



On Finitely Ambiguous Büchi Automata

Christof Löding and Anton Pirogov^(✉)

RWTH Aachen University, Templergraben 55, 52062 Aachen, Germany
{loeding,pirogov}@cs.rwth-aachen.de

Abstract. Unambiguous Büchi automata, i.e. Büchi automata allowing only one accepting run per word, are a useful restriction of Büchi automata that is well-suited for probabilistic model-checking. In this paper we propose a more permissive variant, namely *finitely ambiguous Büchi automata*, a generalisation where *each word has at most k accepting runs, for some fixed k* . We adapt existing notions and results concerning finite and bounded ambiguity of finite automata to the setting of ω -languages and present a translation from arbitrary nondeterministic Büchi automata with n states to finitely ambiguous automata with at most 3^n states and at most n accepting runs per word.

Keywords: Büchi automata · Infinite words · Ambiguity

1 Introduction

Nondeterministic Büchi automata (NBA) [6] are finite automata for infinite words that have applications in logical decision procedures, in particular in the field of model checking [3], as they can succinctly represent many interesting properties of non-terminating systems with infinite execution traces. In some contexts, unrestricted nondeterminism is problematic, e.g. in probabilistic model checking reasoning about probabilities becomes very difficult under nondeterminism and therefore other models are necessary.

One solution is determinisation of Büchi automata. As deterministic Büchi automata are strictly weaker than NBA, this requires a quite complex translation to automata with different acceptance conditions and incurs a state blow-up of order $2^{n \log n}$ in the worst-case [14, 16]. But determinisation can be avoided because some restricted forms of nondeterminism are also suitable for probabilistic model checking, e.g. *limit-deterministic Büchi automata* [8], which can be separated into a subset of states that can never accept, but have nondeterministic transitions, and a deterministic subset that contains all of the accepting states, but cannot reach the nondeterministic states again. Such automata can be used with similar algorithms as deterministic automata and are also suitable for the model checking of Markov decision processes [9].

A. Pirogov—This work is supported by the German research council (DFG) Research Training Group 2236 UnRAVeL.

Another well-studied variant are *unambiguous Büchi automata* (UBA), i.e. automata admitting at most one accepting run for each word, which are known to be as powerful as unrestricted Büchi automata [2], while they can be exponentially smaller than equivalent deterministic automata [4]. On finite words, unambiguous automata form an interesting subclass of nondeterministic automata because they admit a polynomial time inclusion test [15] (while this problem is PSPACE-complete for general nondeterministic automata). This result can be extended to finitely ambiguous automata, which have at most k accepting runs for each input for some fixed number k [15]. It is unknown whether the polynomial time inclusion test can be extended to UBA. However, some positive results have been obtained for simpler types of acceptance conditions [10] and a stronger notion of ambiguity [5]. Furthermore, UBA admit a polynomial time algorithm for quantitative probabilistic model checking based on linear equation systems [4].

While standard translations from the temporal logic LTL to NBA yield unambiguous automata, for the transformation of a given NBA into an UBA only non-trivial constructions [11, 12] roughly of order n^n for an n state NBA are known.

In this article, we study finitely ambiguous NBA. To the best of our knowledge, this model has not been considered before. We show that there is a simple construction for transforming any given NBA with n states into a finitely ambiguous NBA with at most 3^n states. We also present an exponential lower bound of order 2^n for such a construction, which is easily obtained from a corresponding lower bound for finitely ambiguous automata on finite words [13]. Furthermore, we study the possible degrees of ambiguity for NBA. We present a classification of the degree of ambiguity of Büchi automata and the complexity of the corresponding decision problems, based on results for finite words in [1, 7, 17]. While many results can be generalized from finite to infinite words in a straightforward way, there are different types of infinite degree of ambiguity for NBA. We characterize those in terms of state patterns similar to those that are used over finite words to distinguish between polynomial and exponential growth rates of ambiguity.

This paper is organized as follows. After introducing basic notation in Sect. 2 we present the classification degrees of ambiguity of nondeterministic Büchi automata and the complexity of the corresponding decision problems in Sect. 3. In Sect. 4 we present the translation from an arbitrary NBA to a finitely ambiguous NBA, and state a lower bound for such a transformation. In Sect. 5 we conclude. Due to space constraints, some proofs are only sketched.

2 Preliminaries

For a finite alphabet Σ , Σ^* denotes the set of all finite and Σ^ω the set of all infinite words over Σ . For $a_i \in \Sigma$ and a (finite or infinite) word $w = a_1 a_2 \dots$ let $w(i) := a_i$. We denote the prefix of length i with $\text{Pref}_i(w) := a_1 \dots a_i$. The suffix starting at position i is denoted by $\text{Suf}_i(w) := a_i a_{i+1} \dots$. Let $\text{Inf}(w) = \{x \mid w(i) = x \text{ for infinitely many } i\}$ denote the *infinity set* of a word w .

Let a tuple $\mathcal{A} = (Q, \Sigma, \Delta, Q_0, F)$ denote a *finite automaton* with some finite alphabet Σ , finite set of states Q , transition relation $\Delta \subseteq Q \times \Sigma \times Q$ and initial and final states $Q_0, F \subseteq Q$. Let $|\mathcal{A}| := |\Delta|$ denote the size of \mathcal{A} . We write $\Delta(P, x)$ for $\{q \mid p \in P, (p, x, q) \in \Delta\}$, $\Delta_S(P, x) = \Delta(P, x) \cap S$ and $\Delta_{\bar{S}}(P, x) = \Delta(P, x) \setminus S$ for some $S \subseteq Q$. For convenience, we write $\Delta(p, x)$ when we mean $\Delta(\{p\}, x)$. With $\mathcal{A}[S]$ we denote the modified automaton with $Q_0 = S$, $\mathcal{A}[\{s\}]$ can be written as $\mathcal{A}[s]$.

