

# From LTL to Unambiguous Büchi Automata via Disambiguation of Alternating Automata

Simon Jantsch<sup>1</sup>, David Müller<sup>1</sup>, Christel Baier<sup>1</sup>, and Joachim Klein<sup>1</sup>

Technische Universität Dresden, Germany <sup>\*</sup>

**Abstract.** This paper proposes a new algorithm for the generation of unambiguous Büchi automata (UBA) from LTL formulas. Unlike existing tableau-based LTL-to-UBA translations, our algorithm deals with very weak alternating automata (VWAA) as an intermediate representation. It relies on a new notion of unambiguity for VWAA and a disambiguation procedure for VWAA. We introduce optimizations on the VWAA level and new LTL simplifications targeted at generating small UBA. We report on an implementation of the construction in our tool **Duggi** and discuss experimental results that compare the automata sizes and computation times of **Duggi** with the tableau-based LTL-to-UBA translation of the **SPOT** tool set. Our experiments also cover the analysis of Markov chains under LTL specifications, which is an important application of UBA.

## 1 Introduction

Translations from linear temporal logic (LTL) to non-deterministic Büchi automata (NBA) have been studied intensively as they are a core ingredient in the classical algorithmic approach to LTL model checking (see, e.g. [37,4,9]). In the worst case, such translations produce automata that are exponentially larger than the input formula. However, a lot of effort has been put into optimizing the general case, which has turned LTL-to-NBA translations feasible in practice. Two classes of algorithms have emerged as being especially well suited: tableau-based decomposition of the LTL formula into an automaton (see, e.g. [18,11]), as represented by the **SPOT** family of tools [15], and translations via very weak alternating automata (VWAA) [17], where **LTL3BA** [3] is the leading tool currently.

A property that has been studied in many areas of automata theory is *unambiguity* [10]. It allows non-deterministic branching but requires that each input word has at most one accepting run. Prominent cases in which unambiguity can be utilized include the universality check for automata (“Is every word accepted?”) on finite words, which is PSPACE-complete for arbitrary non-deterministic finite automata (NFA), but in P for unambiguous finite automata (UFA) [34]. Another

---

<sup>\*</sup> The authors are supported by the DFG through the Collaborative Research Centers CRC 912 (HAEC), the DFG grant 389792660 as part of TRR 248, the DFG-project BA-1679/12-1, the Cluster of Excellence EXC 2050/1 (CeTI, project ID 390696704, as part of Germany’s Excellence Strategy), and the Research Training Group QuantLA (GRK 1763).

example is model checking of Markov chains, which is in P if the specification is given as an unambiguous Büchi automaton (UBA) [5], and PSPACE-hard for arbitrary NBA [35]. Thus, using UBA leads to a single-exponential algorithm for LTL model checking of Markov chains, whereas using deterministic automata always involves a double-exponential lower bound in time complexity.

Every  $\omega$ -regular language is expressible by UBA [1], but NBA may be exponentially more succinct than UBA [23] and UBA may be exponentially more succinct than any deterministic automaton [7]. Universality and language inclusion are in P for subclasses of UBA [7,22], but the complexity is open for general UBA.

Although producing UBA was not the goal of the early translation from LTL to NBA by Vardi and Wolper [37], their construction is asymptotically optimal and produces *separated* automata, a subclass of UBA where the languages of the states are pairwise disjoint. Separated automata can express all  $\omega$ -regular languages [8], but UBA may be exponentially more succinct [7]. LTL-to-NBA translations have been studied intensively [17,18,16,13], but the generation of UBA from LTL formulas has not received much attention so far. We are only aware of three approaches targeted explicitly at generating UBA or subclasses. The first approach by Couvreur et al. [12] adapts the algorithm of [37], but still generates separated automata. LTL-to-UBA translations that attempt to exploit the advantages of UBA over separated automata have been presented by Benedikt et al. [6] and Duret-Lutz [14]. These adapt tableau-based LTL-to-NBA algorithms ([18] in the case of [6] and [11] in the case of [14]) and rely on transformations of the form  $\varphi \vee \psi \rightsquigarrow \varphi \vee (\neg\varphi \wedge \psi)$  to enforce that splitting disjunctive formulas generates states with disjoint languages, thus ensuring unambiguity.

To the best of our knowledge, the only available tool that supports the translation of LTL formulas to UBA is `ltl2tgba`, which is part of the `SPOT` tool set and implements the LTL-to-UBA algorithm of [14].

Proofs of all theorems and lemmas can be found in the appendix.

