CONCATENATION ON DETERMINISTIC AND ALTERNATING AUTOMATA

Michal Hospodár Galina Jirásková

Mathematical Institute, Slovak Academy of Sciences,
Grešákova 6, 040 01 Košice, Slovakia
hosmich@gmail.com jiraskov@saske.sk

Abstract
We study the state complexity of the concatenation operation on regular languages represented by deterministic and alternating finite automata. For deterministic automata, we show that the upper bound $m2^n - k2^{n-1}$ on the state complexity of concatenation can be met by ternary languages, the first of which is accepted by an m-state DFA with k final states, and the second one by an n-state DFA with ℓ final states for arbitrary integers m, n, k, ℓ. In the case of $k \leq m - 2$, that is, in the case when the first automaton has at least two non-final states, we are able to provide appropriate binary witnesses. We use these witnesses to describe a pair of binary languages meeting the upper bound $2^m + n + 1$ for the concatenation on alternating finite automata. This solves an open problem stated by Fellah, Jürgensen, and Yu [1990, Constructions for alternating finite automata, Intern. J. Computer Math. 35, 117–132], where the upper bound $2^m + n + 1$ is proved.

1. Introduction

Concatenation is a binary operation on formal languages defined as $KL = \{uv \mid u \in K$ and $v \in L\}$. It is known that if a language K is accepted by an m-state deterministic finite automaton (DFA) and L is accepted by an n-state DFA, then the concatenation KL is accepted by a DFA of at most $m2^n - 2^{n-1}$ states [9]. Ternary languages meeting this upper bound were described by Yu, Zhuang, and K. Salomaa [13]. Maslov [9] proposed binary witnesses for concatenation, but he did not provide any proof. The tightness of this upper bound in the binary case was proved in [5].

However, if the minimal DFA recognizing the first language has more that one final state, then the upper bound $m2^n - 2^{n-1}$ on the state complexity of concatenation cannot be met; here, the state complexity of a regular language is the number of states in the minimal DFA for the language, and the state complexity of a regular operation is the number of states that are sufficient and necessary in the worst case for a DFA to recognize the language resulting from the operation. Yu et al. [13] showed that the state complexity of concatenation is at most

Research supported by VEGA grant 2/0084/15 and grant APVV-15-0091.

$m2^n - k2^{n-1}$, where k is the number of final states in the minimal DFA for the first language. The binary languages meeting this upper bound were described for each k with $1 \le k \le m-1$ in [4, Theorem 1], but there are some errors in the proof of this theorem, and one of our aims is to fix these errors. We also show that the witnesses from [9] and [13] meet the upper bound $m2^n - k2^{n-1}$ if we make the k last states final in the DFA for the first language.

Then we study the complexity of concatenation also in the case where the second automaton has more than one final state. Our motivation comes from [3], where the authors consider the concatenation operation on languages represented by alternating finite automata (AFA), and get an upper bound $2^m + n + 1$. They also write:

> *"We conjecture that this number of states is actually necessary in the worst case, but have no proof."*

It is known [3, Theorem 4.1, Corollary 4.2] and [6, Lemma 1, Lemma 2] that a language L is accepted by an n-state AFA if and only if its reversal L^R is accepted by a 2^n-state DFA with 2^{n-1} states final. Hence to get a lower bound for concatenation on AFAs, we need two languages represented by DFAs with half of states final that are hard for concatenation on DFAs.

We first inspect the witnesses from [4, 9, 13] and show that none of them meets the upper bound $m2^n - k2^{n-1}$ if the second automaton has more then one final state. Then we describe ternary languages meeting this bound for all m, n, k, ℓ, where m and k is the number of states and the number of final states in the minimal DFA for the first language, and n and ℓ is the number of states and the number of final states in the minimal DFA for the second language. Finally, in the case of $k \le m-2$, that is, if the first automaton has at least two non-final states, we describe appropriate binary languages. We use these witnesses to define, for every $m, n \ge 2$, binary languages K and L accepted by an m-state and n-state AFAs, respectively, such that the minimal AFA for KL requires $2^m + n + 1$ states. This proves that the upper bound $2^m + n + 1$ from [3] is tight, and solves the open problem stated in [3, Theorem 9.3].

2. Preliminaries

In this section, we give some basic definitions and preliminary results. For details and all unexplained notions, the reader may refer to [11, 12].

Let Σ be a finite alphabet of symbols. Then Σ^* denotes the set of strings over Σ including the empty string ε. A language is any subset of Σ^*. The concatenation of languages K and L is the language $KL = \{uv \mid u \in K \text{ and } v \in L\}$. The cardinality of a finite set A is denoted by $|A|$, and its power-set by 2^A. We define an operator $\ominus$ as follows: If $i, j \in \{0, 1, \dots, n-1\}$, then $j \ominus i = (j - i) \bmod n$, and if $S \subseteq \{0, 1, \dots, n-1\}$, then $S \ominus i = \{j \ominus i \mid j \in S\}$.

A *nondeterministic finite automaton* (NFA) is a quintuple $A = (Q, \Sigma, \cdot, I, F)$, where Q is a finite set of states, Σ is a finite alphabet, $\cdot : Q \times \Sigma \to 2^Q$ is the transition function which is extended to the domain $2^Q \times \Sigma^*$ in the natural way, $I \subseteq Q$ is the set of initial states, and $F \subseteq Q$ is the set of final states. The *language accepted by* A is the set $L(A) = \{w \in \Sigma^* \mid I \cdot w \cap F \neq \emptyset\}$.

For a symbol a, we say that (p, a, q) is a transition in NFA A if $q \in p \cdot a$; we also say that (p, a, q) is an *in-transition* on symbol a going to the state q. Next, we say that (p, a, q) is a unique in-transition on a going to q if there is no state r with $r \neq p$ such that (r, a, q) is a transition in A. We call a state q is *uniquely distinguishable* if there exists a string w which is accepted by A form and only from q, that is, we have $p \cdot w \cap F \neq \emptyset$ iff $p = q$. For a string w, we write $p \xrightarrow{w} q$ if $q \in p \cdot w$.

An NFA A is *deterministic* (DFA) and complete if $|I| = 1$ and $|q \cdot a| = 1$ for each q in Q and each a in Σ. In such a case, we write $q \cdot a = q'$ instead of $q \cdot a = \{q'\}$. The *state complexity* of a regular language L, $\mathrm{sc}(L)$, is the smallest number of states in any DFA for L.

The reversal L^R of a language L is defined as $L^R = \{w^R \mid w \in L\}$, where w^R is the mirror image of the string w. For every finite automaton $A = (Q, \Sigma, \cdot, I, F)$ we can construct the automaton $A^R = (Q, \Sigma, \cdot', F, I)$ where the function $\cdot'$ is defined as follows: $p \in q \cdot' a$ iff $q \in p \cdot a$ for every p, q in Q and every a in Σ. Then it holds that $L(A^R) = (L(A))^R$.

Every NFA $A = (Q, \Sigma, \cdot, I, F)$ can be converted into an equivalent DFA $A' = (2^Q, \Sigma, \cdot', I, F')$, where for every reachable set $R \subseteq Q$ it holds $R \cdot' a = R \cdot a$ and $F' = \{R \in 2^Q \mid R \cap F \neq \emptyset\}$ [10]. The DFA A' is called the *subset automaton* of the NFA A. The subset automaton may not be minimal since some of its states may be unreachable or equivalent to other states. In the following proposition, we provide a sufficient condition for an NFA, which guarantees that the corresponding subset automaton does not have equivalent states.

Proposition 2.1 *Let $N = (Q, \Sigma, \cdot, I, F)$ be an NFA. Assume that for each state q in Q, there is a string w_q in Σ^* which is accepted by N only from the state q, that is, we have $(q \cdot w_q) \cap F \neq \emptyset$, and $(p \cdot w_q) \cap F = \emptyset$ if $p \neq q$. Then the subset automaton of N does not have equivalent states.*

Proof. Two distinct subsets of the subset automaton must differ in a state q, and the string w_q distinguishes the two subsets. □

To describe a string w_q accepted by an NFA only from state q, we use the next observation.

Proposition 2.2 *Let a string w_q be accepted by an NFA N only from state q. If (p, a, q) is the unique in-transition going to state q by symbol a, then the string aw_q is accepted by N only from state p.*

3. Construction of NFA for Concatenation

Let K and L be accepted by minimal DFAs A and B, respectively. Without loss of generality, we may assume that the state set of A is $\{q_0, q_1, \ldots, q_{m-1}\}$ with the initial state q_0, and the state set of B is $\{0, 1, \ldots, n-1\}$ with the initial state 0. Moreover, we denote the transition function in both A and B by $\cdot$; there is no room for confusion since A and B have distinct state sets. We first recall the construction of an NFA for the concatenation of languages K and L.