A transition sequence $\pi = (q_1, a_1, r_1) \dots (q_n, a_n, r_n)$ is called a *path* if q_i equals r_{i-1} for all $1 < i \leq n$. The source and target of the path are denoted by $\text{src}(\pi) = q_1$ and $\text{trg}(\pi) = r_n$ while $\text{lbl}(\pi) = a_1 \dots a_n$ and $\text{st}(\pi) = q_1 \dots q_n r_n$ denote the label and state sequence of π , respectively. For convenience, let $\pi(i) := \text{st}(\pi)(i)$. We also consider infinite paths (with the obvious definition). In general, when speaking of a path or a sequence, we refer to a finite or infinite path or sequence, depending on the context.

The set of all paths from states in Q to states in R labelled with x is denoted by $P(Q, x, R) := \{\pi \mid \text{src}(\pi) \in Q, \text{lbl}(\pi) = x, \text{trg}(\pi) \in R\}$, while $P^\omega(Q, x) := \{\pi \mid \text{src}(\pi) \in Q, \text{lbl}(\pi) = x, x \in \Sigma^\omega\}$ denotes all infinite paths with label x starting in a state from Q . Paths compose in the expected way. We write $p \xrightarrow{x} q$ if $P(p, x, q) \neq \emptyset$ and $p \rightarrow q$ if $p \xrightarrow{x} q$ for some $x \in \Sigma^*$. A *strongly connected component (SCC)* $C \subseteq Q$ of \mathcal{A} as usual is a maximal (w.r.t. inclusion) subset of states such that if $p, q \in C$ then $p \rightarrow q$ and $q \rightarrow p$.

The language of \mathcal{A} when read as NFA is defined as $L(\mathcal{A}_{\text{NFA}}) := \{x \in \Sigma^* \mid P(Q_0, x, F) \neq \emptyset\}$. The ω -language of \mathcal{A} when read as nondeterministic Büchi automaton (NBA) is defined as $L(\mathcal{A}_{\text{NBA}}) := \{x \in \Sigma^\omega \mid \exists \pi \in P^\omega(Q_0, x), \text{Inf}(\text{st}(\pi)) \cap F \neq \emptyset\}$. The set $\text{Runs}(\mathcal{A}, x)$ contains accepting runs of \mathcal{A} on x , i.e. all paths from an initial state that are labelled with x and satisfy the corresponding acceptance condition. We say that a set of runs is *separated (at time i)* when the prefixes of length i of those runs are pairwise different.

We say \mathcal{A} is *trim*, if each path from an initial state is a prefix of an accepting run and if \mathcal{A} is an NBA we additionally require that each accepting state is on a cycle. This means for NFA that an accepting state is always reachable and for NBA that a cycle with an accepting state is always reachable and no state is uselessly marked as accepting. In the following, let $\mathcal{A} = (Q, \Sigma, \Delta, Q_0, F)$ be some arbitrary finite automaton if not specified otherwise.

3 Ambiguity of Büchi Automata

We first give some basic definitions concerning the degree of ambiguity of NFA and NBA. Then we restate some results on the ambiguity of NFA given in [17], and continue with the analysis of degrees of ambiguity for NBA.

By \aleph_0 we denote the cardinality of the natural numbers and by 2^{\aleph_0} the cardinality of the real numbers. The *degree of ambiguity* of automaton \mathcal{A} on a word x is defined as $\text{da}(\mathcal{A}, x) := |\text{Runs}(\mathcal{A}, x)|$ and the degree of ambiguity of an automaton is given by $\text{da}(\mathcal{A}) := \sup_x \{\text{da}(\mathcal{A}, x)\}$ over all possible words x . Note that the result depends on whether we consider \mathcal{A} to be an NFA or an NBA – in

the first case we consider finite input words, in the second case infinite words. If $\text{da}(\mathcal{A}) < \aleph_0$, \mathcal{A} is *finitely ambiguous*. We say \mathcal{A} is *k-ambiguous* if $\text{da}(\mathcal{A}) = k$ and *unambiguous* for $k = 1$. If $\text{da}(\mathcal{A}) \geq \aleph_0$, \mathcal{A} is *infinitely ambiguous*.

For infinitely ambiguous \mathcal{A}_{NFA} let the *degree of polynomial ambiguity* $\text{dpa}(\mathcal{A})$ be the smallest $k \in \mathbb{N}$ such that for all $w \in \Sigma^*$, $|w| = n$, $\text{da}(\mathcal{A}, w) \in \mathcal{O}(n^k)$. \mathcal{A}_{NFA} is *polynomially ambiguous* if $\text{dpa}(\mathcal{A}) < \infty$, otherwise *exponentially ambiguous*.

So polynomial ambiguity on finite words means that there is no constant upper bound on the number of accepting runs that holds for all words, but there is a polynomial function bounding the number of accepting runs for words with a fixed length. Similarly, exponential ambiguity means that no such polynomial bound exists. As shown in [17], the different types of ambiguity for NFA can be characterized by the following state patterns:

Definition 1 (Infinite ambiguity conditions for NFA).