**Contribution.** We describe a novel LTL-to-UBA construction. It relies on an intermediate representation of LTL formulas using VWAA and adapts the known translation from VWAA to NBA by Gastin and Oddoux [17]. We introduce a notion of unambiguity for VWAA, show that the subsequent translation steps preserve it and that checking whether a VWAA is unambiguous is PSPACE-complete (Section 3). To the best of our knowledge, unambiguity for alternating automata has not been considered before.

We present a disambiguation procedure for VWAA that relies on intermediate unambiguity checks to identify ambiguous states and local disambiguation transformations for the VWAA (Section 4). It has the property that an already unambiguous VWAA is not changed. Figure 1 gives an overview of our LTL-to-UBA algorithm. Apart from the main construction, we introduce novel LTL rewrite rules and a heuristic, both of which are aimed at producing small UBA and may also benefit existing tools (see Figure 2). The heuristic is targeted at states with a certain structure, defined using the concepts of *purely-universal* and *alternating* formulas (Section 5). Finally, we report on an implementation of our construction in our tool `Duggi` and compare it to the existing LTL-to-UBA

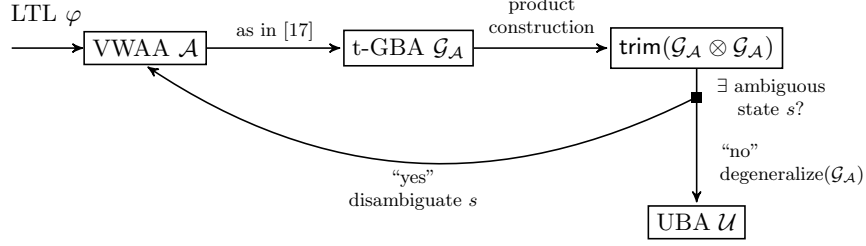


Fig. 1: The LTL-to-UBA step. A sequence of unambiguity checks and disambiguation transformations are applied and ultimately a UBA is returned. We use  $\text{trim}(\mathcal{G}_A \otimes \mathcal{G}_A)$  to check whether unambiguity is achieved or more iterations are necessary.

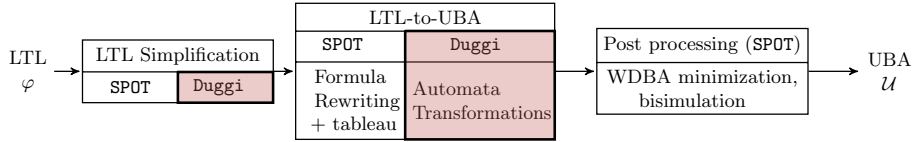


Fig. 2: Overview of the general LTL-to-UBA generation algorithm. The LTL simplification step, the actual LTL-to-UBA translation step, and the automaton post processing step can be combined freely. We propose novel rewriting rules for LTL and a LTL-to-UBA translation, both implemented in our tool *Duggi*.

translator *ltl2tgba*. We also compare *Duggi* with *ltl2tgba* in the context of Markov chain analysis under LTL specifications (Section 6).

## 2 Preliminaries

This section introduces our notation and standard definitions. The set of infinite words over a finite alphabet  $\Sigma$  is denoted by  $\Sigma^\omega$  and we write  $w[i]$  to denote the  $i$ -th position of an infinite word  $w \in \Sigma^\omega$ , and  $w[i..]$  to denote the suffix  $w[i]w[i+1]\dots$ . We write  $\mathcal{B}^+(X)$  to denote the set of positive Boolean formulas over a finite set of variables  $X$ . A *minimal model* of a formula  $f \in \mathcal{B}^+(X)$  is a set  $M \subseteq X$  such that  $M \models f$ , but no  $M' \subset M$  satisfies  $M' \models f$ . LTL is defined using  $\mathcal{U}$  (“Until”) and  $\bigcirc$  (“Next”). Additionally we use syntactical derivations  $\Diamond$  (“Finally”),  $\Box$  (“Globally”), and  $\mathcal{R}$  (“Release”) (see [4,19] for details).

**Alternating automata on infinite words.** An alternating  $\omega$ -automaton  $\mathcal{A}$  is a tuple  $(Q, \Sigma, \Delta, \iota, \Phi)$  where  $Q$  is a non-empty, finite set of states,  $\Sigma$  is a finite alphabet,  $\Delta : Q \times \Sigma \rightarrow \mathcal{B}^+(Q)$  is the transition function,  $\iota \in \mathcal{B}^+(Q)$  is the initial condition and  $\Phi$  is the acceptance condition. Additionally, we define the function  $\delta : Q \times \Sigma \rightarrow 2^{2^Q}$  which assigns to a pair  $(q, a) \in Q \times \Sigma$  the set of minimal models of  $\Delta(q, a)$  and the set  $I \subseteq 2^Q$  as the set of minimal models of  $\iota$ . We denote

by  $\mathcal{A}(\iota')$  the automaton  $(Q, \Sigma, \delta, \iota', \Phi)$  and we write  $\mathcal{A}(Q_0)$  for  $\mathcal{A}(\bigwedge_{q \in Q_0} q)$ , if  $Q_0 \subseteq Q$ . We call the number of the reachable states of an automaton  $\mathcal{A}$  its size.