Construction 3.1 (DFA A and DFA $B \rightarrow$ NFA N for $L(A)L(B)$)
Let $A = (\{q_0, q_1, \ldots, q_{m-1}\}, \Sigma, \cdot, q_0, F_A)$ and $B = (\{0, 1, \ldots, n-1\}, \Sigma, \cdot, 0, F_B)$ be DFAs. We

construct NFA $N = (\{q_0, q_1, \dots, q_{m-1}\} \cup \{0, 1, \dots, n-1\}, \Sigma, \cdot, I, F_B)$ *from DFAs* A *and* B *as follows:*

(a) *for each* a *in* Σ *and each state* q_i *of* A*, if* $q_i \cdot a \in F_A$*, then add the transition* $(q_i, a, 0)$*;*
(b) *the set* I *of initial states of* N *is* $\{q_0\}$ *if* $q_0 \notin F_A$*, and it is* $\{q_0, 0\}$ *otherwise;*
(c) *the set of final state of* N *is* F_B*.*

Using the construction described above, we get an upper bound on the state complexity of concatenation. Notice that the bound depends on the number of final states in the minimal DFA for the first language.

Proposition 3.2 (Concatenation: Upper Bound if $|F_A| = k$**)** *Let* K *and* L *be regular languages with* $\mathrm{sc}(K) = m$ *and* $\mathrm{sc}(L) = n$*, and let the minimal DFA for* K *have* k *final states. Then* $\mathrm{sc}(KL) \le (m-k)2^n + k2^{n-1}$*.*

Proof. Let the languages K and L be accepted by DFAs $A = (\{q_0, q_1, \dots, q_{m-1}\}, \Sigma, \cdot, q_0, F_A)$ and $B = (\{0, 1, \dots, n-1\}, \Sigma, \cdot, 0, F_B)$, respectively. Let $|F_A| = k$. Construct an NFA N for KL as in the construction above, and consider the corresponding subset automaton D. Since A is deterministic and complete, each reachable subset in D is of the form $\{q_i\} \cup S$, where $S \subseteq \{0, 1, \dots, n-1\}$. Moreover, if q_i is a final state of A, then $0 \in S$ since the NFA N has the transition $(q, a, 0)$ whenever a state q of A goes to a final state q_i on a symbol a. It follows that the subset automaton of N has at most $(m-k)2^n + k2^{n-1}$ reachable states. □

Notice that $(m-k)2^n + k2^{n-1} = m2^n - k2^{n-1}$, which is maximal if $k = 1$. We prefer to write the upper bound as $(m-1)2^n + 2^{n-1}$; cf. [9, 13].

Corollary 3.3 (Concatenation: Upper Bound) *Let* K *and* L *be regular languages with* $\mathrm{sc}(K) = m$ *and* $\mathrm{sc}(L) = n$*. Then* $\mathrm{sc}(KL) \le (m-1)2^n + 2^{n-1}$*.*

4. Ternary and Binary Witness Languages

Motivated by the open problem from [3] concerning the tightness of the upper bound $2^m + n + 1$ for concatenation on alternating automata, we study the state complexity of the concatenation of languages represented by deterministic finite automata that have more than one final state. Let us start with the following observation.

Observation 4.1 *Let* $m \ge 1$*,* $k \le m$*. Let* A *be an* m*-state DFA with* k *final states and* B *be a* 1*-state DFA, both over an alphabet* Σ*. Then* $\mathrm{sc}(L(A)L(B)) \le m - k + 1$*, and the bound is tight if* $|\Sigma| \ge 1$*.*

Proof. If a complete DFA B has one state, then either $L(B) = \emptyset$ or $L(B) = \Sigma^*$. Since for every language L holds $L\emptyset = \emptyset$, and hence $\mathrm{sc}(L\emptyset) = 1$, we assume that $L(B) = \Sigma^*$. We construct the DFA for $L(A)L(B)$ from A as follows: for every final state p and every a in Σ, we replace the transition (p, a, q) by the transition (p, a, p). The resulting automaton is deterministic and complete, has m states and k final states. All the final states are equivalent since every string

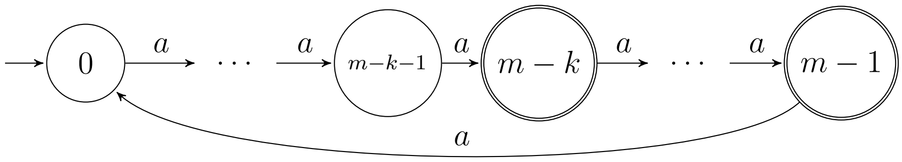

Figure 1: Unary automaton A with $\mathrm{sc}(L(A)\Sigma^*) = m - k + 1$.

is accepted from any of them. It follows that we can merge all final states into a single final state. This gives the upper bound $m - k + 1$. To prove tightness, consider the unary DFA shown in Figure 1. For each final state p, we remove all the transitions going from p, and add the transition (p, a, p). Then we merge all final states into a single final state. The resulting automaton accepts the language $a^{m-k}a^*$ and has $m - k + 1$ states. Since the DFA in Figure 1 meets the upper bound, it must be minimal. □

Next we investigate the automata with $\mathrm{sc}(L(B)) \geq 2$. We inspect three worst-case examples from the literature, and change them by making some states in the first automaton final. To simplify the proofs, we use the property of all these witnesses that the letter a makes the permutation $q_i \cdot a = q_{(i+1) \bmod m}$ in A and a permutation in B. If these two conditions are satisfied, then we get the following observation.

Lemma 4.2 *Consider DFAs* $A = (\{q_0, q_1, \ldots, q_{m-1}\}, \Sigma, \cdot, q_0, \{q_{m-k}, q_{m-k+1}, \ldots, q_{m-1}\})$ *and* $B = (Q_B, \Sigma, \cdot, 0, F_B)$*, where* $Q_B = \{0, 1, \ldots, n-1\}$*. Assume that there is a symbol* a *in* Σ *such that* $q_i \cdot a = q_{(i+1) \bmod m}$ *and* a *performs a permutation on* Q_B*. Let* N *be an NFA for* $L(A)L(B)$ *from the Construction 3.1. Then in the subset automaton of* N*, we have*

(1) *for each* $S \subseteq Q_B$ *with* $0 \in S$*, the set* $\{q_{m-k}\} \cup S$ *is reachable from a set* $\{q_{m-k-1}\} \cup S'$*, where* $S' \subseteq Q_B$ *and* $|S'| = |S| - 1$*;*
(2) *for each* $S \subseteq Q_B$ *and each* $i = 1, 2, \ldots, m-k-1$*, the set* $\{q_i\} \cup S$ *is reachable from a set* $\{q_0\} \cup S'$*, where* $S' \subseteq Q_B$ *and* $|S'| = |S|$*;*
(3) *moreover, if* $0 \cdot a = 0$*, then for each* $S \subseteq Q_B$ *with* $0 \in S$ *and for each* $i = 0, 1, \ldots, m-1$*, the set* $\{q_i\} \cup S$ *is reachable from a set* $\{q_{m-k-1}\} \cup S'$*, where* $S' \subseteq Q_B$ *and* $|S'| = |S| - 1$*.*

Proof. Since a is a permutation on Q_B, we can use $q \cdot a^{-1}$ to denote the state p with $p \cdot a = q$. Next, we can extend a^{-1} to subsets of Q_B and to a^{-i}.

(1) Let $S' = (S \setminus \{0\}) \cdot a^{-1}$. Then $|S'| = |S| - 1$ and the set $\{q_{m-k}\} \cup S$ is reached from $\{q_{m-k-1}\} \cup S'$ by a.
(2) Let $S' = S \cdot a^{-i}$. Then $|S'| = |S|$ and the set $\{q_i\} \cup S$ is reached from $\{q_0\} \cup S'$ by a^i.
(3) Let $S' = (S \setminus \{0\}) \cdot a^{-(k+1+i)}$. Then $|S'| = |S| - 1$ and the set $\{q_i\} \cup S$ is reached from $\{q_{m-k-1}\} \cup S'$ by a^{k+1+i} since $0 \cdot a = 0$.