- \mathcal{A} satisfies IDA (infinite degree of ambiguity) if there are $p, q \in Q, p \neq q, v \in \Sigma^*$ and three paths $\pi_1 \in P(p, v, p), \pi_2 \in P(p, v, q), \pi_3 \in P(q, v, q)$. We call a tuple $(p, q, v, \pi_{1,2,3})$ an IDA pattern.
- \mathcal{A} satisfies EDA (exponential degree of ambiguity) if there is $p \in Q, v \in \Sigma^*$ and two cycles $\pi_1, \pi_2 \in P(p, v, p), \pi_1 \neq \pi_2$. We call a tuple $(p, v, \pi_{1,2})$ an EDA pattern.

The corresponding paths π_i may be omitted when not required.

Theorem 1 ([17]).

1. If \mathcal{A} satisfies EDA, then \mathcal{A} satisfies IDA.
2. \mathcal{A} satisfies EDA $\Leftrightarrow \mathcal{A}_{\text{NFA}}$ is exponentially ambiguous.
3. \mathcal{A} satisfies $\neg\text{EDA}$ and IDA $\Leftrightarrow \mathcal{A}_{\text{NFA}}$ is polynomially ambiguous.
4. \mathcal{A} satisfies $\neg\text{IDA} \Leftrightarrow \mathcal{A}_{\text{NFA}}$ is finitely ambiguous.

A novel aspect when measuring ambiguity of Büchi automata is that there are multiple degrees of infinite ambiguity when considering infinite words, as a single infinite word can have infinitely many different accepting runs, which is not possible with finite words. In fact, for an infinite word the number of accepting runs can even be uncountable. We will see that for each infinite word the cardinality of the set of different accepting runs is either finite, equal to \aleph_0 or equal to 2^{\aleph_0} .

Formally, if there exists $w \in \Sigma^\omega$ with $\text{da}(\mathcal{A}, w) = 2^{\aleph_0}$, i.e. some word w has uncountably many accepting runs, we say that \mathcal{A} is *uncountably ambiguous* and we write $\text{da}(\mathcal{A}) = 2^{\aleph_0}$. If \mathcal{A} is not uncountably ambiguous, but there exists some word w with $\text{da}(\mathcal{A}, w) = \aleph_0$, i.e. w has a countably infinite number of accepting runs, \mathcal{A}_{NBA} is called *strict-countably ambiguous*. Later we will show that these ambiguity cases can be characterized using the following refinements of the state patterns in Definition 1:

Definition 2 (Additional ambiguity conditions for NBA).

- \mathcal{A} satisfies IDA_F if it has an IDA pattern $(p, q, v, \pi_{1,2,3})$ such that $q \in F$.
- \mathcal{A} satisfies EDA_F if it has an EDA pattern $(p, v, \pi_{1,2})$ such that $p \in F$.

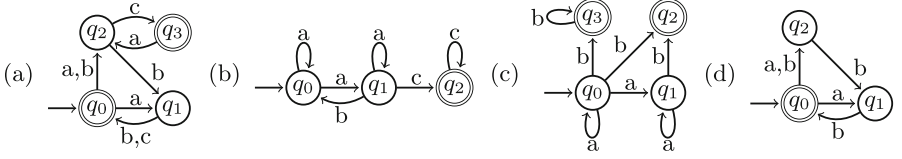


Fig. 1. (a) The word $(ab)^\omega$ is accepted unambiguously, while ab^ω has two accepting runs due to the choice of the first transition. The word $(ac)^\omega$ has strict-countable ambiguity as $q_0 \xrightarrow{ac} q_0, q_0 \xrightarrow{ac} q_3$ and $q_3 \xrightarrow{ac} q_3$, which is an IDA_F pattern. The word $(acabb)^\omega$ has uncountably many accepting runs due to two paths $q_0 \xrightarrow{acabb} q_0$, implying EDA_F and therefore the automaton is uncountably ambiguous. (b) This automaton has an IDA pattern (q_0, q_1, a) and an EDA pattern (q_0, aab) , but no IDA_F nor EDA_F . The word a^*ac^ω is polynomially ambiguous and $(aab)^*ac^\omega$ is exponentially ambiguous, as the corresponding pattern can be traversed in different ways a finite number of times, before reading the first c . (c) Counterexample of Lemma 1 for non-trim automata: $L(\mathcal{A}_{\text{NFA}}) = a^*b^+$ and \mathcal{A}_{NFA} is not finitely ambiguous, because each word a^nb has for $0 < i \leq n$ the accepting runs $q_0^i q_1^{n-i} q_2$, while $L(\mathcal{A}_{\text{NBA}}) = a^*b^\omega$ is unambiguous, as the only accepting path must have a state sequence of the form $q_0^* q_3$. (d) As NFA this automaton is unambiguous, while as NBA there are two accepting runs on ab^ω .

If the ambiguity of the NBA is not finite, but there are also no infinite words with at least \aleph_0 accepting runs, we call the NBA *limit-countably ambiguous*. In this case we can adapt the notions of polynomial and exponential ambiguity. These cases can not be defined in exactly the same way as for NFA because we only consider words that are infinite. But we can still preserve the spirit of the definitions of polynomial and exponential ambiguity by defining them as bounds on the maximal growth of ambiguity with increasing prefix length of words, instead of whole words.

