 \leq k \leq n/2 - 2$ [13], and $\alpha = 2^n - k$, for $5 \leq k \leq 2n-2$ and some coprimality condition for k [14], are non-magic. In [16] it was proven that the integer α is non-magic, if $n \leq \alpha \leq 1+n(n+1)/2$. This result was improved by showing that α is non-magic for $n \leq \alpha \leq 2^{\sqrt[3]{n}}$ in [17]. Further non-magic numbers for two-letter input alphabet were identified in [5] and [23]. It turned out that the problem becomes easier if one allows more input letters. In fact, for exponentially growing alphabets there are no magic numbers at all [16]. This result was improved to less growing alphabets in [5], to constant alphabets of size four in [15], and very recently to three-letter alphabets [19]. Magic numbers for unary NFAs were recently studied in [6] by revising the Chrobak normal-form for NFAs. In the same paper also a brief historical summary of the magic number problem can be found. Further results on the magic number problem (in particular in relation to the operation problem on regular languages) can be found, for example, in [17, 18].

To our knowledge the magic number problem was not systematically studied for subregular languages families, except for unary languages. Several of these subfamilies are well motivated by their representations as finite automata or regular expressions: finite languages (are accepted by acyclic finite automata), star-free languages or regular non-counting languages (which can be described by regular-like expression using only union, concatenation, and complement), prefix-closed languages (are accepted by automata where all states are accepting), suffix-closed (or multiple-entry or fully-initial) languages (are accepted by automata where the computation can start in any state), infix-closed languages (are accepted by automata where all states are both initial and accepting), suffix-free languages (are accepted by non-returning automata, that is, automata where the initial state does not have any in-transition), prefix-free languages (are accepted by non-exiting automata, that is, automata where all out-transitions of every accepting state go to a rejecting sink state), and infix-free languages (are accepted by non-returning and non-exiting automata, where these conditions are necessary, but not sufficient).

The hierarchy of these and some further subregular language families is well known. We study all families mentioned with respect to the magic number problem, and show—except for finite languages, where only some partial results will be presented—that there are only trivial magic numbers, whenever they exist.

2. Definitions

Let Σ^* denote the set of all *words* over a finite alphabet Σ. For $n \geq 0$ we write Σ^n for the set of all words of length n. The *empty word* is denoted by λ and $\Sigma^+ = \Sigma^* \setminus \{\lambda\}$. A *language* L over Σ is a subset of Σ^*. For the length of a word w we write $|w|$. Set inclusion is denoted by $\subseteq$ and strict set inclusion by $\subset$. We write 2^S for the power set and $|S|$ for the cardinality of a set S.

A *nondeterministic finite automaton* (NFA) is a quintuple $A = (Q, \Sigma, \delta, q_0, F)$, where Q is the finite set of *states*, Σ is the finite set of *input symbols*, $q_0 \in Q$ is the *initial state*, $F \subseteq Q$ is the set of *accepting states*, and $\delta : Q \times \Sigma \to 2^Q$ is the *transition function*. As usual the transition function is extended to $\delta : Q \times \Sigma^* \to 2^Q$ reflecting sequences of inputs: $\delta(q, \lambda) = \{q\}$ and $\delta(q, aw) = \bigcup_{q' \in \delta(q,a)} \delta(q', w)$, for $q \in Q$, $a \in \Sigma$, and $w \in \Sigma^*$. A word $w \in \Sigma^*$ is *accepted* by A if $\delta(q_0, w) \cap F \neq \emptyset$. The *language accepted* by A is $L(A) = \{ w \in \Sigma^* \mid w \text{ is accepted by } A \}$.

A finite automaton is *deterministic* (DFA) if and only if $|\delta(q, a)| = 1$, for all $q \in Q$ and $a \in \Sigma$. In this case we simply write $\delta(q, a) = p$ for $\delta(q, a) = \{p\}$ assuming that the transition function is a mapping $\delta : Q \times \Sigma \to Q$. So, any DFA is complete, that is, the transition function is total, whereas for NFAs it is possible that δ maps to the empty set. Note that a sink state is counted for DFAs, since they are always complete, whereas it is not counted for NFAs, since their transition function may map to the empty set. In the sequel we refer to the DFA obtained from an NFA $A = (Q, \Sigma, \delta, q_0, F)$ by the power-set construction as $A' = (2^Q, \Sigma, \delta', \{q_0\}, F')$, where $\delta'(P, a) = \bigcup_{p \in P} \delta(p, a)$, for $P \subseteq Q$ and $a \in \Sigma$, and $F' = \{ P \subseteq Q \mid P \cap F \neq \emptyset \}$.

As already mentioned in the introduction, in [15] it was shown that for all integers n and α such that $n \leq \alpha \leq 2^n$, there exists an n-state nondeterministic finite automaton $A_{n,\alpha}$ whose equivalent minimal deterministic finite automaton has exactly α states. Since some of our constructions rely on this proof and for the sake of completeness and readability we briefly recall the sketch of the construction. In the following we call the NFA $A_{n,\alpha}$ the Jirásek-Jirásková-Szabari automaton, or for short the JJS-automaton. The next result is from [15].

Theorem 1. *For all integers n and α such that $n \leq \alpha \leq 2^n$, there exists an n-state nondeterministic finite automaton $A_{n,\alpha}$ over a four-letter alphabet whose equivalent minimal deterministic finite automaton has exactly α states.*

In the construction for some fixed integer n the cases $\alpha = n$ and $\alpha = 2^n$ are treated separately by appropriate witness languages. For the remaining cases it is first shown that every α satisfying $n < \alpha < 2^n$ can be written as a specific sum of powers of two. In particular, for all integers n and α such that $n < \alpha < 2^n$, there exist integers k and m with $1 \leq k \leq n-1$ and $1 \leq m < 2^k$, such that $\alpha = n - (k+1) + 2^k + m$ and

$$m = (2^{k_1} - 1) + (2^{k_2} - 1) + \cdots + (2^{k_{\ell-1}} - 1) + \begin{cases} (2^{k_\ell} - 1) \\ 2 \cdot (2^{k_\ell} - 1), \end{cases}$$

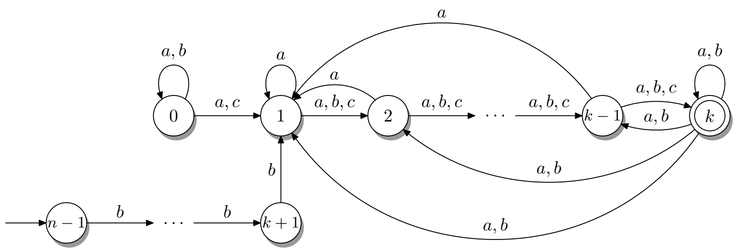


Fig. 1. Jirásek-Jirásková-Szabari's (JJS) nondeterministic finite automaton $A_{n,\alpha}$ with n states (d-transitions are not shown) accepting a language for which the equivalent minimal DFA needs exactly $\alpha = n - (k+1) + m$ states.