A run of  $\mathcal{A}$  for  $w \in \Sigma^\omega$  is a directed acyclic graph (DAG)  $(V, E)$  [27], where

1.  $V \subseteq Q \times \mathbb{N}$ , and  $E \subseteq \bigcup_{0 \leq l} (Q \times \{l\}) \times (Q \times \{l+1\})$ ,
2.  $\{q : (q, 0) \in V\} \in I$ ,
3. for all  $(q, l) \in V : \{q' : ((q, l), (q', l+1)) \in E\} \in \delta(q, w[l])$ ,
4. for all  $(q, l) \in V \setminus (Q \times \{0\})$  there is a  $q'$  such that  $((q', l-1), (q, l)) \in E$ .

We define  $V(i) = \{s : (s, i) \in V\}$ , called the  $i$ -th *layer* of  $V$ . A run is called accepting if every infinite path in it meets the acceptance condition.

A word is accepted by  $\mathcal{A}$  if there exists an accepting run for it. We denote the set of accepted words of  $\mathcal{A}$  by  $\mathcal{L}(\mathcal{A})$ . We distinguish between Büchi, generalized Büchi and co-Büchi acceptance conditions. A Büchi condition is denoted by  $\text{Inf}(Q_f)$  for a set  $Q_f \subseteq Q$ . An infinite path  $\pi = q_0 q_1 \dots$  meets  $\text{Inf}(Q_f)$  if  $Q_f \cap \text{inf}(\pi) \neq \emptyset$ , where  $\text{inf}(\pi)$  denotes the set of infinitely occurring states in  $\pi$ . A co-Büchi condition is denoted by  $\text{Fin}(Q_f)$  and  $\pi$  meets  $\text{Fin}(Q_f)$  if  $Q_f \cap \text{inf}(\pi) = \emptyset$ . An infinite path  $\pi$  meets a generalized Büchi condition  $\bigwedge_{i \in F} \text{Inf}(Q_i)$  if it meets  $\text{Inf}(Q_i)$  for all  $i \in F$ . A *transition-based* acceptance condition uses sets of transitions  $T \subseteq Q \times \Sigma \times Q$  instead of sets of states to define acceptance of paths.

We call a subset  $C \subseteq Q$  a configuration and say that  $C$  is reachable if it is a layer of some run. A configuration  $C$  is reachable from a state  $q$ , also written as  $q \xrightarrow{*} C$ , if  $C$  is a reachable configuration of  $\mathcal{A}(q)$ . Analogously,  $C' \subseteq Q$  is reachable from  $C \subseteq Q$ , or  $C \xrightarrow{*} C'$ , if  $C'$  is a reachable configuration of  $\mathcal{A}(C)$ . A configuration  $C$  is *reachable via*  $u$  if there is a run  $(V, E)$  for a word  $uw$ , with  $u \in \Sigma^*$ ,  $w \in \Sigma^\omega$ , such that  $C = V(|u|)$ . We extend this notion to reachability from states and configurations via finite words in the expected way and write  $q \xrightarrow{u} C$  and  $C \xrightarrow{u} C'$ . We define  $\mathcal{L}(C) = \mathcal{L}(\mathcal{A}(C))$ .

The *underlying graph* of  $\mathcal{A}$  has vertices  $Q$  and edges  $\{(q, q') : \exists a \in \Sigma. \exists S \in \delta(q, a). q' \in S\}$ . We say that  $\mathcal{A}$  is *very weak* if every strongly connected component of its underlying graph consists of a single state and  $\mathcal{A}$  has a co-Büchi acceptance. If  $|C_0| = 1$  for every  $C_0 \in I$ , and  $|C_\delta| = 1$  for every  $C_\delta \in \delta(q, a)$  with  $(q, a) \in Q \times \Sigma$ , we call  $\mathcal{A}$  non-deterministic. As a non-deterministic automaton has only singleton successor sets, its runs are infinite sequences of states. Finally, an automaton  $\mathcal{A}$  is *trimmed* if  $\mathcal{L}(q) \neq \emptyset$  holds for every state  $q$  in  $\mathcal{A}$ , and we write  $\text{trim}(\mathcal{A})$  for the automaton that we get by removing all states with empty language in  $\mathcal{A}$ . For the non-alternating automata types that we consider,  $\text{trim}(\mathcal{A})$  can be computed in linear time using standard graph algorithms.