□

Ternary witness languages meeting the upper bound for concatenationare described in [13, Theorem 2.1]. We generalize these languages to get more final states in the first automaton. Then we provide a proof that if the minimal DFA for the first language has k final states, then the state complexity of the resulting concatenation meets the upper bound $(m - k)2^n + k2^{n-1}$.

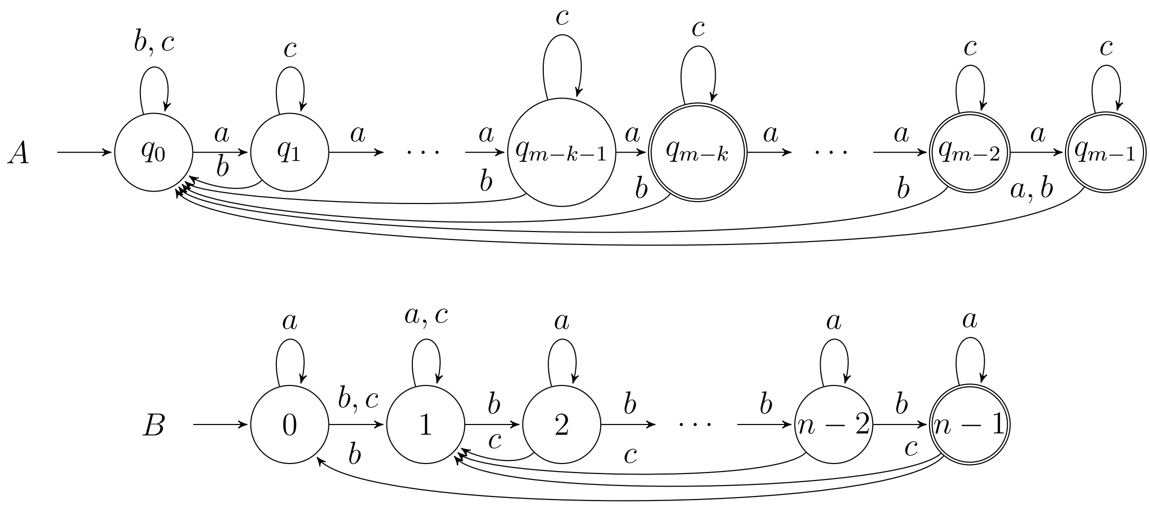

Figure 2: Ternary witnesses for concatenation meeting the bound $(m-k)2^n + k2^{n-1}$; $m, n \geq 2$ and $1 \leq k \leq m-1$.

Lemma 4.3 (Ternary Witness Languages with $|F_A| = k$ and $|F_B| = 1$) *Let $m, n \geq 2$, $1 \leq k \leq m-1$. Let K and L be the ternary languages accepted by DFAs A and B, respectively, shown in Figure 2. Then* $\mathrm{sc}(KL) = (m-k)2^n + k2^{n-1}$.

Proof. Construct an NFA N for KL from DFAs A and B by adding transitions $(q_{i-1}, a, 0)$ and $(q_i, c, 0)$ for $m-k \leq i \leq m-1$ as shown in Figure 3; the initial state of N is q_0, and the set of final states is $\{n-1\}$. Let $\mathcal{R}$ be the following family of $(m-k)2^n + k2^{n-1}$ subsets of states:

$$\begin{aligned}\mathcal{R} = &\big\{\{q_i\} \cup S \mid 0 \leq i \leq m-k-1 \text{ and } S \subseteq Q_B\big\} \cup \\ &\big\{\{q_i\} \cup S \mid m-k \leq i \leq m-1,\ S \subseteq Q_B \text{ and } 0 \in S\big\}.\end{aligned}$$

To prove the lemma, we only need to show that each subset in $\mathcal{R}$ is reachable in the subset automaton of N, and that all these subsets are pairwise distinguishable.

We first prove reachability. The proof is by induction on $|\{q_i\} \cup S|$. The basis, $|\{q_i\} \cup S| = 1$, holds true since $\{q_0\}$ is the initial state and it goes to the subset $\{q_i\}$ by a^i if $1 \leq i \leq m-k-1$. Let $1 \leq t \leq n$, and assume that each subset in $\mathcal{R}$ of size t is reachable. By Lemma 4.2, each set $\{q_i\} \cup S$ of size $t+1$, where $m-k \leq i \leq m-1$ and $S \subseteq Q_B$ with $0 \in S$, can be reached from a set of size t since the symbol a performs the permutation $q_i \cdot a = q_{(i+1) \bmod m}$ on states of A and a permutation on states of B and moreover $0 \cdot a = 0$. Next, by Lemma 4.2, each set $\{q_i\} \cup S$ of size $t+1$ where $1 \leq i \leq m-k-1$ is reached from a set $\{q_0\} \cup S'$ of size $t+1$. Hence it is enough to show the reachability of sets $\{q_0\} \cup S$ for every $S \subseteq Q_B$ such that $|\{q_0\} \cup S| = t+1$. We have

$$\{q_{m-1}\} \cup (S \ominus \min S) \xrightarrow{a} \{q_0\} \cup (S \ominus \min S) \xrightarrow{b^{\min S}} \{q_0\} \cup S,$$

where $0 \in S \ominus \min S$ and the set $\{q_{m-1}\} \cup (S \ominus \min S)$ can be reached from a set of size t by Lemma 4.2. This proves reachability. To prove distinguishability, let $\{q_i\} \cup S$ and $\{q_j\} \cup T$ be two distinct subsets in $\mathcal{R}$. Notice that the state $n-1$ is uniquely distinguishable in NFA N since it is a unique final state. Next, the state $n-1$ is reached from each state of Q_B in the

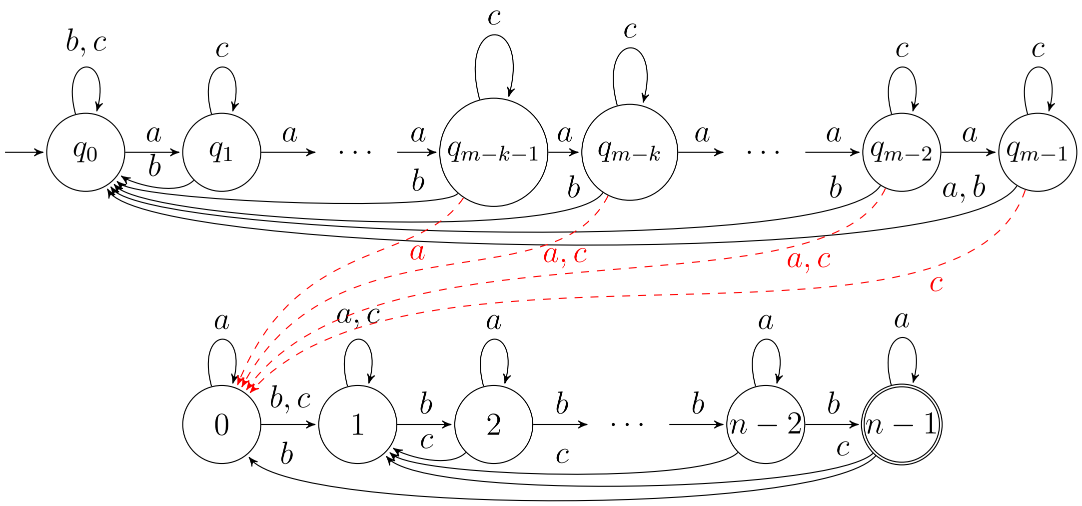

Figure 3: An NFA N for KL, where K and L are accepted by DFAs A and B from Figure 2.

subgraph of unique in-transitions $(j, b, j+1)$ where $0 \leq j \leq n-2$. It follows that each state in Q_B is uniquely distinguishable. By Proposition 2.1, if $S \neq T$, then $\{q_i\} \cup S$ and $\{q_j\} \cup T$ are distinguishable. Now let $S = T$. Then $i \neq j$, and without loss of generality, $i < j$. There are three cases:

(1) Let $i < m-k \leq j$, that is, q_i is non-final and q_j is final in A. Then $0 \notin (\{q_i\} \cup S) \cdot c$, but $0 \in (\{q_j\} \cup S) \cdot c$, so the sets differ in the state 0 and are distinguishable as shown above.
(2) Let $m-k \leq i < j$, that is, both q_i and q_j are final in A. Then we read a^{m-j} and get the sets $\{q_0\} \cup S$ and $\{q_{m-j+i}\} \cup S$ which are considered in case (1).
(3) Let $i < j < m-k$, that is, both q_i and q_j are non-final in A. Then we read a^{m-k-j} and get the sets $\{q_{m-k-j+i}\} \cup S$ and $\{q_{m-k}\} \cup \{0\} \cup S$ which either differ in the state 0 or are considered in case (1).