where $1 \leq \ell \leq k-1$ and $k \geq k_1 > k_2 > \cdots > k_\ell \geq 1$. Then NFAs are constructed such that the powerset construction yields DFAs whose number of states is exactly one of these powers of two, which finally are combined appropriately to lead to a single NFA $A_{n,\alpha}$ with state set $\{0, 1, \ldots, n-1\}$, initial state $q_0 = n-1$ if $k < n-1$, and $q_0 = 1$ otherwise, and sole accepting state k, such that the equivalent minimal DFA has exactly α states. Automaton $A_{n,\alpha}$ is depicted in Figure 1, where the following d-transitions are not shown:

$$\delta(i,d) = \begin{cases} \{0,2,3,4,\ldots,k-k_i+1\} & \text{if } 1 \leq i \leq \ell-1 \\ \{0,1,\ldots,k-k_i+1\} & \text{if } i = \ell \text{ and } m \text{ is of the first form} \\ \{0,2,3,4,\ldots,k-k_i+1\} & \text{if } i = \ell \text{ and } m \text{ is of the second form} \\ \{0,1,\ldots,k-k_i+1\} & \text{if } i = \ell+1 \text{ and } m \text{ is of the second form} \\ \emptyset & \text{otherwise.} \end{cases}$$

Finally, we briefly recall the so-called (extended) *fooling set* technique (see, for example, [1, 7, 11]) that is widely used for proving lower bounds on the number of states necessary for an NFA to accept a given language.

Theorem 2 (Extended Fooling Set Technique) *Let $L \subseteq \Sigma^*$ be a regular language and suppose there exists a set of pairs $S = \{ (x_i, y_i) \mid 1 \leq i \leq n \}$ such that (1) $x_i y_i \in L$, for $1 \leq i \leq n$, and (2) $i \neq j$ implies $x_i y_j \notin L$ or $x_j y_i \notin L$, for $1 \leq i, j \leq n$. Then any nondeterministic finite automaton accepting L has at least n states. Here S is called an (extended)* fooling set *for L.*

3. Results

We systematically investigate the magic number problem for the aforementioned subregular language families. For the remaining theorems of this paper, when speaking of an n-state NFA we always mean a minimal NFA. Given a subregular language family, if $f(n)$ is the number of states that is sufficient and necessary in the worst case for a DFA to accept the language of an n-state NFA belonging to the family, then a number α with $f(n) < \alpha \leq 2^n$ is called a *trivial* magic number. Similarly, if $g(n)$ is the number of states that is necessary for any DFA simulating an arbitrary n-state NFA, then all numbers α with $\alpha < g(n)$ are also called *trivial* magic numbers. For example, for infix-free languages $g(n)$ is shown to be $n+1$ in Theorem 6, while $f(n)$ is known to be $2^{n-2}+2$ [2]. An observation from [2] shows that the magic number problem for *elementary* and *combinational* languages is trivial.

3.1. *Star-free languages and power separating languages*

A language $L \subseteq \Sigma^*$ is *star-free* (or regular *non-counting*) if and only if it can be obtained from the elementary languages $\{a\}$, for $a \in \Sigma$, by applying the Boolean operations union, complementation, and concatenation finitely often. These languages are exhaustively studied, for example, in [24]. Since regular languages are closed under Boolean operations and concatenation, every star-free language is regular. On the other hand, not every regular language is star free.

Here we use an alternative characterization of star-free languages by so called permutation-free automata [24]: a regular language $L \subseteq \Sigma^*$ is star-free if and only if the minimal DFA accepting L is *permutation-free*, that is, there is *no* word $w \in \Sigma^*$ that induces a non-trivial permutation of any subset of the set of states. Here a trivial permutation is simply the identity permutation. Observe that a word uw induces a non-trivial permutation $\{q_0, q_1, \ldots, q_{r-1}\} \subseteq Q$ in a DFA with state set Q and transition function δ if and only if wu induces a non-trivial permutation $\{\delta(q_0, u), \delta(q_1, u), \ldots, \delta(q_{r-1}, u)\}$ in the same automaton. Further, if one finds a non-trivial permutation consisting of multiple disjoint cycles, it suffices to consider a single cycle. Before we show that no magic numbers exist for star-free languages we prove a useful lemma on permutations in (minimal) DFAs obtained by the powerset construction.

Lemma 3. *Let A be a nondeterministic finite automaton with state set Q over alphabet Σ, and assume that A' is the equivalent minimal deterministic finite automaton, which is non-permutation-free. If the word w in Σ^* induces a non-trivial permutation on the state set $\{P_0, P_1, \ldots, P_{r-1}\} \subseteq 2^Q$ of A', that is, $\delta'(P_i, w) = P_{i+1}$, for $0 \leq i < r-1$, and $\delta'(P_{r-1}, w) = P_0$, then there are no two states P_i and P_j with $i \neq j$ such that $P_i \subseteq P_j$.*

Proof. Assume to the contrary that $P_0 \subseteq P_i$ (possibly after a cyclic shift), for some $0 < i \leq r-1$. Then one can show by induction that $\delta'(P_0, v) \subseteq \delta'(P_i, v)$, for every word $v \in \Sigma^*$. In particular, this also holds true for the word w that induces the non-trivial permutation on the state set $\{P_0, P_1, \ldots, P_{r-1}\}$. But then $P_{ki \bmod r} = \delta'(P_0, w^{ki}) \subseteq \delta'(P_i, w^{ki}) = P_{(k+1)i \bmod r}$, for $k \geq 0$, and one finds the chain of inclusions $P_0 \subseteq P_i \subseteq P_{2i \bmod r} \subseteq P_{3i \bmod r} \subseteq \cdots \subseteq P_{ri \bmod r} = P_0$, which implies $P_0 = P_i$, a contradiction. □

Now we are prepared for the main theorem, which utilizes Lemma 3.

Theorem 4. *For all integers n and α such that $n \leq \alpha \leq 2^n$, there exists an n-state nondeterministic finite automaton accepting a* star-free *language whose equivalent minimal deterministic finite automaton has exactly α states.*

Proof. The case $\alpha = 2^n$ is treated in [2], and for $\alpha = n$ consider the unary star-free language $L_n = a^{n-1}a^*$ accepted by an n-state NFA, where any equivalent minimal DFA has also exactly n states. Thus, in the following assume that $n < \alpha < 2^n$. We show that the JJS-automaton $A_{n,\alpha}$ accepts a star-free language, by proving that $A'_{n,\alpha}$ is permutation free. Let $\alpha = n - (k+1) + 2^k + m$ be as described in Section 2 and let $Q = \{0, 1, \ldots, n-1\}$ and $\Sigma = \{a, b, c, d\}$.