**From LTL to NBA.** We use the standard translation from LTL to VWAA where the states of the VWAA correspond to subformulas of  $\varphi$  and the transition relation follows the Boolean structure of the state and the LTL expansion laws [36,30]. It has been used as a first step in an LTL-to-NBA translation in [17], whose construction we follow. We recall this construction in Appendix A. Additionally, we use the optimizations proposed in [3]. We also maintain the following invariant, as proposed in [17]: for all  $(q, a) \in Q \times \Sigma$  and successor sets  $S_1, S_2 \in \delta(q, a)$ , such that  $S_1 \neq S_2$ , it holds that  $S_1 \not\subseteq S_2$ .

A VWAA  $\mathcal{A}$  can be transformed into a transition-based generalized Büchi automaton (t-GBA) by a powerset-like construction, where the non-deterministic choices of  $\mathcal{A}$  are captured by non-deterministic choices of the t-GBA, and the universal choices are captured by the powerset.

**Definition 1.** Let  $\mathcal{A} = (Q, \Sigma, \Delta, \iota, \text{Fin}(Q_f))$  be a VWAA. The t-GBA  $\mathcal{G}_{\mathcal{A}}$  is the tuple  $(2^Q, \Sigma, \delta', I, \bigwedge_{f \in Q_f} \text{Inf}(\mathcal{T}_f))$ , where

- $\delta'(C, a) = \bigotimes_{q \in C} \delta(q, a)$ , where  $T_1 \otimes T_2 = \{C_1 \cup C_2 : C_1 \in T_1, C_2 \in T_2\}$
- $\mathcal{T}_f = \{(C, a, C') : f \notin C' \text{ or there exists } Y \in \delta(f, a) \text{ and } f \notin Y \subseteq C'\}$

**Theorem 2 (Theorem 2 of [17]).** Let  $\mathcal{A}$  be a VWAA and  $\mathcal{G}_{\mathcal{A}}$  be as in Definition 1. Then,  $\mathcal{L}(\mathcal{A}) = \mathcal{L}(\mathcal{G}_{\mathcal{A}})$ .

The size of  $\mathcal{G}_{\mathcal{A}}$  may be exponential in  $|Q|$  and the number of Büchi conditions of  $\mathcal{G}_{\mathcal{A}}$  is  $|Q_f|$ . Often a Büchi automaton with a (non-generalized) Büchi acceptance is desired. For this step we follow the construction of [17], which translates  $\mathcal{G}_{\mathcal{A}}$  into an NBA  $\mathcal{N}_{\mathcal{G}_{\mathcal{A}}}$  of at most  $|Q_f| \cdot 2^{|Q|}$  reachable states.

### 3 Unambiguous VWAA

In this section we introduce a notion of unambiguity for VWAA and show that unambiguous VWAA are translated to UBA by the translation presented in Section 2. We define unambiguity in terms of configurations of the VWAA, which are strongly related to the states of the resulting NBA. Let  $\mathcal{A} = (Q, \Sigma, \Delta, \iota, \text{Fin}(Q_f))$  be a fixed VWAA for the rest of this section.

**Definition 3.**  $\mathcal{A}$  is unambiguous if it has no distinct configurations  $C_1, C_2$  that are reachable via the same word  $u \in \Sigma^*$  and such that  $\mathcal{L}(C_1) \cap \mathcal{L}(C_2) \neq \emptyset$ .

The standard definition of unambiguity is that an automaton is unambiguous if it has at most one accepting run for any word. In our setting runs are DAG's and we do allow multiple accepting runs for a word, as long as they agree on the configurations that they reach for each prefix. In this sense it is a weaker notion. However, the notions coincide on non-deterministic automata as the edge relation of the run is then induced by the sequence of visited states.

**Theorem 4.** Let  $\mathcal{N}_{\mathcal{G}_{\mathcal{A}}}$  be the NBA for  $\mathcal{A}$ , obtained by the translation from Section 2. If  $\mathcal{A}$  is unambiguous, then  $\mathcal{N}_{\mathcal{G}_{\mathcal{A}}}$  is unambiguous.

We show that every step in the translation from VWAA to NBA preserves unambiguity. First, we establish the following correspondance:

**Lemma 5.** If  $\mathcal{A}$  is unambiguous, then for every accepting run  $r = Q_0 Q_1 \dots$  of  $\mathcal{G}_{\mathcal{A}}$  for  $w \in \Sigma^\omega$  there exists an accepting run  $\rho = (V, E)$  of  $\mathcal{A}$  for  $w$  such that  $Q_i = V(i)$  for all  $i \geq 0$ .