So formally, if \mathcal{A}_{NBA} is not finitely ambiguous and not at least strict-countably ambiguous, it is limit-countably ambiguous. More specifically, if the function f is an upper bound such that for all $w \in L(\mathcal{A}_{\text{NBA}})$ and $i \in \mathbb{N}$ we have $|\{\pi \in P(Q_0, \text{Pref}_i(w), Q) \mid \text{Runs}(\mathcal{A}[\text{trg}(\pi)]_{\text{NBA}}, \text{Suf}_{i+1}(w)) \neq \emptyset\}| \leq f(i)$, we say that \mathcal{A}_{NBA} is polynomially ambiguous if one can choose $f(i) := c \cdot i^d$ with constants c and d , and exponentially ambiguous otherwise.

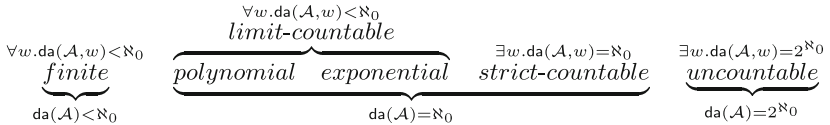


Fig. 2. Illustration of the ambiguity hierarchy for NBA. The five classes (without “limit-countable”) are pairwise disjoint. The depicted order reflects the meaning of e.g. “at most polynomial amb.”, which includes finite and polynomial ambiguity, or “at least strict-countable amb.”, which includes strict-countable and uncountable ambiguity.

Consider Fig. 1(a) and (b) to see examples of different ambiguity types on infinite words. In the following we will justify the ambiguity claims in the examples, by relating the patterns from Definition 2 to the various cases that emerge for NBA. The resulting hierarchy is illustrated in Fig. 2 and summarized in Theorem 2. We start by establishing that the notion of finite ambiguity of an automaton is closely related for NFA and NBA under the condition that the considered automaton must be trim as NBA.

Lemma 1. *Let \mathcal{A} be trim as NBA. Then \mathcal{A}_{NFA} is finitely ambiguous if and only if \mathcal{A}_{NBA} is finitely ambiguous.*

Proof. For the first claim, assume \mathcal{A} is not a finitely ambiguous NBA. Then for all $k \in \mathbb{N}$ there is a word with at least k different runs. Pick a word $w \in \Sigma^\omega$ with at least $k|Q|$ different runs and a time where all these runs are separated. Then there are at least k runs that are in the same state $p \in Q$. We can extend these prefixes by a path from p to some $q \in F$, obtaining a word that is accepted by at least k runs of the corresponding NFA.

The second claim is similar, with the difference that k different finite runs on some finite word that end in the same state are extended to accepting infinite runs (this requires that \mathcal{A} is trim). \square

See Fig. 1(c) for a non-trim counterexample and also notice that in (d) the automaton has a different finite ambiguity as NFA than as NBA. Hence, in general $\text{da}(\mathcal{A}_{\text{NBA}}) \neq \text{da}(\mathcal{A}_{\text{NFA}})$ and the calculation of the exact degree must also be adapted to the NBA setting. We will sketch a corresponding procedure later in the context of Theorem 4.

We now state some technical lemmas for establishing the connection between the state patterns and the degrees of ambiguity summarized in Theorem 2 further below.

Lemma 2. *If \mathcal{A} satisfies EDA_F , then \mathcal{A} satisfies IDA_F .*

Proof (sketch). Similar argument as for the implication $\text{EDA} \Rightarrow \text{IDA}$ in [17]. \square

Lemma 3. *If \mathcal{A} satisfies $\neg \text{EDA}_F$, then for all $q \in F, w \in \Sigma^\omega$ the number of infinite paths of $\mathcal{A}[q]$ visiting q infinitely often is at most $|Q|$.*

Proof (sketch). Shown by simple construction of an EDA_F pattern in case of more than $|Q|$ such paths. \square

Now we can relate the extended patterns to the new infinite ambiguity cases.

Lemma 4. *\mathcal{A} satisfies $\text{EDA}_F \Leftrightarrow \mathcal{A}_{\text{NBA}}$ is uncountably ambiguous. Furthermore, if \mathcal{A} does not satisfy EDA_F , then \mathcal{A}_{NBA} is at most strict-countably ambiguous.*

Proof. For one direction, let $(p, v, \pi_{1,2})$ be an EDA pattern satisfying EDA_F . Pick some $u \in \Sigma^*, \pi_0 \in P(Q_0, u, p)$. Clearly, $uv^\omega \in L(\mathcal{A}_{\text{NBA}})$. Observe that for $v^\omega = v_0v_1\dots$ each v_i can be consumed by taking either π_1 or π_2 . Hence, the number of runs on uv^ω is uncountable, i.e., $\text{da}(\mathcal{A}_{\text{NBA}}) = \text{da}(\mathcal{A}_{\text{NBA}}, uv^\omega) = 2^{\aleph_0}$.

For the other direction, let $w \in L(\mathcal{A}_{\text{NBA}})$ and assume EDA_F does not hold. For each $i \in \mathbb{N}$, let $x = \text{Pref}_i(w)$ and notice that the number of different paths $P_x = P(Q_0, x, q)$ must be finite for each $q \in F$, as x is finite. By Lemma 3, for each such path there are at most $|Q|$ continuations to infinite runs that visit q infinitely often. It follows that \mathcal{A}_{NBA} is at most strict-countably ambiguous. \square

Lemma 5. *\mathcal{A} satisfies $\text{IDA}_F \Leftrightarrow \mathcal{A}_{\text{NBA}}$ is at least strict-countably ambiguous.*

Proof (sketch). For one direction, let $(p, q, v, \pi_{1,2,3})$ be an IDA pattern satisfying IDA_F . Let $u \in \Sigma^*$, $\pi_0 \in P(Q_0, u, p)$. Clearly, $uv^\omega \in L(\mathcal{A}_{\text{NBA}})$. Observe that for each $i \in \mathbb{N}$, π_1 can be taken i times before using π_2 and then taking path π_3 forever. Hence, $\text{da}(\mathcal{A}_{\text{NBA}}) \geq \text{da}(\mathcal{A}_{\text{NBA}}, uv^\omega) \geq \aleph_0$.

