# Nondeterministic Biautomata and Their Descriptional Complexity

## (Extended Abstract)*

Markus Holzer and Sebastian Jakobi

Institut für Informatik, Universität Giessen,
Arndtstr. 2, 35392 Giessen, Germany
{holzer,jakobi}@informatik.uni-giessen.de

Biautomata were recently introduced in [4] as a generalization of ordinary deterministic finite automata. A biautomaton consists of a *deterministic* finite control, a read-only input tape, and two reading heads, one reading the input from left to right, and the other head reading the input from right to left. An input word is accepted by a biautomaton, if there is an accepting computation starting with the heads on the two ends of the word and meeting somewhere in an accepting state. Although the choice of reading a symbol by either head is nondeterministic, the determinism of the biautomaton is enforced by two properties, which will be described later. Descriptional complexity issues for deterministic biautomata were addressed in [3]. We focus on the descriptional complexity of *nondeterministic* biautomata, which are defined as follows: a *nondeterministic biautomaton* is a sixtuple $A = (Q, \Sigma, \cdot, \circ, I, F)$, where $Q$ is a finite set of *states*, $\Sigma$ is an *alphabet*, $\cdot\colon Q \times \Sigma \to 2^Q$ is the *forward transition function*, $\circ\colon Q \times \Sigma \to 2^Q$ is the *backward transition function*, $I \subseteq Q$ is the set of *initial states*, and $F \subseteq Q$ is the set of *final states*. The transition functions $\cdot$ and $\circ$ are extended to words in the following way, for every word $v \in \Sigma^*$ and letter $a \in \Sigma$:

$$q \cdot \lambda = \{q\}, \quad q \cdot av = \bigcup_{p \in (q \cdot a)} p \cdot v, \quad \text{and} \quad q \circ \lambda = \{q\}, \quad q \circ va = \bigcup_{p \in (q \circ a)} p \circ v,$$

and further, both $\cdot$ and $\circ$ can be extended to sets of states $S \subseteq Q$, and $w \in \Sigma^*$ by $S \cdot w = \bigcup_{p \in S} p \cdot w$, and $S \circ w = \bigcup_{p \in S} p \circ w$. The biautomaton $A$ *accepts* the word $w \in \Sigma^*$, if and only if $w = u_1 u_2 \dots u_k v_k \dots v_2 v_1$, for some words $u_i, v_i \in \Sigma^*$ with $1 \leq i \leq k$, such that $[((\dots((((I \cdot u_1) \circ v_1) \cdot u_2) \circ v_2) \dots) \cdot u_k) \circ v_k] \cap F \neq \emptyset$. The *language accepted* by $A$ is defined as $L(A) = \{\, w \in \Sigma^* \mid A \text{ accepts } w \,\}$. Moreover, the biautomaton $A$ is *deterministic*, if $|I| = 1$, and $|q \cdot a| = |q \circ a| = 1$ for all states $q \in Q$ and letters $a \in \Sigma$. The automaton $A$ has the $\diamond$-*property*, if $(q \cdot a) \circ b = (q \circ b) \cdot a$, for every state $q \in Q$ and $a, b \in \Sigma$. Further, $A$ has the $F$-property, if $q \cdot a \cap F \neq \emptyset$ if and only if $q \circ a \cap F \neq \emptyset$, for every state $q \in Q$ and letter $a \in \Sigma$. A deterministic biautomaton as defined above

* This is an extended abstract of: M. Holzer, S. Jakobi. *Nondeterministic Biautomata and Their Descriptional Complexity.* In H. Jürgensen, R. Reis (eds.), *Proc.* 15*th DCFS*, volume 8031 in LNCS, pages 112–123, London, Ontario, Canada, July 2013. Springer.

that has both the $\diamond$- and the $F$-property is exactly what is called a biautomaton in [4]. Nondeterministic biautomata characterize the family of linear context-free languages [5], while nondeterministic biautomata that have the $\diamond$-property accept only regular languages. In fact, we are able to prove the following result.

**Theorem 1.** *The trade-off between deterministic or nondeterministic biautomata with or without the $F$-property and deterministic or nondeterministic biautomata that satisfy at least the $\diamond$-property is non-recursive.* $\square$

By a straight-forward power-set construction, one can convert any $n$-state nondeterministic biautomaton into an equivalent deterministic biautomaton having at most $2^n$ states—this construction preserves both the $\diamond$- and the $F$-property. An exponential lower bound for the determinization of nondeterministic biautomata with both properties is given in the following result.

**Theorem 2.** *For all integers $n \geq 1$ there is a binary regular language $L_n$ accepted by a nondeterministic biautomaton with $\diamond$-, and $F$-property that has $3n+2$ states, and for which every equivalent deterministic biautomaton with $\diamond$- and $F$-property needs at least $2^{2n}+1$ states.* $\square$

We also study the costs for the conversions from deterministic or nondeterministic finite automata (DFAs, NFAs), syntactic monoids, and regular expressions into nondeterministic biautomata that have the $\diamond$- and $F$-property. To prove lower bounds for such nondeterministic biautomata, the following generalization of the fooling set technique [1] is useful.