Table 1: The adapted expansion laws for  $\mathcal{U}$  and  $\mathcal{R}$  are the result of applying the disjunction rule to the classic expansion laws.

	expansion law	adapted expansion law
$\varphi \mathcal{U} \psi$	$\Gamma \equiv \psi \vee (\varphi \wedge \bigcirc \Gamma)$	$\Gamma \equiv \psi \vee (\varphi \wedge \neg \psi \wedge \bigcirc \Gamma)$
$\varphi \mathcal{R} \psi$	$\Gamma \equiv \psi \wedge (\varphi \vee \bigcirc \Gamma)$	$\Gamma \equiv \psi \wedge (\varphi \vee (\neg \varphi \wedge \bigcirc \Gamma))$

Intuitively, the lemma states that if  $\mathcal{A}$  is unambiguous, then every accepting run  $r$  of  $\mathcal{G}_{\mathcal{A}}$  can be matched by an accepting run  $\rho$  of  $\mathcal{A}$  such that the states of  $r$  are the layers of  $\rho$ . The proof is not immediate and requires  $\mathcal{A}$  to be unambiguous.

A direct consequence of Lemma 5 is that if  $\mathcal{A}$  is unambiguous, then so is  $\mathcal{G}_{\mathcal{A}}$ . The degeneralization construction in [17] makes  $|Q_f| + 1$  copies of  $\mathcal{G}_{\mathcal{A}}$ . As the next copy is uniquely determined by the current state and word label, it preserves unambiguity. In combination with Lemma 7 we obtain Theorem 4.

We now show that deciding whether a VWAA is unambiguous is PSPACE-complete. The idea for proving hardness is to reduce LTL satisfiability, which is known to be PSPACE-hard, to VWAA emptiness (this follows directly by the  $\text{LTL} \rightarrow \text{VWAA}$  translation) and VWAA emptiness to VWAA unambiguity. The second step uses the following trick: a VWAA  $\mathcal{A}$  accepts the empty language if and only if the disjoint union of  $\mathcal{A}$  with itself is unambiguous.

To check whether VWAA is unambiguous we first show that for every accepting run of  $\mathcal{A}$ , we find a matching accepting run of  $\mathcal{G}_{\mathcal{A}}$ , which follows directly from the definition of  $\mathcal{G}_{\mathcal{A}}$ :

**Lemma 6.** *For every accepting run  $\rho = (V, E)$  of  $\mathcal{A}$  for  $w \in \Sigma^\omega$  there exists an accepting run  $r = Q_0 Q_1 \dots$  of  $\mathcal{G}_{\mathcal{A}}$  for  $w$ , such that  $Q_i = V(i)$  for all  $i \geq 0$ .*

Lemma 5 and Lemma 6 give us the following:

**Lemma 7.**  *$\mathcal{A}$  is unambiguous if and only if  $\mathcal{G}_{\mathcal{A}}$  is unambiguous.*

However, checking whether  $\mathcal{G}_{\mathcal{A}}$  is unambiguous can be done in space polynomial in the size of  $\mathcal{A}$ , and we conclude:

**Theorem 8.** *Deciding whether a VWAA is unambiguous is PSPACE-complete.*

## 4 Disambiguating VWAA

Our disambiguation procedure is inspired by the idea of “separating” the language of successors for every non-deterministic branching. A disjunction  $\varphi \vee \psi$  is transformed into  $\varphi \vee (\neg \varphi \wedge \psi)$  by this principle. The rules for  $\mathcal{U}$  and  $\mathcal{R}$  are derived by applying the disjunction rule to the expansion law of the corresponding operator (see Table 1). These rules are applied by `1t12tgba` in its tableau-based algorithm to guarantee that the resulting automaton is unambiguous, and have also been proposed in [6].

In our approach we define corresponding transformations for non-deterministic branching in the VWAA. Furthermore, we propose to do this in an “on-demand” manner: instead of applying these transformation rules to every non-deterministic split, we identify ambiguous states during the translation and only apply the transformations to them. This guarantees that we return the automaton produced by the core translation, without disambiguation, in case it is already unambiguous.

The main steps of our disambiguation procedure are the following:

1. A preprocessing step that computes a complement state  $\tilde{s}$  for every state  $s$ .
2. A procedure that identifies ambiguous states.
3. Local transformations that remove the ambiguity.