Assume to the contrary that a word $w \in \Sigma^*$ induces a non-trivial permutation on the state set $\{P_0, P_1, \ldots, P_{r-1}\}$ of $A'_{n,\alpha}$, for some $r \geq 2$. By the construction of $A_{n,\alpha}$ it suffices to consider the case where $P_i \subseteq \{0, 1, \ldots, k\}$, for $0 \leq i \leq r-1$, since after reading $n - (k+1)$ letters one cannot return to the states $k+1, k+2, \ldots, n-1$ in the tail. We consider two cases, namely whether w contains at least one letter d or not: (1) Without loss of generality, we assume that one letter d appears as last letter in w, that is, $w = w'd$, for some $w' \in \Sigma^*$. Then there are state sets $M_i \subseteq Q$ such that $\delta'(M_i, d) = P_i$, for $0 \leq i \leq r-1$. Since

$$\delta(1, d) \subseteq \delta(2, d) \subseteq \cdots \subseteq \delta(\ell, d)[\subseteq \delta(\ell+1, d)],$$

by the construction of the JJS-automaton we may conclude that there are at least two states P_i and P_j from the non-trivial permutation in $A'_{n,\alpha}$ such that $P_i \subseteq P_j$. The inequality must hold true, since the value of each $\delta'(M_i, d)$, for $0 \leq i \leq r-1$, depends only on the largest element on which a d-transition is defined. But then we obtain a contradiction by Lemma 3 and, thus, word w cannot contain any letter d. (2) Next let $w \in \{a, b, c\}^*$. Let us show that in this case, we can safely concentrate on the states $\{1, 2, \ldots, k\}$. First, if one P_i contains state 0, then all P_j with $0 \leq j \leq r-1$ must contain 0, too, and reading a c is not allowed because without a d one cannot return to state 0 in $A_{n,\alpha}$ by construction. Second, the only state reachable from $\{0\}$ is $\{0, 1\}$, but from any other state containing one of $\{1, 2, \ldots, k\}$ one reaches state 1 by reading a, too. Therefore, if there is a non-trivial permutation $\{P_0, P_1, \ldots, P_{r-1}\}$ with $0 \in P_i$, for $0 \leq i \leq r-1$ in $A'_{n,\alpha}$, then there is also another one in which element 0 does not appear.

We distinguish two further subcases: let $w \in \{a, b\}^*$. The non-trivial permutation under consideration contains r states from $A'_{n,\alpha}$, which for instance implies $\delta'(P_0, w^{rk}) = P_0$. Since w contains at least one letter, both letters a and b move state i, with $1 \leq i < k$, one step closer to k, and $\delta(k, a) = \delta(k, b) = \{1, 2, \ldots, k\}$ in $A_{n,\alpha}$, we conclude that $\delta'(P_0, w^{rk}) = \{1, 2, \ldots, k\}$. Thus, P_0 is a superset of all P_i, for $0 \leq i \leq r-1$, which contradicts Lemma 3.

It remains to consider the case that w contains a letter c. Without loss of generality, we may assume that $w = cw'$, for some $w' \in \{a, b, c\}^*$. Then we find the following situations:

(a) For every i with $0 \leq i \leq r-1$, after reading a prefix v of w^{rk} long enough, from state P_i in $A'_{n,\alpha}$ a state containing k is reached, since all three letters move state j, with $1 \leq j < k$, one step closer to k in $A_{n,\alpha}$.
(b) As soon as $k \in \delta(P_i, v)$, for some prefix v of w and some state P_i of $A'_{n,\alpha}$, the prefix v must be extended by c (to become a prefix of w^{rk} again), since otherwise $\delta'(P_i, va)$ must be $\{1, 2, \ldots, k\}$, which contradicts Lemma 3.
(c) Since all states in the non-trivial permutation are different, it must contain states P_i and P_j in $A'_{n,\alpha}$ and there is a prefix v of w^{rk} such that exactly one state $\delta'(P_i, v)$ or $\delta'(P_j, v)$ contains element k.

Now we turn to prove that this case neither succeeds. To this end we define $\langle R, S\rangle = |R \setminus S| + |S \setminus R|$, for $R, S \in 2^Q$. For two arbitrary states P_i and P_j from the non-trivial permutation we have:

(1) If both states P_i and P_j do not contain element k, then $\langle P_i, P_j\rangle$ is equal to $\langle \delta'(P_i, a), \delta'(P_j, a)\rangle$. The same holds true for the letters b and c.
(2) If k is in *exactly* one of the states P_i or P_j, then one finds $\langle P_i, P_j\rangle = \langle \delta'(P_i, c), \delta'(P_j, c)\rangle - 1$. The cases when reading a and b cannot appear due to the first property above.
(3) Finally, if both states P_i and P_j contain element k, then $\langle P_i, P_j\rangle = \langle \delta'(P_i, c), \delta'(P_j, c)\rangle$. Again reading a or b is impossible by a similar reasoning as above.

In particular, this shows that $\langle P_i, P_j\rangle$ is never increased. Since for all states P_i, for $0 \leq i \leq r-1$, in the non-trivial permutation $\delta'(P_i, w^{rk}) = P_i$ we have $\langle P_i, P_j\rangle = \langle \delta'(P_i, w^{rk}), \delta'(P_j, w^{rk})\rangle$. According to the above mentioned third property there exists a situation while reading the word w^{rk} such that the value of $\langle P_i, P_j\rangle$ is strictly decreased. Since $\langle P_i, P_j\rangle$ is never increased again, we obtain a contradiction. Thus, there is no word w that induces a non-trivial permutation on a subset of states from $A'_{n,\alpha}$. □

The previous theorem generalizes to all language families that are a superset of the family of star-free languages such as, for example, the family of power separating languages introduced in [29].

3.2. *Stars and comet languages*

Consider languages over an alphabet Σ. A language L is a *star language* if and only if $L = H^*$, for some regular language H, and L is a *comet language* if and only if it can be represented as concatenation G^*H of a regular star language G^* and a regular language H, such that $G \neq \{\lambda\}$ and $G \neq \emptyset$. Star languages and comet languages were introduced in [3] and [4]. Next, a language L is a *two-sided comet language* if and only if $L = EG^*H$, for a regular star language G^* and regular languages E and H, such that $G \neq \{\lambda\}$ and $G \neq \emptyset$. So, (two-sided) comet languages are always infinite. Clearly, every star language not equal to $\{\lambda\}$ is also a comet language and every comet is a two-sided comet language, but the converse is not true in general.

Theorem 5. *For all integers n and α such that $n \leq \alpha \leq 2^n$, there exists an n-state nondeterministic finite automaton accepting a* star *language whose equivalent minimal deterministic finite automaton has exactly α states. The statement remains valid for* (two-sided) comet *languages.*

Proof. For $\alpha = n$ we consider the language $L_n = (a^n)^*$, which can be accepted by a cyclic n-state DFA. Furthermore, n states are also necessary for any NFA to accept L_n, since $\{ (a^i, a^{n-i}) \mid 0 \leq i \leq n-1 \}$ is a fooling set for L_n. In case $\alpha = 2^n$ we take the NFA A from [25]. Its single accepting state is also the initial state, so it is $L(A)^* = L(A)$, and then $L(A)$ is a star language.