For the other direction, let $w \in L(\mathcal{A}_{\text{NBA}})$ and assume IDA_F does not hold. We show that the number of runs on w is finite, and thus the degree of ambiguity of \mathcal{A}_{NBA} is at most limit-countable. Since IDA_F does not hold, EDA_F does not hold by Lemma 2. Then by Lemma 3 a run can separate into at most $|Q|$ different accepting runs that visit $q \in F$ infinitely often, after seeing q the first time. Assume that there is some state $q \in F$ such that there are infinitely many accepting runs that visit q infinitely often. Then there must be such runs for which the first visit to q happens arbitrarily late. From that observation one can construct an IDA_F pattern, a contradiction. But then the number of accepting runs on w must be finite. \square

The relationship of the different state patterns in \mathcal{A} and the corresponding ambiguity classes of the Büchi automaton are summed up in the following theorem:

Theorem 2. *Let \mathcal{A} be a trim NBA.*

1. \mathcal{A} satisfies $\text{EDA}_F \Leftrightarrow \mathcal{A}_{\text{NBA}}$ is uncountably ambiguous.
2. \mathcal{A} satisfies $\neg\text{EDA}_F$ and $\text{IDA}_F \Leftrightarrow \mathcal{A}_{\text{NBA}}$ is strict-countably ambiguous.
3. \mathcal{A} satisfies $\neg\text{IDA}_F$ and $\text{EDA} \Leftrightarrow \mathcal{A}_{\text{NBA}}$ is exponentially ambiguous.
4. \mathcal{A} satisfies $\neg\text{IDA}_F, \neg\text{EDA}$ and $\text{IDA} \Leftrightarrow \mathcal{A}_{\text{NBA}}$ is polynomially ambiguous.
5. \mathcal{A} satisfies $\neg\text{IDA}_F$ and $\text{IDA} \Leftrightarrow \mathcal{A}_{\text{NBA}}$ is limit-countably ambiguous.
6. \mathcal{A} satisfies $\neg\text{IDA} \Leftrightarrow \mathcal{A}_{\text{NBA}}$ is finitely ambiguous.

Proof.

(1+2): By Lemma 4 and Lemmas 4 and 5, respectively.

(3+4): We show both directions by establishing that the difference in growth of ambiguity of \mathcal{A}_{NFA} in the length of finite words and the growth of ambiguity of \mathcal{A}_{NBA} in the length of finite prefixes is bounded by constants.

(\Rightarrow): $\neg\text{IDA}_F$ implies $\neg\text{EDA}_F$ by Lemma 2 and by Lemma 5 \mathcal{A}_{NBA} is not strict-countably ambiguous, hence no word has infinitely many runs. But by Theorem 1 [17] there is a family of finite words with increasing length that w.l.o.g. terminate in the same $q \in F$ and witness the polynomial (exponential) ambiguity of \mathcal{A}_{NFA} . As \mathcal{A} is trim, for each such word u with k accepting runs, those runs can be extended by the same loop from q

to q labelled with some v , hence uv^ω has at least k (and by Lemma 3 at most $k|Q|$) accepting runs in \mathcal{A} as NBA. Hence \mathcal{A}_{NBA} has asymptotically the same limit-countable ambiguity.

(\Leftarrow) : Pick some $w \in L(\mathcal{A}_{\text{NBA}})$ with $\text{da}(\mathcal{A}_{\text{NBA}}, w) = k$ and pick i such that after reading the prefix of length i all accepting runs have separated. Notice that there must be at least $\lfloor \frac{k}{|Q|} \rfloor$ different runs that are in the same state p . As \mathcal{A} is trim, there is some finite word x with $|x| \leq |Q|$ leading from p to some $q \in F$. Let $a \leq |Q|^{|Q|}$ be the maximum number of different paths of length $|Q|$ in \mathcal{A} that have the same source, target and label and let $\hat{w} = \text{Pref}_i(w)x$. Then on any x each run can separate into at most a different runs and therefore we have that $\lfloor \frac{k}{|Q|} \rfloor \leq \text{da}(\mathcal{A}_{\text{NFA}}, \hat{w}) \leq ak$. Hence, by picking for each i an infinite word w with the largest number of separated accepting runs after reading $\text{Pref}_i(w)$, we can construct a family of finite words with the number of accepting runs growing asymptotically in the same way as the maximum number of separated accepting runs grows on prefixes of infinite words. By Theorem 1 [17] this implies that \mathcal{A} must satisfy IDA or EDA, respectively.

(5): By (3), (4) and Theorem 1, as EDA implies IDA.

(6): Shown in Theorem 1 for trim NFA. By Lemma 1 this extends to trim NBA. \square

The ambiguity class of an NBA from the hierarchy in Theorem 2 can also be determined efficiently:

Theorem 3.