If no ambiguity is found in step 2, the VWAA is unambiguous. The high-level overview is also depicted in Figure 1. In what follows we fix a VWAA  $\mathcal{A} = (Q, \Sigma, \Delta, \iota, \text{Fin}(Q_f))$  and assume that it has a single initial state.

**Complement states.** The transformations we apply for disambiguation rely on the following precondition: for every state  $s$  of  $\mathcal{A}$  there should be another state  $\tilde{s}$  such that  $\mathcal{L}(\tilde{s}) = \overline{\mathcal{L}(s)}$ . We compute these complement states in a preprocessing step and add them to  $\mathcal{A}$ . Complementing alternating automata can be done without any blow up by dualizing both the acceptance condition and transition structure, as shown by Muller and Schupp [31]. As dualizing the acceptance condition and complementing the set of final states yields an equivalent VWAA, we can keep the co-Büchi acceptance when complementing.

The complement automaton has the same underlying graph and is therefore also very weak. Furthermore, no state  $s$  is reachable from its own complement state  $\tilde{s}$ , which is an invariant that we maintain and which ensures that very weakness is preserved in the construction.

**Source configurations and source states.** To characterize ambiguous situations we define *source configurations* and *source states*. A source configuration of  $\mathcal{A}$  is a reachable configuration  $C$  such that there exist two different configurations  $C_1, C_2$  that are reachable from  $C$  via some  $a \in \Sigma$  and  $\mathcal{L}(C_1) \cap \mathcal{L}(C_2) \neq \emptyset$ . By definition,  $\mathcal{A}$  is not unambiguous if a source configuration exists.

Let  $C$  be a source configuration of  $\mathcal{A}$  and let  $C_1, C_2$  be the successor configurations as described above. A source state of  $C$  is a state  $s \in C$  with two transitions  $S_1, S_2 \in \delta(s, a)$  such that  $S_i \subseteq C_i$ , for  $i \in \{1, 2\}$ ,  $S_1 \neq S_2$  and  $(S_1 \cup S_2) \setminus (C_1 \cap C_2) \neq \emptyset$ . The last condition ensures that either  $S_1$  or  $S_2$  contains a state that is not common to  $C_1$  and  $C_2$ . By Definition 1,  $C_i = \bigcup_{q \in C} S_q$  with  $S_q \in \delta(a, q)$  for all  $q \in C$ , and thus  $C$  must contain a source state.

**Ambiguity check and finding source states.** For the analysis of source configurations and source states we use the standard product construction  $\mathcal{G}_1 \otimes \mathcal{G}_2$ , which returns a t-GBA such that  $\mathcal{L}(\mathcal{G}_1 \otimes \mathcal{G}_2) = \mathcal{L}(\mathcal{G}_1) \cap \mathcal{L}(\mathcal{G}_2)$  for two given t-GBA  $\mathcal{G}_1$  and  $\mathcal{G}_2$ . Specifically, we consider the self product  $\mathcal{G}_{\mathcal{A}} \otimes \mathcal{G}_{\mathcal{A}}$  of  $\mathcal{G}_{\mathcal{A}}$ . It helps to identify ambiguity:  $\mathcal{G}_{\mathcal{A}}$  is not unambiguous if and only if there exists a reachable state  $(C_1, C_2)$  in  $\text{trim}(\mathcal{G}_{\mathcal{A}} \otimes \mathcal{G}_{\mathcal{A}})$  with  $C_1 \neq C_2$ .

The pair of configurations  $(C_1, C_2)$  is a witness to ambiguity of  $\mathcal{A}$ . We look for a symbol  $a \in \Sigma$  and a configuration  $C$  such that  $(C, C) \xrightarrow{a} (C'_1, C'_2) \rightarrow^* (C_1, C_2)$

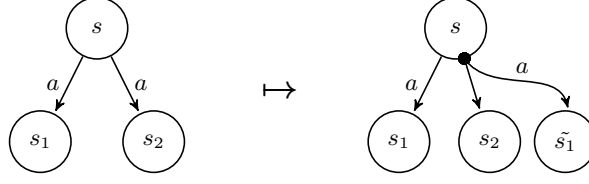


Fig. 3: Disambiguation scheme for a source state  $s$  with successors  $s_1$  and  $s_2$  in the VWAA. Transitions with successor set of size  $\geq 1$  are conjoined by a  $\bullet$ .

is a path in  $\text{trim}(\mathcal{G}_{\mathcal{A}} \otimes \mathcal{G}_{\mathcal{A}})$  and  $C'_1 \neq C'_2$ . Such a configuration must exist as we have assumed that  $\mathcal{A}$  has a single initial state  $q_i$ , which implies that  $\text{trim}(\mathcal{G}_{\mathcal{A}} \otimes \mathcal{G}_{\mathcal{A}})$  has a single initial state  $(\{q_i\}, \{q_i\})$ .  $C$  is a source configuration and therefore must contain a source state which we can find by inspecting all pairs of transitions of states in  $C$ .