Now assume $n < \alpha < 2^n$. We adapt the n-state JJS-automaton $A_{n,\alpha}$ over the alphabet $\{a, b, c, d\}$ with state set Q, initial state q_0, and transition function δ as follows. Define the NFA $B_{n,\alpha} = (Q, \{a, b, c, d, \#\}, \delta', k, \{k\})$. The transition function δ' of $B_{n,\alpha}$ is obtained from δ by adding a transition on the input symbol $\#$ from state k to state q_0. Then $B_{n,\alpha}$ accepts a star language, since its single accepting state is the initial state. The minimality of $B_{n,\alpha}$ can be shown by similar means as the minimality of $A_{n,\alpha}$ by slightly adapting the appropriate fooling set S of $A_{n,\alpha}$ to a fooling set $S' = \{ (\#x, y) \mid (x, y) \in S \}$ for $B_{n,\alpha}$. Finally, proving that the equivalent minimal DFA $B'_{n,\alpha}$ has exactly α states is done similarly as in the original proof for the DFA $A'_{n,\alpha}$: any state that $A'_{n,\alpha}$ reaches on the input $w \in \{a, b, c, d\}^*$ is reachable in $B'_{n,\alpha}$ on the input $\#w$. Further, reading $\#$ in any state of $B'_{n,\alpha}$ always leads to states $\{q_0\}$ or $\emptyset$, which are both reachable in $A'_{n,\alpha}$. □

3.3. *Subword specific languages*

In this section we consider languages for which for every word in the language either all or none of its prefixes, suffixes or infixes belong to the same language. Again, there are only trivial magic numbers. We start with subword-free languages.

A language $L \subseteq \Sigma^*$ is *prefix-free* if and only if $y \in L$ implies $yz \notin L$, for all words $z \in \Sigma^+$, *infix-free* if and only if $y \in L$ implies $xyz \notin L$, for all $xz \in \Sigma^+$, and *suffix-free* if and only if $y \in L$ implies $xy \notin L$, for all $x \in \Sigma^+$. Note that DFAs accepting non-empty prefix- or suffix-free languages must have a non-accepting sink state [8, 9], thus n is trivially magic in these cases.

Theorem 6. *Let A be a* minimal *n-state nondeterministic finite automaton accepting a non-empty prefix-, suffix- or infix-free language. Then any equivalent minimal deterministic finite automaton accepting language $L(A)$ needs at least $n + 1$ states.*

In the following we show that no non-trivial magic numbers exist for subword-free languages. The upper bound for the deterministic blow-up in prefix- and suffix-free languages is $2^{n-1} + 1$ and for infix-free languages it is $2^{n-2} + 2$, so all numbers above are trivially magic.

Theorem 7. *For all integers n and α such that $n < \alpha \leq 2^{n-1} + 1$, there exists an n-state nondeterministic finite automaton accepting a* prefix-free *language whose equivalent minimal deterministic finite automaton has exactly α states. The statement remains true for nondeterministic finite automata accepting* suffix-free *languages.*

Proof. The case $\alpha = 2^{n-1} + 1$ for prefix- and suffix-free languages is treated in [2]. This also covers the statement for $n \leq 2$. Let $3 \leq n < \alpha \leq 2^{n-1}$, then the JJS-automaton $A_{n-1,\alpha-1} = (Q, \Sigma, \delta, q_0, \{k\})$ with states $Q = \{0, 1, \ldots, n-2\}$ can be constructed because $(n-1) < \alpha - 1 < 2^{n-1}$. This automaton can be extended with a new initial state (accepting state) such that the accepted language is suffix-free (prefix-free). More precisely, in the suffix-free case we construct $B_{n,\alpha} = (Q \cup \{s\}, \Sigma \cup \{\#\}, \delta_1, s, \{k\})$ with $\delta_1(s, \#) = \{q_0\}$, and $\delta_1(q, a) = \delta(q, a)$ for all $q \in Q$ and $a \in \Sigma$. Then $L(B_{n,\alpha}) \subseteq \{\#\}\{a, b, c, d\}^+$ is suffix-free. The minimality of $A_{n-1,\alpha-1}$ transfers to $B_{n,\alpha}$: let S be the fooling set for $A_{n-1,\alpha-1}$, then $S' \cup \{(\lambda, \#b^{n-k-1}c^{k-1})\}$ is a fooling set for $B_{n,\alpha}$, where $S' = \{\, (\#x, y) \mid (x, y) \in S \,\}$. In the powerset automaton $B'_{n,\alpha}$ all $\alpha - 1$ states from $A'_{n-1,\alpha-1}$ and the additional initial state $\{s\}$ are reachable and distinguishable, since $\{s\}$ is the only state that does not lead to $\emptyset$ on $\#$. Note that $\emptyset$ is already reachable in $A'_{n,\alpha}$, so $\{s\}$ is the only state in $B'_{n,\alpha}$ that is not in $A'_{n,\alpha}$. Thus, we have exactly n states in the NFA and α states in the DFA. An automaton accepting a prefix-free language can be constructed analogously: let $C_{n,\alpha} = (Q \cup \{t\}, \Sigma \cup \{\#\}, \delta_2, q_0, \{t\})$ with $\delta_2(k, \#) = \{t\}$ and $\delta_2(q, a) = \delta(q, a)$ for $q \in Q \setminus \{k\}$ and $a \in \Sigma$. Then $L(C_{n,\alpha}) \subseteq \{a, b, c, d\}^+\{\#\}$ is prefix-free, $C_{n,\alpha}$ has n states and $C'_{n,\alpha}$ has α states. □

For infix-free regular languages the situation is slightly different compared to above. Note that the only nonempty infix-free regular language accepted by a nondeterministic finite automaton with a single state is $\{\lambda\}$, whose minimal deterministic finite automaton has two states.

Theorem 8. *For all integers n and α such that $2 \leq n < \alpha \leq 2^{n-2} + 2$, there exists an n-state nondeterministic finite automaton accepting an* infix-free *language whose equivalent minimal deterministic finite automaton has exactly α states.*

Proof. For $\alpha = 2^{n-2} + 2$ we refer once more to [2], which covers all possible values of α for $2 \leq n \leq 3$. So, let $4 \leq n < \alpha < 2^{n-2} + 2$ and $A_{n-2,\alpha-2} = (Q, \Sigma, \delta, q_0, \{k\})$ be the JJS-automaton with state set $Q = \{0, 1, \ldots, n-3\}$. Combining both constructions from the proof of Theorem 7, we get $B_{n,\alpha} = (Q \cup \{s, t\}, \Sigma \cup \{\#\}, \delta_1, s, \{t\})$ with additional transitions $\delta_1(s, \#) = \{q_0\}$, $\delta_1(k, \#) = \{t\}$, and $\delta_1(q, a) = \delta(q, a)$ for $q \in Q$ and $a \in \Sigma$. This automaton accepts an infix-free language and gives a blow-up of α states. □

Next, we consider prefix-, infix-, and suffix-closed languages. A language $L \in \Sigma^*$ is *prefix-closed* if and only if $xy \in L$ implies $x \in L$, for $x \in \Sigma^*$, *infix-closed* if and only if $xyz \in L$ implies $y \in L$, for $x, z \in \Sigma^*$, and *suffix-closed* if and only if $yz \in L$ implies $z \in L$, for $z \in \Sigma^*$. We use the following results from [20].