1. Uncountable ambiguity of an NBA \mathcal{A} can be decided in $\mathcal{O}(|\mathcal{A}|^2)$.
2. The ambiguity class of an NBA \mathcal{A} can be computed in $\mathcal{O}(|\mathcal{A}|^3)$.
3. $\text{dpa}(\mathcal{A}_{\text{NBA}})$ can be computed in $\mathcal{O}(|\mathcal{A}|^3)$.

Proof. A straightforward modification of the corresponding algorithms for NFA from [1], which use a depth-first search in products of \mathcal{A} with itself. EDA_F and IDA_F can easily be checked by adding the corresponding restriction to the IDA / EDA pattern, i.e., that a specific state must additionally be accepting. In case of polynomial ambiguity, the algorithm to calculate $\text{dpa}(\mathcal{A}_{\text{NBA}})$ can be used without changes after excluding EDA and IDA_F . \square

Computing the exact degree of finite ambiguity, similar to the case for NFA [7], is a difficult problem:

Theorem 4. *Deciding whether $\text{da}(\mathcal{A}_{\text{NBA}}) > d$ for a given automaton \mathcal{A} and $d \in \mathbb{N}$ is a PSPACE-complete problem.*

Proof (sketch). We adapt the algorithm presented in [7] to show PSPACE-completeness of this problem from NFA to NBA. The nondeterministic algorithm for NFA guesses a word with at least $d+1$ accepting runs and evaluates the product of the transition matrices for each symbol along the word (which yields the number of different runs), while bounding the growth of the numbers. First we

show that we can restrict ourselves to ultimately periodic words of the form uv^ω with $u, v \in \Sigma^*$ and then we argue that it suffices to guess u and v accordingly. We show completeness by a simple reduction from the NFA to the NBA problem that maps each finite word w accepted by the NFA one-to-one to a word $w\#^\omega$ accepted by the NBA with the same number of accepting runs. \square

4 Translation from NBA to Finitely Ambiguous NBA

We present a construction that converts a given NBA with n states into a finitely ambiguous NBA with degree of ambiguity at most n . For explaining the intuition of the construction and for proving its correctness, we first introduce in Sect. 4.1 the notion of the reduced split tree as defined in [11]. This tree, defined for an NBA \mathcal{A} and an infinite word w , collects runs of \mathcal{A} on w in a specific way.

The translation of an NBA into a finitely ambiguous NBA is presented in Sect. 4.2. The construction uses two disjoint sets for tracking runs of the given NBA, and its description does not rely on the notion of reduced split tree. However, for understanding the role of the two subsets in the construction, reduced split trees are a valuable tool.

4.1 Reduced Split Trees

An X -labelled binary tree is a partial function $T : \{0, 1\}^* \rightarrow X$ such that the domain N_T of T is a non-empty prefix-closed set. The elements of N_T are called the nodes of T . The root node is ε (since N_T is prefix closed, $\varepsilon \in N_T$). For a node $u \in \{0, 1\}^*$, the node $u0$ is the *left child* of u , and $u1$ is the *right child* of u . For two nodes $u, v \in \{0, 1\}^*$ we say that u and v are on the same level if $|u| = |v|$, and we further say that u is to the left of v if u is lexicographically smaller than v .

An *infinite path* through such a tree corresponds to an infinite sequence $\pi \in \{0, 1\}^\omega$ (the nodes on the path are the finite prefixes of π). We say that π is *left-recurring* if it contains infinitely many 0 (which means that it moves to the left successor infinitely often).

Let $\mathcal{A} = (Q, \Sigma, \Delta, Q_0, F)$ be an NBA and $w \in \Sigma^\omega$ be an infinite word. The following definitions are illustrated by simple examples in Fig. 3(a)–(c).

The *split tree* $T_s^{\mathcal{A}, w}$ is a 2^Q -labelled binary tree with node set $N_T = \{0, 1\}^*$ defined as follows: The root is labelled with the set of initial states $T_s^{\mathcal{A}, w}(\varepsilon) := Q_0$. If $u \in \{0, 1\}^*$ with $|u| = i$ is labelled with $T_s^{\mathcal{A}, w}(u) = P \subseteq Q$, then $T_s^{\mathcal{A}, w}(u0) := \Delta_F(P, w(i))$ and $T_s^{\mathcal{A}, w}(u1) := \Delta_{\bar{F}}(P, w(i))$.

Observe that $T_s^{\mathcal{A}, w}$ is an infinite complete binary tree that encodes all runs of \mathcal{A} on w . It groups the runs by their visits to the set F by continuing runs that pass through an accepting state to the left, and the other ones to the right. Note that $T_s^{\mathcal{A}, w}$ can have nodes with label \emptyset , and that a state can occur in many labels on each level.

The *reduced left-right tree* $T_{rs}^{\mathcal{A}, w}$ is obtained from $T_s^{\mathcal{A}, w}$ by keeping for each state only the leftmost occurrence on each level (i.e., if u is to the left of v , and

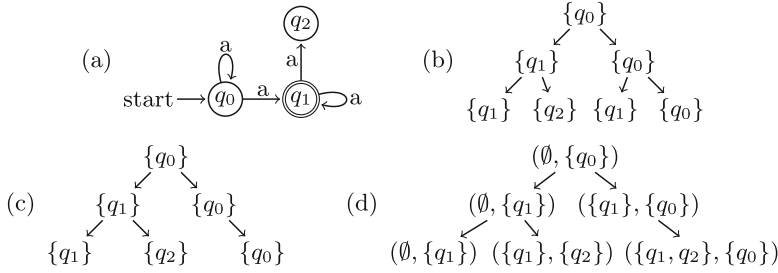


Fig. 3. (a) An NBA \mathcal{A} (b) First levels of $T_s^{\mathcal{A},w}$ for $w = a^\omega$ (c) First levels of $T_{rs}^{\mathcal{A},w}$ for $w = a^\omega$ (d) First two steps of possible transitions of \mathcal{A}'

q occurs in the labels of u and v , then it is removed from the label of v) and then removing vertices labelled with \emptyset . This results in a tree because if the label of a node becomes empty by the above operation, then also the labels of its successors become empty.