This proves distinguishability and concludes our proof. □

Yu et al. [13] left the binary case open. Later, a paper by Maslov [9] was found, in which the author describes binary witnesses. However, he does not provide any proof. Let us show that his witnesses, generalized to have k final states in A, shown in Figure 4, work for $n \geq 3$.

Lemma 4.4 (Binary Witness Languages with $|F_A| = k$ and $|F_B| = 1$) *Let $m \geq 2$, $n \geq 3$, and $1 \leq k \leq m-1$. Let K and L be the binary languages accepted by DFAs A and B shown in Figure 4. Then* $\mathrm{sc}(KL) = (m-k)2^n + k2^{n-1}$.

Proof. Construct an NFA N for KL from DFAs A and B by adding transitions $(q_{i-1}, a, 0)$ and $(q_i, b, 0)$ for $m-k \leq i \leq m-1$, the initial state of N is q_0, and the set of final states is $\{n-1\}$. Let $\mathcal{R}$ be the same family of $(m-k)2^n + k2^{n-1}$ subsets as in the previous proof. We need to show that each set in $\mathcal{R}$ is reachable in the subset automaton. The proof is by induction on $|\{q_i\} \cup S|$. The basis, $|\{q_i\} \cup S| = 1$, holds true since $\{q_0\}$ is the initial state and it goes to the subset $\{q_i\}$ by a^i if $1 \leq i \leq m-k-1$. Let $1 \leq t \leq n$, and assume that each subset in $\mathcal{R}$ of size t is reachable. By Lemma 4.2, it is enough to show the reachability of sets $\{q_0\} \cup S$ for every

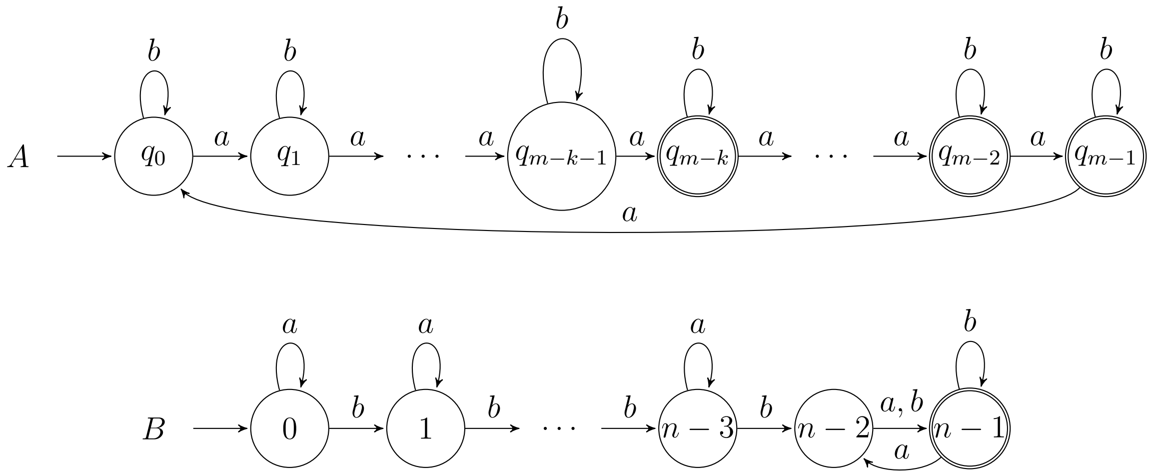

Figure 4: Binary witnesses for concatenation meeting the bound $(m-k)2^n + k2^{n-1}$; $m \geq 2, n \geq 3$, and $1 \leq k \leq m-1$.

$S \subseteq Q_B$ such that $|\{q_0\} \cup S| = t+1$. We have

$$\{q_{m-1}\} \cup (S \ominus \min S) \cdot a^{-1} \xrightarrow{a} \{q_0\} \cup (S \ominus \min S) \xrightarrow{b^{\min S}} \{q_0\} \cup S,$$

where $0 \in S \ominus \min S$ and the set $\{q_{m-1}\} \cup (S \ominus \min S) \cdot a^{-1}$ can be reached from a set of size t by Lemma 4.2. This proves reachability.

To prove distinguishability, let $\{q_i\} \cup S$ and $\{q_j\} \cup T$ be two distinct subsets in $\mathcal{R}$. Notice that the state $n-1$ is uniquely distinguishable since it is a unique final state in B. Next, the state $n-1$ is uniquely reachable by each state in Q_B in the subtree of unique in-transitions $0 \xrightarrow{b} 1 \xrightarrow{b} \ldots \xrightarrow{b} n-2 \xrightarrow{a} n-1$. It follows that each state in Q_B is uniquely distinguishable. By Proposition 2.1, if $S \neq T$, then $\{q_i\} \cup S$ and $\{q_j\} \cup T$ are distinguishable. Now let $S = T$. Then $i \neq j$, and without loss of generality, $i < j$. There are four cases:

(1) Let $i < m-k \leq j$, so $0 \in S$. Then we read b and get $\{q_i\} \cup (S \cdot b)$ and $\{q_j\} \cup \{0\} \cup (S \cdot b)$, which differ in state 0 since $0 \notin S \cdot b$.
(2) If $m-k \leq i < j$, then we read a^{m-j} and get $\{q_{m-j+i}\} \cup (S \cdot a^{m-j})$ and $\{q_0\} \cup (S \cdot a^{m-j})$, which are considered in case (1).
(3) If $i < j < m-k$ and $0 \in S$, then we read a^{m-k-j} and get $\{q_{m-k-j+i}\} \cup (S \cdot a^{m-k-j})$ and $\{q_{m-k}\} \cup (S \cdot a^{m-k-j})$, which are considered in case (1).
(4) If $i < j < m-k$ and $0 \notin S$, then we read a^{m-k-j} and get $\{q_{m-k-j+i}\} \cup (S \cdot a^{m-k-j})$ and $\{q_{m-k}\} \cup \{0\} \cup (S \cdot a^{m-k-j})$, which differ in state 0.

This proves distinguishability and concludes our proof. □

While the ternary witnesses from Lemma 4.3 require the complexities $m \geq 2$ and $n \geq 2$, for the binary case, the witnesses from Lemma 4.4 do not work if $n = 2$. In [4, Theorem 1], binary witnesses for $m \geq 1$ and $n \geq 2$ are described. However, the proof of [4, Theorem 1] does not work: For example, it is claimed that the set $\{q_{m-k-1}, j_2 - 1, \ldots, j_s - 1\}$ goes to $\{q_{m-k+1}, 0, j_2, \ldots, j_s\}$ by aab^{n-1}; cf. line -4 on page 515. In fact it goes to $\{q_{m-k+1}, 0\}$. Such an error occurs several times in the proof, namely, on line -2 on page 515, and on lines 2 and 8 on page 516. The authors overlooked that ab^{n-1} does not perform an identity on $\{0, 1, \dots, n-1\}$, but moves this set to $\{0\}$. Here we provide a correct proof.