Theorem 9. *(1) A nonempty regular language is prefix-closed if and only if it is accepted by some nondeterministic finite automaton with all states accepting. (2) A nonempty regular language is infix-closed if and only if it is accepted by some nondeterministic finite automaton with multiple initial states with all states both initial and accepting.*

Prefix-closed languages reach the upper bound of 2^n states, and for infix-closed languages it is $2^{n-1}+1$. We will show that up to these bounds the only magic number for both language families is n (except for $n = 1$). The bound for suffix-closed languages is $2^{n-1}+1$, and we show that up to this, no magic numbers exist.

Theorem 10. *For all integers n and α such that $n < \alpha \leq 2^{n-1}+1$, there exists an n-state nondeterministic finite automaton accepting an* infix-closed *language whose equivalent minimal deterministic finite automaton has exactly α states. The case $\alpha = n$ can only be reached for $n = 1$.*

Proof. For the second statement, note that each DFA accepting a language $L \neq \Sigma^*$ needs a non-accepting state, which the minimal NFA cannot have, due to Theorem 9. So, Σ^* is the only infix-closed language, for which the size of the minimal DFA equals the size of an equivalent minimal NFA. Both have a single state. The case $\alpha = 2^{n-1}+1$, trivially covering the statement for $n = 2$, is discussed in [2]. For the remaining, assume $3 \leq n < \alpha \leq 2^{n-1}$. In this case, the JJS-automaton $A_{n,\alpha} = (Q, \Sigma, \delta, n-1, \{k\})$ has a non-empty initial tail of states, that is, the initial state is equal to state $n-1$. From $A_{n,\alpha}$ we construct an automaton $A_1 = (Q, \Sigma \cup \{\#, \$\}, \delta_1, Q, Q)$ with all states initial and accepting and transition function $\delta_1(k, \#) = \{k\}$, $\delta_1(q, \$) = \{q-1\}$ if $k+2 \leq q \leq n-1$, $\delta_1(k+1, \$) = \{1\}$ and $\delta_1(q, a) = \delta(q, a)$ for $0 \leq q \leq k$ and $a \in \Sigma$. This NFA with multiple initial states can be converted into an equivalent NFA A_2 with initial state $n-1$ and transition function $\delta_2(n-1, a) = \bigcup_{q \in Q} \delta_1(q, a)$ and $\delta_2(q, a) = \delta_1(q, a)$ for all $a \in \Sigma \cup \{\#, \$\}$ and $q \in Q \setminus \{n-1\}$. With $S_1 = \{\, (\$^i, \$^{n-(k+1)-i}c^{k-1}) \mid 0 \leq i \leq n-(k+1) \,\}$, $S_2 = \{\, (\$^{n-(k+1)}c^i, c^{k-1-i}) \mid 1 \leq i \leq k-1 \,\}$ and $S_3 = \{(\$^{n-(k+1)}d, c^k)\}$, let us check, that $S = S_1 \cup S_2 \cup S_3$ is a fooling set for $L(A_2)$. Different pairs from S_1 result in a word beginning with more than $n-(k+1)$ \$-symbols, pairs from S_2 result in too many c-symbols, c^k from S_3 cannot be combined with any other word and mixing pairs from S_1 and S_2 either results in a word containing the infix $\$c^i\$$ or, if $(\$^{n-(k+1)}, c^{k-1})$ is chosen from S_1, in $\$^{n-(k+1)}c^{i+k-1}$ with too many c-symbols.

In the corresponding powerset automaton A_2', by reading prefixes of $\$^{n-(k+1)}$, one reaches the $n-(k+1)$ states $\{n-1\}$, $\{n-2, \ldots, k+1, 1\}$, $\ldots$, $\{k+1, 1\}$. After reading $\$^{n-(k+1)}$, A_2' is in state $\{1\}$ and from there, according to the JJS-construction, 2^k+m states from $2^{\{0,1,\ldots,k\}}$ are reachable. So we have exactly α states. To see that no further states can be reached, note that the transition function differs from the one of the JJS-automaton only in states $k+1, \ldots, n-1$ and state k. The #-transition in state k gives no new reachable states and reading \$ always leads to either a state $\{n-i, \ldots, k+1, 1\}$, for some $1 \leq i \leq n-(k+1)$, or to state $\{1\}$ or the empty set. So, the only interesting transitions are those of the initial state $\{n-1\}$ on the input symbols a, b, c, and d. Reading a or b leads to $\{0, 1, \ldots, k\}$, reading c to $\{1, 2, \ldots, k\}$ and on input d, A'_2 enters the state $\delta(q, d)$ for the largest $q \in Q$ for which this transition is defined. All these states were already counted.

To prove that any two distinct states $M, N \subseteq Q \setminus \{n-1\}$ are pairwise inequivalent, without loss of generality, pick an element $q \in M \setminus N$. If $q \leq k$, the word $c^{k-q}\#$ distinguishes M and N. Otherwise, if $q \geq k+1$, one can drive it to state 1 by reading \$-symbols, and then c^{k-1} distinguishes the two states. Finally, state $\{n-1\}$ is inequivalent with any state $N \subseteq Q \setminus \{n-1\}$ by the input word $\$^{n-(k+1)}c^{k-1}$. □

The family of infix-closed languages is a subset of the family of suffix-closed languages, so the previous theorem generalizes to the latter language family, except for n which is not magic for $n \geq 1$ anymore.

Corollary 11. *For all integers n and α such that $n \leq \alpha \leq 2^{n-1}+1$, there exists an n-state nondeterministic finite automaton accepting a* suffix-closed *language whose equivalent minimal deterministic finite automaton has exactly α states.*

Proof. Since all infix-closed languages are suffix-closed, we only have to prove the case $\alpha = n$. The witness language for the case $n = 1$ is Σ^* again. For $n > 1$, consider the deterministic finite automaton $A = (\{1, 2, \ldots, n\}, \{a, b\}, \delta, 1, \{1, 2, \ldots, n-1\})$ with transitions $\delta(q, a) = q+1$ if $q < n$, $\delta(n, a) = n$ and $\delta(q, b) = 1$ for all $q \in Q$. The automaton is depicted in Figure 2. A word w is not accepted by A, if and only if its last $n-1$ letters are a. So w is accepted if $|w| < n-1$ or $w = w_1 b w_2$ with $|w_2| < n-1$. In both cases all suffixes of w also satisfy these conditions and, therefore, are accepted by A. Thus, $L(A)$ is suffix-closed.