**Theorem 3.** *A set $S = \{\, (x_i, y_i, z_i) \mid x_i, y_i, z_i \in \Sigma^*, 1 \leq i \leq n \,\}$ is a* bi-fooling set *for a language $L \subseteq \Sigma^*$ if the following two properties hold:*

1. *for $1 \leq i \leq n$ it is $x_i y_i z_i \in L$, and*
2. *for $1 \leq i, j \leq n$, with $i \neq j$, it is $x_i y_j z_i \notin L$ or $x_j y_i z_j \notin L$.*

*If $S$ is a bi-fooling set for the language $L$, then any nondeterministic biautomaton with both the $\diamond$-property and the $F$-property that accepts the language $L$ has at least $|S|$ states.* $\square$

Table 1 summarizes our results on the costs of the above mentioned conversions. For comparison we also list the results from [3] for the conversions from DFAs, NFAs, and syntactic monoids to deterministic biautomata with both properties. Except for the conversion from regular expressions to nondeterministic biautomata, the indicated bounds are tight bounds, i.e., matching lower and upper bounds. We exemplarily present the results on the conversion from regular expressions to biautomata. We measure the size of a regular expression $r$ by its *alphabetic width*, which is the number of occurrences of symbols from the underlying alphabet in the expression. The following upper bound is obtained by adapting the Glushkov construction [2] to biautomata.

**Theorem 4.** *Let $r$ be a regular expression of alphabetic width $n$. Then there is a nondeterministic biautomaton $A$ with $L(A) = L(r)$ that has $(n+1)^2$ states. Further, $A$ has the $\diamond$- and the $F$-property.* $\square$

| Convert from ... | ... to Biautomaton | |
|---|---|---|
| | deterministic, $\diamond$, $F$ | nondeterministic, $\diamond$, $F$ |
| DFA | $n \cdot 2^n - 2 \cdot (n-1)$ | $n^2$ |
| NFA | $2^{2n} - 2 \cdot (2^n - 1)$ | $n^2$ |
| syntactic monoid | $n^2$ | $n$ |
| regular expression | | $n^2 \leq \cdot \leq (n+1)^2$ |

**Table 1.** Tight bounds for conversions from different models describing regular languages to deterministic or nondeterministic biautomata with the $\diamond$-property and the $F$-property. The results on deterministic biautomata are from [3]. For the conversions starting from a DFA or NFA, the integer $n$ is the number of states of the finite automaton, when starting from a syntactic monoid, the number $n$ is the size of the monoid, and for regular expressions, the integer $n$ is the alphabetic width of the expression.

The following result provides a lower bound for this conversion.

**Theorem 5.** *For all integers $n \geq 1$ there is a binary language $L_n$ with alphabetic width $n$, such that any nondeterministic biautomaton with the $\diamond$- and the $F$-property needs $n^2$ states to accept the language $L_n$.* □

The results and constructions for nondeterministic biautomata with $\diamond$- and $F$-property are evidence that this automaton model is a reasonable nondeterministic counterpart of the model of biautomata, as introduced in [4]. Concerning the $F$-property, its influence on the size of the biautomata is yet to be studied. By close inspection of the proof that biautomata with $\diamond$-property accept regular languages, one can deduce a quadratic upper bound for converting biautomata with $\diamond$-property into equivalent nondeterministic finite automata. With the conversions from finite automata to biautomata one can obtain upper bounds for enforcing the $F$-property, while preserving the $\diamond$-property—in case of nondeterministic biautomata, the bound is polynomial, and for deterministic biautomata it is exponential. The search for tight bounds for these conversions is left as an open problem.

## References


1. J.-C. Birget. Intersection and union of regular languages and state complexity. *Inf. Process. Lett.*, 43:185–190, 1992.
2. V. M. Glushkov. The abstract theory of automata. *Russ. Math. Surv.*, 16:1–53, 1961.
3. G. Jirásková and O. Klíma. Descriptional complexity of biautomata. In M. Kutrib, N. Moreira, and R. Reis, editors, *Proc.* 14*th DCFS*, volume 7386 of *LNCS*, pages 196–208, Braga, Portugal, July 2012. Springer.
4. O. Klíma and L. Polák. On biautomata. RAIRO – Theo. Inf. Appl. 46(4):573–592, Oct. 2012.
5. R. Loukanova. Linear context free languages. In C. B. Jones, Z. Liu, and J. Woodcock, editors, *Proc.* 4*th ICTAC*, volume 4711 of *LNCS*, pages 351–365, Macau, China, Sept. 2007. Springer.