Lemma 4.5 ([4], Binary Witness Languages with $|F_A| = k$ and $|F_B| = 1$) *Let $m \geq 1$, $n \geq 2$. Let $k = 1$ if $m = 1$, and $1 \leq k \leq m-1$ otherwise. There exist binary regular languages K and L with $\mathrm{sc}(K) = m$, $\mathrm{sc}(L) = n$ such that the minimal DFA for K has k final states and $\mathrm{sc}(KL) = (m-k)2^n + k2^{n-1}$.*

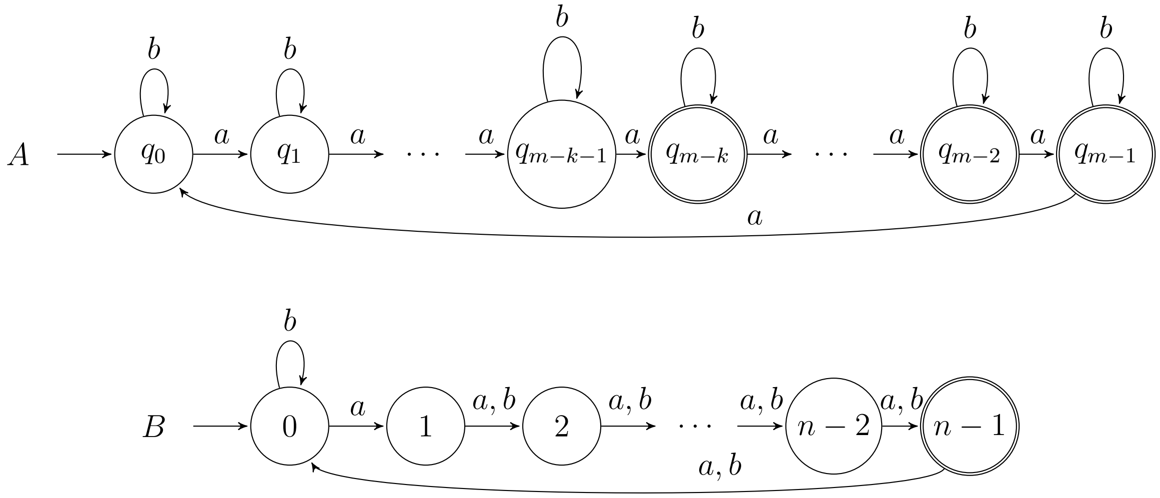

Figure 5: Binary witnesses for k final states meeting the bound $(m-k)2^n + k2^{n-1}$; $m \geq 1, n \geq 2$ [4, Jirásek, Jirásková, Szabari 2005].

Proof. First let $m = 1$, so $K = \{a, b\}^*$ and let L be the language accepted by the DFA B shown in Figure 5 (bottom). Construct an NFA N for KL from B by adding the transition $(0, a, 0)$. In the subset automaton of N, the singleton set $\{0\}$ is the initial subset, and each subset S of size $t+1$ such that $0 \in S$ is reached from the subset $(S \setminus \{0\}) \ominus \min(S \setminus \{0\})$ of size t by the string $ab^{\min(S \setminus \{0\})-1}$. Since the state $n-1$ is uniquely distinguishable and uniquely reachable from every other state in $\{0, 1, \dots, n-1\}$, by Proposition 2.1 all the states of the subset automaton of N are pairwise distinguishable. Hence $\mathrm{sc}(KL) = 2^{n-1}$.

Now let $m \geq 2$. Let K and L be the languages accepted by DFAs A and B shown in Figure 5. Construct an NFA N for KL from DFAs A and B as in the Construction 3.1. Let $\mathcal{R}$ be the same family of $(m-k)2^n + k2^{n-1}$ subsets as in the proof of Lemma 4.3. Let us show that each subset $\{q_i\} \cup S$ in $\mathcal{R}$ is reachable in the subset automaton of N. The proof is by induction on $|\{q_i\} \cup S|$. The basis, with $|\{q_i\} \cup S| \leq 2$, holds true, since we have

$$\{q_0\} \xrightarrow{a^i} \{q_i\} \ (1 \leq i \leq m-k-1),$$
$$\{q_{m-k-1}\} \xrightarrow{a} \{q_{m-k}, 0\} \xrightarrow{(ab^n)^i} \{q_{m-k+i}, 0\} \ (1 \leq i \leq k-1),$$
$$\{q_{m-1}, 0\} \xrightarrow{a} \{q_0, 1\} \xrightarrow{b^{j-1}} \{q_0, j\} \ (2 \leq j \leq n-1), \text{ and } \{q_0, n-1\} \xrightarrow{b} \{q_0, 0\},$$
$$\{q_0, j \ominus i\} \xrightarrow{a^i} \{q_i, j\} \ (1 \leq i \leq m-k-1).$$

bound	2	4	6	[9]	2	4	6	[13]	2	4	6	[4]	2	4	6
2	6	24	96	2	5	4	18	2	6	14	27	2	6	22	84
4	12	48	192	4	10	5	35	4	12	28	54	4	12	42	156
6	18	72	288	6	15	6	52	6	18	42	81	6	18	63	225

Table 1: The state complexity of concatenation if the witness languages from [4, 9, 13] have the second half of their states final; in rows we have m, in columns n.

Let $1 \le t \le n$, and assume the each set in $\mathcal{R}$ of size t is reachable. By Lemma 4.2, every set $\{q_{m-k}\} \cup S$ in $\mathcal{R}$ of size $t+1$ is reachable from a set in $\mathcal{R}$ of size t. Now we reach the sets $\{q_i\} \cup S$ in $\mathcal{R}$ of size $t+1$ for all $i = m-k+1, m-k+2, \dots, m-1$. Let $S' = S \setminus \{0\}$. Then

$$\{q_{i-1}\} \cup (S' \ominus \min S') \xrightarrow{a} \{q_i\} \cup \{0\} \cup (S' \ominus (\min S' - 1)) \xrightarrow{b^{\min S' - 1}} \{q_i\} \cup S,$$

notice that $0 \in S' \ominus \min S'$. This proves the reachability of sets $\{q_i\} \cup S$ if $m-k \le i \le m-1$. Next we prove that for every $S \subseteq Q_B$, the set $\{q_0\} \cup S$ of size $t+1$ is reachable.

If $0 \notin S$, we have $\{q_{m-1}\} \cup (S \ominus \min S) \xrightarrow{a} \{q_0\} \cup (S \ominus (\min S - 1)) \xrightarrow{b^{\min S - 1}} \{q_0\} \cup S$.

If $0 \in S$, let $S' = S \setminus \{0\}$. We have

$$\{q_{m-1}\} \cup (S' \ominus \min S') \cup \{n-1\} \xrightarrow{a} \{q_0\} \cup \{0\} \cup (S' \ominus (\min S' - 1)) \xrightarrow{b^{\min S' - 1}} \{q_0\} \cup S.$$

In both cases, the first set has size $t+1$ and, as shown above, it is reachable. The sets $\{q_i\} \cup S$ in $\mathcal{R}$ of size $t+1$ where $1 \le i \le m-k-1$ are reachable by Lemma 4.2.

To prove distinguishability, let $\{q_i\} \cup S$ and $\{q_j\} \cup T$ be two distinct subsets in $\mathcal{R}$. Notice that the state $n-1$ is uniquely distinguishable since it is a unique final state. Also the state $n-1$ is uniquely reachable by states in Q_B since for every $j = 0, 1, \dots, n-2$ the transition $(j, a, j+1)$ is unique in-transition. By Proposition 2.1, if $S \ne T$, then the sets $\{q_i\} \cup S$ and $\{q_j\} \cup T$ are distinguishable. Now let $S = T$, so $i < j$. If $S = \emptyset$, we read a^{m-k-j} and get $\{q_{m-k-j+i}\}$ and $\{q_{m-k}, 0\}$. If $S \ne \emptyset$, we first read $a^{m-1-j}b^n$ to get $\{q_{m-1-j+i}, 0\}$ and $\{q_{m-1}, 0\}$. When we read a, there are two subcases:

(1) If $m-j+i \ge m-k$, then we get $\{q_{m-j+i}, 0, 1\}$ and $\{q_0, 1\}$ which are distinguishable.
(2) If $m-j+i < m-k$, then we get $\{q_{m-j+i}, 1\}$ and $\{q_0, 1\}$. Then we read $a^{(m-k-1)-(m-j+i)}b^n$ and get $\{q_{m-k-1}, 0\}$ and $\{q_{j-i-k-1}, 0\}$. Finally we read a and get $\{q_{m-k}, 0, 1\}$ and $\{q_{j-i-k}, 1\}$, which are distinguishable.

This proves distinguishability and concludes our proof. □

Our next goal is to describe, for all m, n, k, ℓ with $n \ge 2$, two DFAs of m and n states, and k and ℓ final states, respectively, meeting the upper bound $(m-k)2^n + k2^{n-1}$ on the complexity of the concatenation of their languages. We try to modify the witness automata in all cases, by making the second half of their states final. The upper bound in such a case is $3m \cdot 2^{n-2}$. Table 1 shows that none of the three witnesses presented in [4, 9, 13] meets this bound. Even making two states final in DFA B, results in a complexity of concatenation less that $(m-2)2^n + 2 \cdot 2^{n-1}$ in all three cases. Therefore we present a new pair of witness languages. We use a modified witness from [4, Theorem 1]. However, to cover all possible values of m, n, k, ℓ, we add transitions on a new letter c.