To prove the minimality of A, assume there is an NFA $B = (Q, \{a, b\}, \delta', q_0, F)$ accepting $L(A)$, with $|Q| < n$ states. Since $a^j \notin L(A)$ for all $j \geq n-1$, but the words a^i belong to $L(A)$, for all $i < n-1$, the automaton B hast to count up to $n-2$. It follows, that Q contains at least $n-1$ states. Then, $\delta'(q_0, a^{n-1}) = \emptyset$ and so $\delta'(q_0, a^{n-1}b) = \emptyset$, which contradicts $a^{n-1}b \in L(A)$. Therefore, there is no NFA accepting $L(A)$ with less than n states. □

Since the upper bound for the deterministic blow-up of prefix-closed languages is greater than that of infix-closed languages, we need to treat them separately here.

Theorem 12. *For all integers n and α such that $n < \alpha \leq 2^n$, there exists an n-state nondeterministic finite automaton accepting a* prefix-closed *language whose equivalent minimal deterministic finite automaton has exactly α states. The case $\alpha = n$ can only be reached for $n = 1$.*

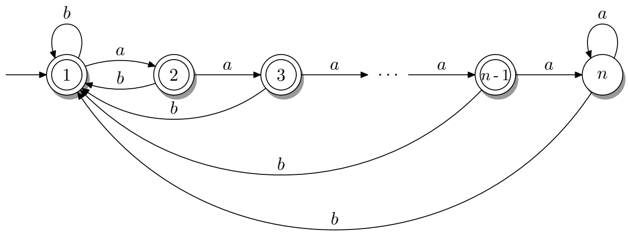


Fig. 2. The minimal n-state DFA A from Corollary 11 accepts a suffix-closed language and is also a minimal NFA.

Proof. The second statement is proven as in Theorem 10, and for the case $\alpha = 2^n$ we refer to [2]. Thus, let $n < \alpha < 2^n$ and $A_{n,\alpha} = (Q, \Sigma, \delta, q_0, \{k\})$ be the JJS-automaton. Define $B_{n,\alpha} = (Q, \Sigma \cup \{\#\}, \delta_1, q_0, Q)$ where the transition function δ_1 is the same as δ except for the additional transition $\delta_1(k, \#) = k$. Since all states of $B_{n,\alpha}$ are accepting, $L(B_{n,\alpha})$ is prefix-closed by Theorem 9. In the powerset automaton $B'_{n,\alpha}$, there are exactly α reachable states, since the only new transition does not affect the reachability. Minimality of $B_{n,\alpha}$ and $B'_{n,\alpha}$ is shown by adding the suffix # to every word used in the proof of minimality of $A_{n,\alpha}$ and $A'_{n,\alpha}$. □

3.4. *Finite languages*

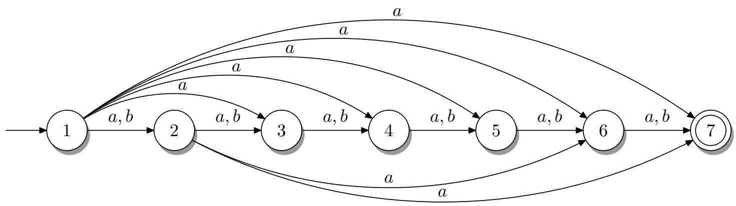
For finite languages the magic number problem turns out to be more challenging which seems to coincide with the fact, that the upper bounds for the deterministic blow-up of finite languages differ much from these of infinite language families. In [28] it was shown that for each n-state NFA over an alphabet of size k, there is an equivalent DFA with at most $O(k^{n/(\log(k)+1)})$ states. This matches an earlier result of $O(2^{n/2})$ for finite languages over binary alphabets [22].

In this section we give some partial results for finite languages over a binary alphabet, that is, we show that a roughly quadratic interval beginning at $n+1$ contains only non-magic numbers and that numbers of some exponential form $2^{(n-1)/2} + 2^i$ are non-magic, too. Note that for finite languages, n is a trivial magic number, since any DFA needs a non-accepting sink state which is not necessary for an NFA.

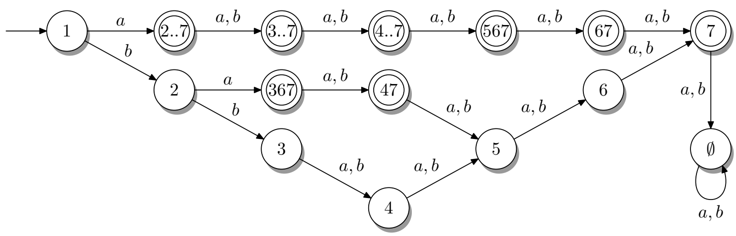
Theorem 13. *For all integers n and α such that $n + 1 \leq \alpha \leq (\frac{n}{2})^2 + \frac{n}{2} + 1$ if n is even, and $n + 1 \leq \alpha \leq (\frac{n-1}{2})^2 + n + 1$ if n is odd, there exists an n-state nondeterministic finite automaton accepting a* finite *language over a binary alphabet whose equivalent minimal deterministic finite automaton has exactly α states.*

Proof. The case $\alpha = n + 1$ can be seen with the witness language $\{a, b\}^{n-1}$, so let $\alpha > n + 1$. Then there is an integer k with $0 \leq k \leq \lfloor \frac{n}{2} \rfloor - 1$ such that

$$1 + \sum_{i=0}^{k}(n - 2i) < \alpha \leq 1 + \sum_{i=0}^{k+1}(n - 2i),$$



(a) A 7-state NFA accepting a finite language.



(b) The equivalent DFA with 15 states.

Fig. 3. An example for the construction in the proof of Theorem 13 with $n = 7$ and $\alpha = 15$. We have $k = 1$ and $m = 2$, so on input a, the NFA goes from state 1 to $\{2, 3, \ldots, 7\}$ and from state 2 to $\{3, 6, 7\}$. For all other states, the NFA goes from state q to state $q + 1$—except for $q = 7$.

and so for some integer m with $1 \leq m \leq n - 2(k + 1)$ we have

$$\alpha = 1 + \sum_{i=0}^{k}(n - 2i) + m.$$

Let $A = (\{1, 2, \ldots, n\}, \{a, b\}, \delta, 1, \{n\})$ be an NFA with transition function δ as follows—in order to simplify our presentation we use the notation $[i]$ for the state set $\{i, i+1, \ldots, n\}$: state n goes to the empty set on both inputs and all other states i go to $\{i + 1\}$ on input b. Also on input a, state i goes to $\{i + 1\}$ if $k + 1 < i < n$. For input a, state $k + 1$ goes to $T = \{k + 2\} \cup [n - m + 1]$, and if $i \leq k$, then state i goes to $S_i = \{i + 1\} \cup [2i + 1]$. Note that the transitions on input b and the sole accepting state n ensure minimality of A and its powerset automaton A'. For an example consider Figure 3.