The interesting properties of $T_{rs}^{\mathcal{A},w}$ are summarized in the following lemma, which is shown in [11].

Lemma 6 ([11]).

1. The word w is accepted by \mathcal{A} if, and only if, $T_{rs}^{\mathcal{A},w}$ contains an infinite left-recurring path.
2. $T_{rs}^{\mathcal{A},w}$ contains at most $|Q|$ infinite paths.

The second claim is obvious because each state of \mathcal{A} is contained in at most one label of each level. The first claim is shown by picking a run of \mathcal{A} on w that always visits the next accepting states as early as possible (details can be found in [11]).

4.2 Construction of a Finitely Ambiguous NBA

We construct from \mathcal{A} a new NBA \mathcal{A}' whose infinite runs on a word w are in one-to-one correspondence with the infinite paths of $T_{rs}^{\mathcal{A},w}$. The run moves into a final state if the corresponding path branches to the left. Then Lemma 6 implies that \mathcal{A}' is equivalent to \mathcal{A} , and its degree of ambiguity is bounded by $|Q|$.

In order to implement this idea, \mathcal{A}' guesses an infinite path through $T_{rs}^{\mathcal{A},w}$ by constructing its label sequence. If the current label is S (corresponding to some node u in $T_{rs}^{\mathcal{A},w}$), then the labels S_0 and S_1 of $u0$ and $u1$ in $T_{rs}^{\mathcal{A},w}$ can be constructed based on the next input letter, the transition relation of \mathcal{A} , and the knowledge which states occur in the labels to the left of u in $T_{rs}^{\mathcal{A},w}$. This latter information is tracked in a second set P . Thus, in \mathcal{A}' a state (P, S) for $P, S \subseteq Q$ is reachable by reading the first i letters of an input word w if S is the label of some node u on level i in $T_{rs}^{\mathcal{A},w}$, and P is the union of the labels of nodes to the left of u (on the same level).

Formally, we define the new automaton in the following way. Given NBA \mathcal{A} , let $\mathcal{A}' = (Q', \Sigma, Q'_0, \Delta', F')$ be defined by

- $Q' = \{(P, S) \in 2^Q \times 2^Q \mid P \cap S = \emptyset \text{ and } S \neq \emptyset\}$
 - $Q'_0 = (\emptyset, Q_0)$
 - $\Delta' = \Delta'_0 \cup \Delta'_1$ with
 - $\Delta'_0 = \{((P, S), a, (P', S')) \mid P' = \Delta(P, a), S' = \Delta_F(S, a) \setminus P'\},$
 - $\Delta'_1 = \{((P, S), a, (P', S')) \mid P' = \Delta(P, a) \cup \Delta_F(S, a), S' = \Delta_{\overline{F}}(S, a) \setminus P'\}$
- Note that the transitions are only defined if $S' \neq \emptyset$ by definition of Q' .
- $F' = \{(P, S) \mid S \subseteq F\}$

Figure 3(d) illustrates the possible transitions of \mathcal{A}' .

The following lemma is a direct consequence of the definition of $T_{rs}^{\mathcal{A},w}$ and the construction of \mathcal{A}' (by an induction on i).

Lemma 7. *Let x be the finite prefix of length i of w , and assume that \mathcal{A}' can reach the state (P, S) by reading x from its initial state. Then there is a node u on level i of $T_{rs}^{\mathcal{A},w}$ with label S , and P is the union of the labels of the nodes to the left of u on level i .*

Lemma 7 implies that the infinite runs of \mathcal{A}' on w are in one-to-one correspondence with the infinite paths of $T_{rs}^{\mathcal{A},w}$ (and therefore \mathcal{A}' has degree of ambiguity at most $|Q|$). The final states of \mathcal{A}' are those in which the second component is a subset of F , and therefore correspond to a left successor on the corresponding path in $T_{rs}^{\mathcal{A},w}$. Hence, \mathcal{A}' has an accepting run of w if, and only if, $T_{rs}^{\mathcal{A},w}$ has a left-recurring path. By Lemma 6, this implies that $L(\mathcal{A}) = L(\mathcal{A}')$.

The number of disjoint pairs of subsets of Q is not larger than $3^{|Q|}$. In summary, we obtain the following result.

Theorem 5. *Let \mathcal{A} be an NBA with n states. Then there exists an automaton \mathcal{A}' with at most 3^n states accepting the same language such that $\text{da}(\mathcal{A}') \leq n$.*

An exponential lower bound for the construction of finitely ambiguous NBA can be inferred from a corresponding lower bound for NFA.