Theorem 4.6 (Ternary Witness Languages with $|F_A| = k$ and $|F_B| = \ell$) *Let $m \geq 1$ and $n \geq 2$. Let $k = 1$ if $m = 1$ and $1 \leq k \leq m-1$ otherwise. Let $1 \leq \ell \leq n-1$. There exists a ternary DFA A with m states and k final states and a ternary DFA B with n states and ℓ final states such that* $\mathrm{sc}(L(A)L(B)) = (m-k)2^n + k2^{n-1}$.

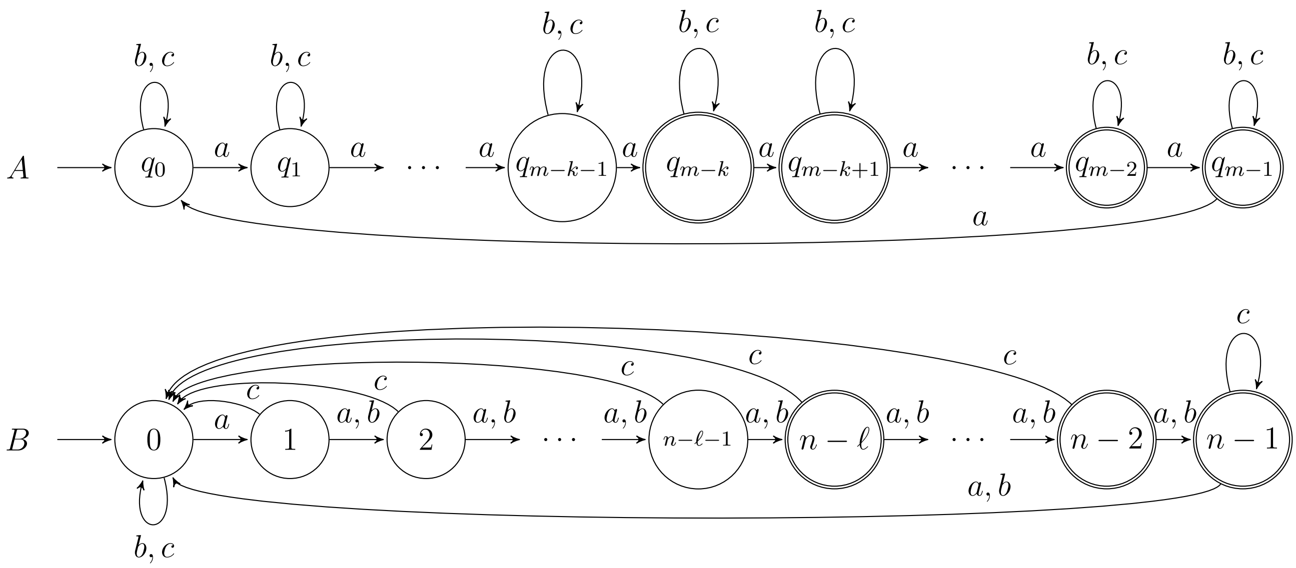

Figure 6: The DFAs with half of states final and $\mathrm{sc}(L(A)L(B)) = (m-k)2^n + k2^{n-1}$.

Proof. Let A and B be the ternary DFAs shown in Figure 6. Notice that the transitions on symbols a and b are the same as in binary DFAs in the proof of Lemma 4.5 shown in Figure 5. On input c, each state q_i of A goes to itself, and each state j of B goes to the state 0, except for $n-1$ which goes to itself. Construct an NFA for $L(A)L(B)$ from DFAs A and B as described in Construction 3.1. Since the transitions on a and b are the same as in DFAs in the proof of Lemma 4.5, the proof of reachability is the same; notice that making some states final in DFA B does not play any role in the proof of reachability.

We only need to prove distinguishability. To this aim, let $\{q_i\} \cup S$ and $\{q_j\} \cup T$ be two distinct reachable subsets. Notice that the state $n-1$ is uniquely distinguishable by the string c since we have $\ell \leq n-1$, so state 0 is not final. Next, the state $n-1$ is uniquely reachable from all states in Q_B by in-transitions on a. It follows that $\{q_i\} \cup S$ and $\{q_j\} \cup T$ are distinguishable if $S \neq T$. Now let $S = T$. In this case we continue exactly the same way as in the proof of Lemma 4.5. □

We have proved that for every number of states in A and B, except for one state in B, and every number of final states in A and B, except for none or all, there exist ternary automata meeting the upper bound $(m-k)2^n + k2^{n-1}$ for concatenation of their languages. However, we might ask whether there are binary languages with more final states in B meeting this bound.

We provide a positive answer in the next theorem. However, notice that we require $k \le m-2$ here, that is, the first DFA must have at least two non-final states.

Theorem 4.7 *Let $m \ge 3$, $n \ge 4$, $1 \le k \le m-2$, and $1 \le \ell \le n-1$. There exists a binary DFA A with m states and k final states and a binary DFA B with n states and ℓ final states such that* $\mathrm{sc}(L(A)L(B)) = (m-k)2^n + k2^{n-1}$.

Proof. Define an m-state DFA $A = (Q_A, \{a,b\}, \cdot, q_0, F_A)$, where $Q_A = \{q_0, q_1, \dots, q_{m-1}\}$, $F_A = \{q_{m-k}, q_{m-k+1}, \dots, q_{m-1}\}$, and the transition function $\cdot$ is defined as follows:

$q_i \cdot a = q_{(i+1) \bmod m}$, $q_0 \cdot b = q_0$ and $q_i \cdot b = q_{i-1}$ if $1 \le i \le m-1$.

Next, define an n-state DFA $B = (Q_B, \{a,b\}, \cdot, 0, F_B)$, where $Q_B = \{0, 1, \dots, n-1\}$, $F_B = \{n-\ell, n-\ell+1, \dots, n-1\}$ if $\ell \le n-2$, and $F_B = Q_B \setminus \{1\}$ if $\ell = n-1$. The transition function $\cdot$ is defined as follows:

$0 \cdot a = 0$, $j \cdot a = j+1$ if $1 \le j \le n-2$, and $(n-1) \cdot a = 1$,

$0 \cdot b = 1$, $1 \cdot b = 2$, and $j \cdot b = j$ if $2 \le j \le n-1$.

The DFAs A and B are shown in Figure 7; notice that the DFA B is the same as in [2]. Since $k \le m-2$, the states q_0 are q_1 are never final. By the definition of B, state 1 is not final either.

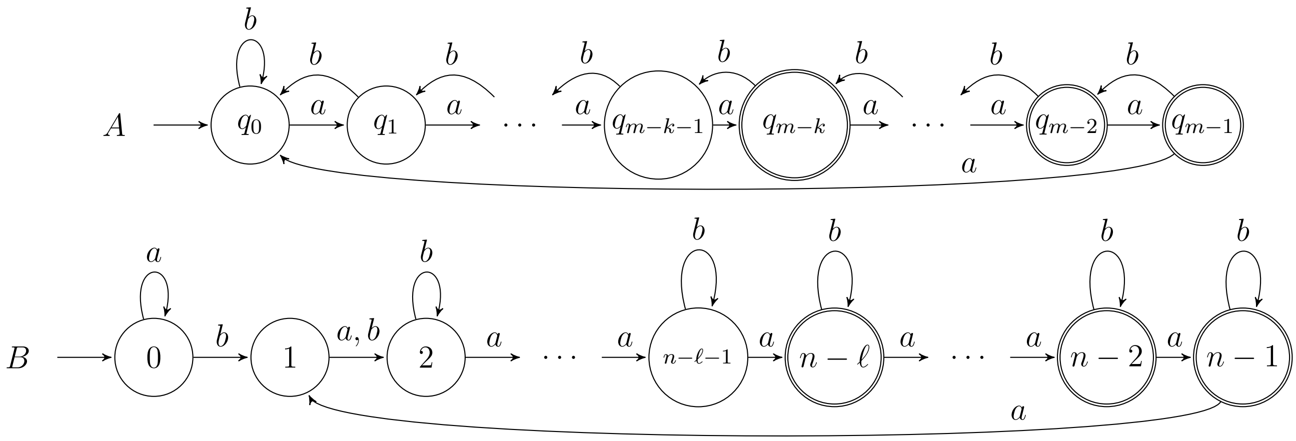

Figure 7: The DFAs with $\mathrm{sc}(L(A)L(B)) = (m-k)2^n + k2^{n-1}$, modified from [2].