For counting all reachable states in A', let $S_{i,j} = \delta'(S_i, b^j)$ and $T_j = \delta'(T, b^j)$, where δ' denotes the transition function of A'. We will now show that in A', we reach exactly the following α states: $\emptyset$, $\{h\}$ with $1 \leq h \leq n$, $S_{i,j}$ with $1 \leq i \leq k$ and $0 \leq j \leq n - 2i - 1$, and T_ℓ with $0 \leq \ell \leq m - 1$. Note that for these ranges of i, j, and ℓ, the sets $S_{i,j} = \{i + j + 1\} \cup [2i + j + 1]$ and $T_\ell = \{k + \ell + 2\} \cup [n - m + 1 + \ell]$ have a cardinality of at least two, because both $n - m + 1 + \ell$ and $2i + 1 + j$ are at most n. Further, the distance between the two smallest elements of $S_{i,j}$ is $i \leq k$ and for T_ℓ, this is $n - m - k - 1 \geq k + 1$. So all these sets are pairwise distinct and we have not multiply counted any state. The fact that all these states are reachable is seen as follows: the singletons and $\emptyset$ are reachable with words from b^* and the sets $S_{i,j}$ and T_ℓ are reachable by definition.

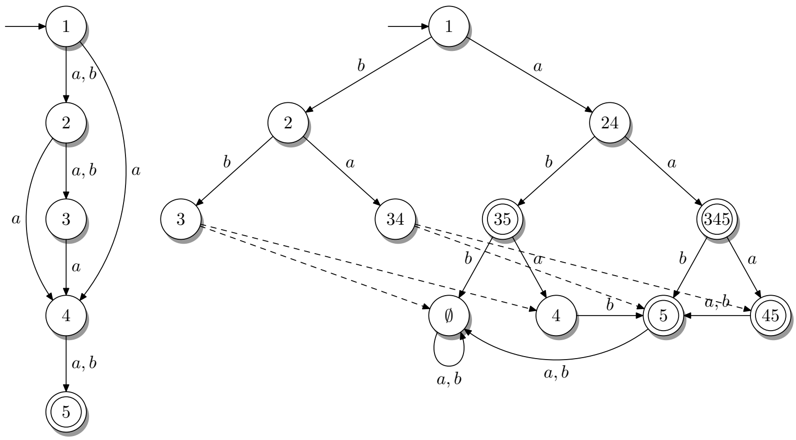


Fig. 4. The NFA A_n from [22] and its powerset automaton that builds a binary tree. In the DFA on the right the transitions of states $\{3\}$ and $\{3, 4\}$ are the same as for $\{3, 5\}$ and $\{3, 4, 5\}$, respectively.

To conclude the proof, we need to show that no further states can be reached. On input b, note that a singleton state goes to another singleton or to the empty set. This also holds for input a in states $\{i\}$ with $i > k+1$. For the other sets note that all elements of T_ℓ are greater than $k+1$, so we have $\delta'(T_\ell, a) = \delta'(T_\ell, b) = T_{\ell+1}$, or T_ℓ goes to a singleton. Next we show that also states $S_{i,j}$ map to the same state on both inputs. If $i+1+j > k+1$, then $\delta(S_{i,j}, a) = \delta(S_{i,j}, b)$, since all elements in $S_{i,j}$ are greater than $k+1$. If $i+1+j < k+1$, then we have

$$\delta(S_{i,j}, a) = \{i+j+2\} \cup [2i+j+2] \cup [2(i+j+1)+1] = S_{i,j+1} = \delta(S_{i,j}, b),$$

because $[2i+j+2] \supseteq [2(i+j+1)+1]$. Finally, for $i+1+j = k+1$ we have

$$\delta(S_{i,j}, a) = \{i+j+2\} \cup [2i+j+2] \cup [n-m+1] = S_{i,j+1} = \delta(S_{i,j}, b),$$

because $[2i+j+2] \supseteq [n-m+1]$, for $m \leq n-2(k+1)$ and $i+j=k$. □

For our last theorem we use the following results presented in [22]: for an integer n let $k = \lceil \frac{n}{2} \rceil$ and $A_n = (\{1, 2, \ldots, n\}, \{a, b\}, \delta, 1, \{n\})$ be an NFA with transitions $\delta(q, a) = \{q+1, k+1\}$ if $q < k$, $\delta(q, a) = \{q+1\}$ if $k \leq q < n$, and $\delta(q, b) = \{q+1\}$ if $q < n$ and $q \neq k$. Then in [22] it is shown that A_n is minimal and that the minimal equivalent DFA has exactly $2^{(n/2)+1} - 1$ states if n is even, and $3 \cdot 2^{(n+1)/2-1} - 1$ states if n is odd. This minimal DFA spans a binary tree on inputs a and b as depicted in Figure 4.

Theorem 14. *For all integers n and α such that $\alpha = 3 \cdot 2^{(n/2)-1} + \beta$ if n is even and $\alpha = 2^{(n+1)/2} + \beta$ if n is odd, with $\beta = 2^i - 1$ for some integer $1 \leq i \leq \lfloor \frac{n-1}{2} \rfloor$, there exists an n-state nondeterministic finite automaton accepting a* finite *language* *over a binary alphabet whose equivalent minimal deterministic finite automaton has exactly α states.*

Proof. Let n, α and β be as required and $x = n - \log(\beta + 1)$. We construct an automaton $B_{n,\beta}$ adapting $A_{n-1} = (\{1, 2, \ldots, n-1\}, \{a, b\}, \delta_1, 1, \{n-1\})$ from above with $k = \lceil \frac{n-1}{2} \rceil$, by taking a new initial state 0 and setting the transition function δ to $\delta(0, b) = \{1\}$, $\delta(0, a) = \{1, x\}$, and $\delta(q, c) = \delta_1(q, c)$, for $1 \le q \le n - 1$ and $c \in \{a, b\}$. The minimality of $B_{n,\beta}$ can be shown with the fooling set

$$S = \{\, (b^i, b^{k-i}ab^{n-k-2}) \mid 0 \le i \le k \,\} \cup \{\, (b^k ab^i, b^{n-k-2-i}) \mid 0 \le i \le n - k - 2 \,\}.$$

Let A'_{n-1} and $B'_{n,\beta}$ be the powerset automata of A_{n-1} and $B_{n,\beta}$. Then, by reading words bw for $w \in \{a, b\}^*$, all states of A'_{n-1} are reachable in $B'_{n,\beta}$. Together with the initial state $\{0\}$, these are $2^{(n-1)/2+1}$ states if n is odd, and $3 \cdot 2^{n/2-1}$ states if n is even. For words of the form aw, for $w \in \{a, b\}^*$, let $\ell = \lfloor \frac{n-1}{2} \rfloor$. From $1 \le \beta \le 2^\ell - 1$ it follows that $k + 1 \le x \le n - 1$ and we reach the states $\delta'(\{0\}, aw) = \delta'(\{1, x\}, w) = \delta'(\{1\}, w) \cup \delta'(\{x\}, w)$. Note that $\delta'(\{x\}, w)$ is $\emptyset$ if $|w| > n-1-x$ and $\{x+|w|\}$ otherwise. Further $\max(\delta(\{1\}, w)) = k+|w| < x+|w|$, so the states $\delta'(\{0\}, aw)$ differ from the ones in A'_{n-1} as long as $\delta'(\{x\}, w) \neq \emptyset$, and this holds if and only if $|w| \le n - 1 - x$. These are $2^{n-x} - 1 = \beta$ additional states.