**Disambiguating a source state.** The general scheme for disambiguating source states is depicted in Figure 3. Assume that we have identified a source state  $s$  with successor sets  $S_1$  and  $S_2$  as explained above. The LTL-to-VWAA construction guarantees  $S_1 \not\subseteq S_2$  and  $S_2 \not\subseteq S_1$ . We need to distinguish the looping successor sets (i.e. those  $S_i$  that contain  $s$ ) from the non-looping. Technically, we consider two cases: either  $S_1$  or  $S_2$  do not contain  $s$  or both sets contain  $s$ . In the first case we assume, w.l.o.g., that  $s \notin S_1$ . The successor set  $S_2$  is split into the  $|S_1|$  new successor sets  $\{(S_2 \cup \{\tilde{s}_1\}) : s_1 \in S_1\}$ . The new sets of states are added to  $\delta(s, a)$  and the successor set  $S_2$  is removed. If both  $S_1$  and  $S_2$  contain  $s$ , we proceed as in the first case but do not add the successor set  $S_2 \cup \{\tilde{s}\}$  to  $\delta(s, a)$ .

This transformation does not guarantee that  $s$  is not a source state anymore. However, it removes the ambiguity that stems from the non-deterministic choice of transitions  $S_1, S_2 \in \delta(a, s)$ . If  $s$  is still a source state it will be identified again for another pair of transitions. After a finite number of iterations all successor sets of  $s$  for any symbol in  $\Sigma$  will accept pairwise disjoint languages, in which case  $s$  cannot be a source state anymore. The transformation preserves very weakness as it only adds transitions from  $s$  to complement states of successors of  $s$  and by assumption there is no path between a state and its complement state.

**Iterative algorithm.** Putting things together, our algorithm works as follows: it searches for source configurations of  $\mathcal{A}$  (using  $\mathcal{G}_{\mathcal{A}}$ ), applies the local disambiguation transformations to  $\mathcal{A}$  as described and recurses (see Figure 1). As rebuilding the t-GBA may become costly, in our implementation we identify which part of the t-GBA has to be recomputed due to the changes in  $\mathcal{A}$ , and rebuild only this part. If no source configuration is found, we know that both  $\mathcal{A}$  and  $\mathcal{G}_{\mathcal{A}}$  are unambiguous and we can apply degeneralization to obtain a UBA.

**Complexity of the procedure.** The VWAA-to-t-GBA translation that we adapt produces a t-GBA  $\mathcal{G}_{\mathcal{A}}$  of size at most  $2^n$  for a VWAA  $\mathcal{A}$  of size  $n$ . In our disambiguation procedure we enlarge  $\mathcal{A}$  by adding complement states for every state in the original automaton, yielding a VWAA of size  $2n$ . Thus, a first



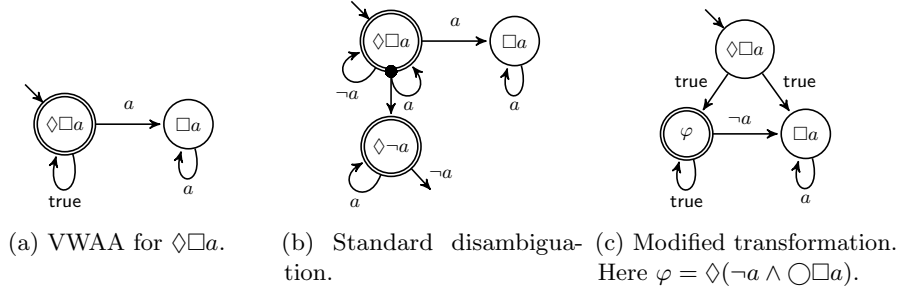


Fig. 4: Three VWAA for  $\Diamond \Box a$ . The automaton in (b) is the result of standard disambiguation and (c) is the result of the modified transformation applied to (a). The automaton in (c) is non-deterministic and has two looping states, whereas (b) is not non-deterministic and has three looping states.

size estimate of  $\mathcal{G}_{\mathcal{A}}$  in our construction is  $4^n$ . However, no state in  $\text{trim}(\mathcal{G}_{\mathcal{A}})$  can contain both  $s$  and  $\tilde{s}$  for any state  $s$  of  $\mathcal{A}$ . The reason is that the language of a state in  $\mathcal{G}_{\mathcal{A}}$  is the intersection of the languages of the VWAA-states it contains, and  $\mathcal{L}(s) \cap \mathcal{L}(\tilde{s}) = \emptyset$ . Thus,  $\text{trim}(\mathcal{G}_{\mathcal{A}})$  has at most  $3^n$  states.