To prove the theorem, we construct an NFA N for $L(A)L(B)$ as described in Construction 3.1. We prove that the subset automaton of N has $(m-k)2^n + k2^{n-1}$ reachable and pairwise distinguishable states. We first prove reachability by induction on $|\{q_i\} \cup S|$. The base, with $|\{q_i\} \cup S| = 1$, holds true since $q_0 \xrightarrow{a^i} q_i$ for $1 \le i \le m-k-1$. By Lemma 4.2, we need only to prove that every set $\{q_0\} \cup S$ of size $t+1$ and $0 \notin S$ is reachable from a set $\{q_0\} \cup S'$ of size $t+1$. Let $S' = ((S \ominus (\min S - 1)) \setminus \{1\}) \cup \{0\}$. Then $|S'| = |S|$ and $0 \in S$. Next we have $\{q_0\} \cup S' \xrightarrow{b(ab)^{\min S - 1}} \{q_0\} \cup S$; notice that $1 \in S \ominus (\min S - 1)$ and $q_0 \xrightarrow{ab} q_0$, because $q_1 \notin F_A$ since $k \le m-2$. This proves reachability.

To prove distinguishability, we use the string $w = \prod_{i=0}^{n-4} a^{n-3-i} b^m a^{i+2}$. Notice that we have

$$\{2\} \cdot w = \{2\}; \quad (Q_B \setminus \{2\}) \cdot w = \{1\}; \quad Q_A \cdot w \subseteq Q_A \cup \{0,1\}.$$

We now use these properties to prove distinguishability. We consider several cases:

(1) If $2 \in S$ and $2 \notin T$, then $\{q_i\} \cup S \xrightarrow{wab^m} \{q_0\} \cup \{2,3\} \xrightarrow{(ba)^{n-2}} \{q_1\} \cup \{1,3\}$, and $\{q_j\} \cup T \xrightarrow{wab^m} \{q_0\} \cup \{2\} \xrightarrow{(ba)^{n-2}} \{q_1\} \cup \{1\}$.
If $3 \in F_B$, we have distinguished the sets. If not, we read $a^{n-\ell-3}$ and distinguish the sets since all states j with $j < n - \ell$ are non-final.
(2) If $1 \leq s \leq n-1$, $s \in S$ and $s \notin T$, we read a^{n+1-s} to get the case (1).
(3) If $0 \in S$ and $0 \notin T$, we read b to get the case (2).
(4) If $S = T$ and $1 \leq i < m-k \leq j$, we read ba and get $\{q_i\} \cup S \cdot ba$ and $\{q_j\} \cup \{0\} \cup S \cdot ba$.
If $0 = i < m-k \leq j$, we read ba and get $\{q_1\} \cup S \cdot ba$ and $\{q_j\} \cup \{0\} \cup S \cdot ba$, notice that $0 \notin S \cdot ba$. We get the case (3).
(5) If $m-k \leq i < j$, we read a^{m-j} and get $\{q_{m-j+i}\} \cup S \cdot a^{m-j}$ and $\{q_0\} \cup S \cdot a^{m-j}$, which is considered in the case (4).
(6) If $i < j < m-k$, we read a^{m-k-j} and get $\{q_{m-k-j+i}\} \cup S \cdot a^{m-k-j}$ and $\{q_{m-k}\} \cup \{0\} \cup S \cdot a^{m-k-j}$. If $0 \notin S \cdot a^{m-k-j}$, we get the case (3). If $0 \in S \cdot a^{m-k-j}$, we get the case (4).

This proves distinguishability and concludes our proof. □

5. Concatenation on Alternating Finite Automata

In this section, we consider the concatenation operation on alternating finite automata (AFAs). Our aim is to describe languages K and L accepted by an m-state and n-state AFA, respectively, such that the minimal AFA for the language KL requires $2^m + n + 1$ states. This solves an open problem stated by Fellah, Jürgensen, and Yu in [3], where the upper bound is proved to be the same. First, let us give some basic definitions and notations. For details, we refer the reader to [1, 3, 6, 7, 8, 11, 12].

An *alternating finite automaton* (AFA) is a quintuple $A = (Q, \Sigma, \delta, s, F)$, where Q is a finite non-empty set of states, $Q = \{q_1, \ldots, q_n\}$, Σ is an input alphabet, δ is the transition function that maps $Q \times \Sigma$ into the set $\mathcal{B}_n$ of boolean functions over the n variables $q_1, \ldots, q_n$, $s \in Q$ is the initial state, and $F \subseteq Q$ is the set of final states. For an example, consider AFA $A_1 = (\{q_1, q_2\}, \{a, b\}, \delta, q_1, \{q_2\})$, where transition function δ is given in Table 2. The transition function δ is extended to the domain $\mathcal{B}_n \times \Sigma^*$ as follows: For all g in $\mathcal{B}_n$, a in Σ, and w in Σ^*,

$\delta(g, \varepsilon) = g$;
if $g = g(q_1, \ldots, q_n)$, then $\delta(g, a) = g(\delta(q_1, a), \ldots, \delta(q_n, a))$;
$\delta(g, wa) = \delta(\delta(g, w), a)$.

Next, let $f = (f_1, \ldots, f_n)$ be the boolean vector with $f_i = 1$ iff $q_i \in F$. The language accepted by the AFA A is the set $L(A) = \{w \in \Sigma^* \mid \delta(s, w)(f) = 1\}$. In our example we have $\delta(q_1, ab) = \delta(\delta(q_1, a), b) = \delta(q_1 \vee q_2, b) = q_1 \vee (\overline{q_1} \wedge q_2) = q_1 \vee q_2$. To determine whether $ab \in L(A_1)$, we evaluate $\delta(q_1, ab)$ at the vector $f = (0, 1)$. We obtain 1, hence $ab \in L(A_1)$.

Recall that the state complexity of a regular language L, $\mathrm{sc}(L)$, is the smallest number of states in any DFA accepting L. Similarly, the alternating state complexity of L, $\mathrm{asc}(L)$, is the

δ	a	b
q_1	$q_1 \vee q_2$	q_1
q_2	q_2	$\overline{q_1} \wedge q_2$

Table 2: The transition function of the alternating finite automaton A_1.

smallest number of states in any AFA for L. It follows from [3, Theorem 4.1, Corollary 4.2] and [6, Lemma 1, Lemma 2] that a language L is accepted by an n-state AFA if and only if L^R is accepted by a DFA with 2^n states and 2^{n-1} final states. As this is a crucial observation for this section, we restate these results and provide proof ideas.

Lemma 5.1 ([3, 6]) *Let L be a language accepted by an n-state AFA. Then the reversal L^R is accepted by a DFA of 2^n states, of which 2^{n-1} are final.*

Proof Idea. Let $A = (\{q_1, q_2, \ldots, q_n\}, \Sigma, \delta, q_1, F)$ be an n-state AFA for L. Construct a 2^n-state NFA $A' = (\{0,1\}^n, \Sigma, \delta', S, \{f\})$, where

- for every $u = (u_1 \ldots, u_n) \in \{0,1\}^n$ and every $a \in \Sigma$, $\delta'(u, a) = \{u' \in \{0,1\}^n \mid \delta(q_i, a)(u') = u_i \text{ for } i = 1, \ldots, n\}$;
- $S = \{(b_1, \ldots, b_n) \in \{0,1\}^n \mid b_1 = 1\}$;
- $f = (f_1, \ldots, f_n) \in \{0,1\}^n$ with $f_i = 1$ iff $q_i \in F$.

Then $L(A) = L(A')$, NFA A' has 2^{n-1} initial states and $(A')^R$ is deterministic. It follows that L^R is accepted by a DFA with 2^n states, of which 2^{n-1} are final. □

Corollary 5.2 *For every regular language L, we have* $\operatorname{asc}(L) \geq \lceil \log(\operatorname{sc}(L^R)) \rceil$.

Lemma 5.3 (cf. [6], Lemma 2) *Let L^R be accepted by a DFA A of 2^n states, of which 2^{n-1} are final. Then L is accepted by an n-state AFA.*

Proof Idea. Consider 2^n-state NFA A^R for L which has 2^{n-1} initial states and exactly one final state. Let the state set Q of A^R be $\{0, 1, \ldots, 2^n - 1\}$ with initial states $\{2^{n-1}, \ldots, 2^n - 1\}$ and final state k. Let δ be the transition function of A^R. Moreover, for every $a \in \Sigma$ and for every $i \in Q$, there is exactly one state j such that j goes to i on a in A^R. For a state $i \in Q$, let $\operatorname{bin}(i) = (b_1, \ldots, b_n)$ be the binary n-tuple such that $b_1 b_2 \ldots b_n$ is the binary notation of i on n digits with leading zeros if necessary.