The minimality of $B'_{n,\beta}$ can be seen as follows. Let M and N be distinct subsets of $\{0, 1, \ldots, n-1\}$. If M and N differ in an element $q \ge k+1$, then the word b^{n-1-q} distinguishes both sets. Otherwise, note that both M and N contain at most one element from the set $\{0, 1, \ldots, k\}$. So we may assume without loss of generality, that $p = \min(M) < \min(N)$. Then the word $b^{k-p}ab^{n-2-k}$ is accepted from state M but not from N. □

References

[1] J. C. Birget. Intersection and union of regular languages and state complexity. *Inform. Process. Lett.*, 43:185–190, 1992.

[2] H. Bordihn, M. Holzer, and M. Kutrib. Determinization of finite automata accepting subregular languages. *Theoret. Comput. Sci.*, 410:3209–3222, 2009.

[3] J. A. Brzozowski. Roots of star events. *J. ACM*, 14:466–477, 1967.

[4] J. A. Brzozowski and R. Cohen. On decompositions of regular events. In *Symposium on Switching and Automata Theory (SWAT 1967)*, pages 255–264, IEEE Computer Society Press, 1967.

[5] V. Geffert. (Non)determinism and the size of one-way finite automata. In *Descriptional Complexity of Formal Systems (DCFS 2005)*, pages 23–37, Universita degli Studi di Milano, 2005.

[6] V. Geffert. Magic numbers in the state hierarchy of finite automata. *Inform. Comput.*, 205:1652–1670, 2007.

[7] I. Glaister and J. Shallit. A lower bound technique for the size of nondeterministic finite automata. *Inform. Process. Lett.*, 59:75–77, 1996.

[8] Y. S. Han and K. Salomaa. State complexity of basic operations on suffix-free regular languages. *Theoret. Comput. Sci.*, 410:2537–2548, 2009.

[9] Y. S. Han, K. Salomaa, and D. Wood. Operational State complexity of prefix-free regular languages. In *Automata, Formal Languages, and Related Topics*, pages 99-115, Institute of Informatics, University of Szeged, Hungary, 2009.
[10] M. Holzer and M. Kutrib. Descriptional and computational complexity of finite automata. In *Language and Automata Theory and Applications (LATA 2009)*, number 5457 in LNCS, pages 23–42, Springer, 2009.
[11] M. Holzer and M. Kutrib. Nondeterministic finite automata—recent results on the descriptional and computational complexity. *Internat. J. Found. Comput. Sci.*, 20: 563–580, 2009.
[12] M. Holzer and M. Kutrib. Descriptional complexity—An introductory survey. In *Scientific Applications of Language Methods*, Imperial College Press, 2010.
[13] K. Iwama, Y. Kambayashi, and K. Takaki. Tight bounds on the number of states of DFAs that are equivalent to n-state NFAs. *Theoret. Comput. Sci.*, 237:485–494, 2000.
[14] K. Iwama, A. Matsuura, and M. Paterson. A family of NFAs which need $2^n - \alpha$ deterministic states. *Theoret. Comput. Sci.*, 301:451–462, 2003.
[15] J. Jirásek, G. Jirásková, and A. Szabari. Deterministic blow-ups of minimal nondeterministic finite automata over a fixed alphabet. *Internat. J. Found. Comput. Sci.*, 19:617–632, 2008.
[16] G. Jirásková. Deterministic blow-ups of minimal NFAs. *RAIRO Theoret. Informatics Appl.*, 40:485–499, 2006.
[17] G. Jirásková. On the state complexity of complements, stars, and reversal of regular languages. In *Developments in Language Theory (DLT 2008)*, number 5257 in LNCS, pages 431–442, Springer, 2008.
[18] G. Jirásková. Concatenation of regular languages and descriptional complexity. In *Computer Science Symposium in Russia (CSR 2009)*, number 5675 in LNCS, pages 203–214, Springer, 2009.
[19] G. Jirásková. Magic numbers and ternary alphabet. *Internat. J. Found. Comput. Sci.*, 22:331–344, 2011.
[20] J. Y. Kao, N. Rampersad, and J. Shallit. On NFAs where all states are final, initial, or both. *Theoret. Comput. Sci.*, 410:5010–5021, 2009.
[21] O. B. Lupanov. A comparison of two types of finite sources. *Problemy Kybernetiki*, 9:321–326, 1963 (in Russian). German translation: Über den Vergleich zweier Typen endlicher Quellen, Probleme der Kybernetik 6:328–335, 1966.
[22] R. Mandl. Precise bounds associated with the subset construction on various classes of nondeterministic finite automata. In *Conference on Information and System Sciences*, pages 263–267, 1973.
[23] A. Matsuura and Y. Saito. Equivalent transformation of minimal finite automata over a two-letter alphabet. In *Descriptional Complexity of Formal Systems (DCFS 2008)*, pages 224–232, University of Prince Edward Island, 2008.
[24] R. McNaughton and S. Papert. *Counter-free automata.* Number 65 in Research monographs, MIT Press, 1971.
[25] A. R. Meyer and M. J. Fischer. Economy of description by automata, grammars, and formal systems. In *Symposium on Switching and Automata Theory (SWAT 1971)*, pages 188–191, IEEE Computer Society Press, 1971.
[26] F. R. Moore. On the bounds for state-set size in the proofs of equivalence between deterministic, nondeterministic, and two-way finite automata. *IEEE Trans. Comput.*, 20:1211–1214, 1971.
[27] M. O. Rabin and D. Scott. Finite automata and their decision problems. *IBM J. Res. Dev.*, 3:114–125, 1959.

[28] K. Salomaa and S. Yu. NFA to DFA transformation for finite language over arbitrary alphabets. *J. Autom., Lang. Comb.*, 2:177–186, 1997.

[29] H.-J. Shyr and G. Thierrin. Power-separating regular languages. *Math. Systems Theory*, 8:90–95, 1974.