Theorem 6. *For each $n \in \mathbb{N}$ there exists an NBA with n states such that each finitely ambiguous NBA accepting the same language has at least $2^n - 1$ states.*

Proof (sketch). In [13] a family $\{\mathcal{A}_i\}_{i \in \mathbb{N}}$ of NFA is presented such that each \mathcal{A}_n has n states, and each at most polynomially ambiguous NFA equivalent to \mathcal{A}_n has $2^n - 1$ states. So the lower bound also holds for finitely ambiguous NFA. This lower bound can be lifted to NBA by considering the languages $L'_n = (L(\mathcal{A}_n)\#)^\omega$, for which one easily obtains n state NBA from the NFA \mathcal{A}_n . Furthermore, from a finitely ambiguous NBA for L'_n one can extract a finitely ambiguous NFA for $L(\mathcal{A}_n)$ with the same number of states. \square

Our construction is already close to this bound. As it is tracking just two subsets, it is also much simpler than the translations used to obtain unambiguous automata presented in [11, 12] that have upper bounds of $4(3n)^n$ and $n(0.76n)^n$ respectively, and are obtained at the cost of much more involved constructions. To the best of our knowledge, it is not known whether there is a stronger lower bound for the construction of unambiguous NBA than the one in Theorem 6.

5 Conclusion

In this paper we presented a fine classification for the ambiguity of nondeterministic Büchi automata by lifting results known for NFA and extending them to precisely capture the subtle differences in the case of infinite ambiguity. Finally we presented and discussed a translation from NBA to finitely ambiguous NBA. In future work we plan to investigate how this partial disambiguation can be applied in the setting of probabilistic model checking and look for cases in which using finitely ambiguous NBA could have an advantage over full disambiguation or determinisation.

References

1. Allauzen, C., Mohri, M., Rastogi, A.: General algorithms for testing the ambiguity of finite automata. In: Ito, M., Toyama, M. (eds.) DLT 2008. LNCS, vol. 5257, pp. 108–120. Springer, Heidelberg (2008). https://doi.org/10.1007/978-3-540-85780-8_8
2. Arnold, A.: Rational ω -languages are non-ambiguous. *Theor. Comput. Sci.* **26**(1–2), 221–223 (1983)
3. Baier, C., Katoen, J.: *Principles of Model Checking*. MIT Press, Cambridge (2008)
4. Baier, C., Kiefer, S., Klein, J., Klüppelholz, S., Müller, D., Worrell, J.: Markov chains and unambiguous Büchi automata. In: Chaudhuri, S., Farzan, A. (eds.) CAV 2016. LNCS, vol. 9779, pp. 23–42. Springer, Cham (2016). https://doi.org/10.1007/978-3-319-41528-4_2
5. Bousquet, N., Löding, C.: Equivalence and inclusion problem for strongly unambiguous Büchi automata. In: Dediu, A.-H., Fernau, H., Martín-Vide, C. (eds.) LATA 2010. LNCS, vol. 6031, pp. 118–129. Springer, Heidelberg (2010). https://doi.org/10.1007/978-3-642-13089-2_10
6. Büchi, J.R.: On a decision method in restricted second order arithmetic. In: *Studies in Logic and the Foundations of Mathematics*, vol. 44, pp. 1–11. Elsevier (1966)
7. Chan, T.H., Ibarra, O.H.: On the finite-valuedness problem for sequential machines. *Theor. Comput. Sci.* **23**(1), 95–101 (1988)
8. Courcoubetis, C., Yannakakis, M.: The complexity of probabilistic verification. *J. ACM* **42**(4), 857–907 (1995)
9. Couvreur, J.-M., Saheb, N., Sutre, G.: An optimal automata approach to LTL model checking of probabilistic systems. In: Vardi, M.Y., Voronkov, A. (eds.) LPAR 2003. LNCS (LNAI), vol. 2850, pp. 361–375. Springer, Heidelberg (2003). https://doi.org/10.1007/978-3-540-39813-4_26
10. Isaak, D., Löding, C.: Efficient inclusion testing for simple classes of unambiguous ω -automata. *Inf. Process. Lett.* **112**(14–15), 578–582 (2012)
11. Kähler, D., Wilke, T.: Complementation, disambiguation, and determinization of Büchi automata unified. In: Aceto, L., Damgård, I., Goldberg, L.A., Halldórsson, M.M., Ingólfssdóttir, A., Walukiewicz, I. (eds.) ICALP 2008. LNCS, vol. 5125, pp. 724–735. Springer, Heidelberg (2008). https://doi.org/10.1007/978-3-540-70575-8_59
12. Karmarkar, H., Joglekar, M., Chakraborty, S.: Improved upper and lower bounds for Büchi disambiguation. In: Van Hung, D., Ogawa, M. (eds.) ATVA 2013. LNCS, vol. 8172, pp. 40–54. Springer, Cham (2013). https://doi.org/10.1007/978-3-319-02444-8_5

13. Leung, H.: Separating exponentially ambiguous finite automata from polynomially ambiguous finite automata. *SIAM J. Comput.* **27**(4), 1073–1082 (1998)
14. Safra, S.: On the complexity of omega-automata. In: *Foundations of Computer Science 29th Annual Symposium on 1988*, pp. 319–327. IEEE (1988)
15. Stearns, R.E., Hunt III, H.B.: On the equivalence and containment problems for unambiguous regular expressions, regular grammars and finite automata. *SIAM J. Comput.* **14**(3), 598–611 (1985)
16. Thomas, W.: Languages, Automata, and Logic. In: Rozenberg, G., Salomaa, A. (eds.) *Handbook of Formal Languages*, pp. 389–455. Springer, Heidelberg (1997). https://doi.org/10.1007/978-3-642-59126-6_7
17. Weber, A., Seidl, H.: On the degree of ambiguity of finite automata. *Theor. Comput. Sci.* **88**(2), 325–349 (1991)