Define an n-state AFA $A' = (Q', \Sigma, \delta', q_1, F')$, where $Q' = \{q_1, \ldots, q_n\}$, $F' = \{q_\ell \mid \operatorname{bin}(k)_\ell = 1\}$, and for each i in Q and a in Σ, $(\delta'(q_1, a), \ldots, \delta'(q_n, a))(\operatorname{bin}(i)) = \operatorname{bin}(j)$ where $i \in \delta(j, a)$. Then $L(A') = L(A^R)$. □

By Corollary 5.2, we have $\operatorname{asc}(L) \geq \lceil \log(\operatorname{sc}(L^R)) \rceil$. The upper bound for concatenation on AFAs is $2^m + n + 1$, as proved by Fellah et al. [3, Theorem 9.3]. They conjectured that this bound is tight. In [6], the lower bound $2^m + n$ was proved, however, the witnesses from [4, Theorem 1] with half of states final in both automata were used. As we mentioned above, cf. Table 1, these witness languages do not meet the upper bound for concatenation on DFAs. Hence the proof in [6, Theorem 5] is not correct, so the problem is still open. Our next aim is to prove the tightness of the upper bound $2^m + n + 1$ for concatenation on AFAs. We might use the ternary witness from Theorem 4.6, but, as we show below, for $\operatorname{asc}(K) \geq 2$, $\operatorname{asc}(L) \geq 2$, it is sufficient to use the binary witness languages described in the proof of Theorem 4.7 to get languages that meet the upper bound $2^m + n + 1$ for concatenation on AFAs. The following theorem not only proves the claim in [6, Theorem 5], but also solves the open problem mentioned above.

Theorem 5.4 (Concatenation on AFAs) *Let* $m, n \geq 2$. *Let* $K, L \subseteq \Sigma^*$ *and* $\operatorname{asc}(K) = m$ *and* $\operatorname{asc}(L) = n$. *Then* $\operatorname{asc}(KL) \leq 2^m + n + 1$, *and this bound is tight if* $|\Sigma| \geq 2$.

Proof. The upper bound on the complexity concatenation of AFA languages is known to be 2^m+n+1 [3, Theorem 9.3]. Let L^R be the binary regular language accepted by the minimal DFA A from the proof of Theorem 4.7, with 2^n states and 2^{n-1} final states. Let K^R be the binary regular language accepted by the minimal DFA B from the proof of Theorem 4.7, with 2^m states and 2^{m-1} final states. Then, by Lemma 5.3, we have $\operatorname{asc}(K) \leq m$ and $\operatorname{asc}(L) \leq n$. Using Theorem 4.7, we get $\operatorname{sc}((KL)^R) = \operatorname{sc}(L^R K^R) = 2^{n-1} \cdot 2^{2^m} + 2^{n-1} \cdot 2^{2^m-1} = 2^{n-1} \cdot 2^{2^m}(1 + 1/2)$. By Corollary 5.2, we have $\operatorname{asc}(KL) \geq \lceil \log(2^{n-1} \cdot 2^{2^m}(1 + 1/2)) \rceil = 2^m + n$.

Our next aim is to show that $\operatorname{asc}(KL) \geq 2^m + n + 1$. Suppose for a contradiction that KL is accepted by an AFA of $2^m + n$ states. Then $(KL)^R$ is accepted by a 2^{2^m+n}-state DFA with 2^{2^m+n-1} final states. It follows that every minimal DFA for $(KL)^R$ has at most 2^{2^m+n-1} final states. However, the minimal DFA for $(KL)^R$ has $2^{n-1}2^{2^m} + 2^{n-1}2^{2^m-1}$ states, of which $2^{n-1}2^{2^{m-1}} + 2^{n-1}2^{2^{m-1}-1}$ are non-final; notice that $\{q_i\} \cup S$ is non-final iff $i \leq 2^{n-1} - 1$ and $S \subseteq \{0, 1, \ldots, 2^{m-1} - 1\}$ or $2^{n-1} \leq i \leq 2^n - 1$ and $S \subseteq \{0, 1, \ldots, 2^{m-1} - 1\}$ with $0 \in S$. Thus the number of final states in the minimal DFA for $(KL)^R$ is $2^{n-1}(2^{2^m}+2^{2^m-1}) - 2^{n-1}(2^{2^{m-1}}+2^{2^{m-1}-1})$, and since $m \geq 2$, we get

$$
\begin{aligned}
&2^{n-1}(2^{2^m} + 2^{2^m-1}) - 2^{n-1}(2^{2^{m-1}} + 2^{2^{m-1}-1}) = \\
&2^{2^m} 2^{n-1}\left(1 + \frac{1}{2} - \frac{1}{2^{2^{m-1}}} - \frac{1}{2^{2^{m-1}+1}}\right) > \\
&2^{2^m+n-1}\left(1 + \frac{1}{2} - \frac{1}{4} - \frac{1}{4}\right) = 2^{2^m+n-1}.
\end{aligned}
$$

Hence, the minimal DFA for $(KL)^R$ has more than 2^{2^m+n-1} final states, a contradiction. It follows that $\operatorname{asc}(KL) \geq 2^m + n + 1$, which proves the theorem. □

6. Conclusions

We studied the state complexity of the concatenation of languages represented by deterministic and alternating finite automata.

First, we looked at ternary languages in [13, Theorem 2.1] and provided a proof that if we give k final states to the first automaton, then the concatenation still meets the upper bound $(m-k)2^n + k2^{n-1}$ if $m, n \geq 2$. Similarly, we proved that if we make more states final in the first automaton in [9], then the resulting concatenation meets the upper bound if $n \geq 3$. Then we fixed errors in the proof of [4, Theorem 1] to get another pair of binary languages meeting this upper bound if $m \geq 1$ and $n \geq 2$.Our computations showed that none of these witnesses meets the upper bound $(m-k)2^n + k2^{n-1}$ if we make k states final in the first automaton, and more than one state final in the second one. Thus, we had to define new ternary languages meeting this upper bound also in the case when the second automaton has more than one final state. We also defined binary languages meeting the same bound in the case when the first automaton has at least two non-final states and both automata have at least four states. Using this result, we were able to describe two languages accepted by an m-state and an n-state alternating finite automata meeting the bound $2^m + n + 1$ for their concatenation. This fixes some errors in [6, Theorem 5] and solves the open problem from [3, Theorem 9.3].

References

[1] J. A. BRZOZOWSKI, E. L. LEISS, On equations for regular languages, finite automata, and sequential networks. *Theoret. Comput. Sci.* 10 (1980), 19–35.

[2] E. DORČÁK, *Concatenation of regular languages and state complexity*. Bachelor thesis (in Slovak), University of P. J. Šafárik in Košice, 2015.

[3] A. FELLAH, H. JÜRGENSEN, S. YU, Constructions for alternating finite automata. *Intern. J. Computer Math.* 35 (1990), 117–132.

[4] J. JIRÁSEK, G. JIRÁSKOVÁ, A. SZABARI, State complexity of concatenation and complementation. *Int. J. Found. Comput. Sci.* 16 (2005) 3, 511–529.

[5] G. JIRÁSKOVÁ, State complexity of some operations on binary regular languages. *Theoret. Comput. Sci.* 330 (2005) 2, 287–298.

[6] G. JIRÁSKOVÁ, Descriptional complexity of operations on alternating and boolean automata. In: *Computer Science - Theory and Applications - 7th International Computer Science Symposium in Russia, CSR 2012. Proceedings*. 2012, 196–204.

[7] E. LEISS, On generalized language equations. *Theoret. Comput. Sci.* 14 (1981) 1, 63 – 77.

[8] E. LEISS, Succinct representation of regular languages by boolean automata. *Theoret. Comput. Sci.* 13 (1981) 3, 323 – 330.

[9] A. MASLOV, Estimates of the number of states of finite automata. *Soviet Math. Doklady* 11 (1970), 1373–1375.

[10] M. O. RABIN, D. SCOTT, *Finite automata and their decision problems*. Technical report, IBM Res. Develop. 3, 1959.

[11] M. SIPSER, *Introduction to the theory of computation*. PWS Publishing Company, 1997.

[12] S. YU, *Chapter 2: Regular languages. In: Handbook of Formal Languages - Vol. I.*. 1, Springer, 1997.

[13] S. YU, Q. ZHUANG, K. SALOMAA, The state complexities of some basic operations on regular languages. *Theoret. Comput. Sci.* 125 (1994) 2, 315 – 328.