The amount of ambiguous situations that we identify is bounded by the number of non-deterministic splits in the VWAA, which may be exponential in the length of the input LTL formula. In every iteration we check ambiguity of the new VWAA, which can be done in exponential time. Thus, our procedure computes a UBA in time exponential in the length of the formula.

## 5 Heuristics for purely-universal formulas

In this section we introduce alternative disambiguation transformations for special source states representing formulas  $\varphi \mathcal{U} \nu$ , where  $\nu$  is *purely-universal*. The class of purely-universal formulas is a syntactically defined subclass of LTL-formulas with suffix-closed languages. These transformations reduce the size of the resulting UBA and often produce automata of a simpler structure. The idea is to decide whether  $\nu$  holds whenever moving to a state representing  $\varphi \mathcal{U} \nu$  and, if not, finding the *last* position where it does not hold.

*Example 9.* Consider the formula  $\Diamond \Box a$ . A VWAA for it is shown in Figure 4a. It is ambiguous, as a word satisfying  $\Box a$  may loop in the initial state for an arbitrary amount of steps before moving to the next state.

In the standard disambiguation transformation the state  $\Diamond \neg a$  is added to the self loop of the initial state (Figure 4b). The automaton in Figure 4c, on the other hand, makes the following case distinction: either a word satisfies  $\Box a$ , in which case we move to that state directly, or there is a suffix that satisfies  $\neg a$  and  $\bigcirc \Box a$ . The state  $\varphi$  is used to find the *last* occurrence of  $\neg a$ , which is unique.

To generalize this idea and identify the situations where it is applicable we use the syntactically defined subclasses of *purely-universal* ( $\nu$ ), *purely-eventual*

( $\mu$ ) and *alternating* ( $\xi$ ) formulas ([16,3]). In the following definition  $\varphi$  ranges over arbitrary LTL formulas:

$$\begin{aligned}\nu &::= \Box\varphi \mid \nu \vee \nu \mid \nu \wedge \nu \mid \bigcirc\nu \mid \nu\mathcal{U}\nu \mid \varphi\mathcal{R}\nu \mid \Diamond\nu \\ \mu &::= \Diamond\varphi \mid \mu \vee \mu \mid \mu \wedge \mu \mid \bigcirc\mu \mid \varphi\mathcal{U}\mu \mid \mu\mathcal{R}\mu \mid \Box\mu \\ \xi &::= \Box\mu \mid \Diamond\nu \mid \xi \vee \xi \mid \xi \wedge \xi \mid \bigcirc\xi \mid \varphi\mathcal{U}\xi \mid \varphi\mathcal{R}\xi \mid \Diamond\xi \mid \Box\xi\end{aligned}$$

Formulas that fall into these classes define suffix closed ( $\nu$ ), prefix closed ( $\mu$ ) and prefix invariant ( $\xi$ ) languages respectively:

**Lemma 10** ([16,3]). *For all  $u \in \Sigma^*$  and  $w \in \Sigma^\omega$ :*

- *If  $\nu$  is purely-universal, then  $uw \models \nu \implies w \models \nu$ .*
- *If  $\mu$  is purely-eventual, then  $w \models \mu \implies uw \models \mu$ .*
- *If  $\xi$  is alternating, then  $w \models \xi \iff uw \models \xi$ .*

Let  $\nu$  be purely-universal. We want to find a formula  $\mathbf{g}(\nu)$ , called the *goal* of  $\nu$ , that is simpler than  $\nu$  and satisfies  $\mathbf{g}(\nu) \wedge \bigcirc\nu \equiv \nu$ . If  $\nu$  does not hold initially for some word  $w$  we can identify the last suffix  $w[i..]$  where it does not hold, given that such an  $i$  exists, by checking if  $w[i..]$  satisfies  $\neg\mathbf{g}(\nu) \wedge \bigcirc\nu$ .

It is not clear how to define  $\mathbf{g}(\nu)$  for purely-universal formulas of the form  $\nu_1 \vee \nu_2$  or  $\nu_1\mathcal{U}\nu_2$ . We therefore introduce the concept of *disjunction-free* purely-universal formulas in which all occurrences of  $\vee$  and  $\mathcal{U}$  appear in the scope of some  $\Box$ . As  $\varphi\mathcal{R}\nu \equiv \nu$  if  $\nu$  is purely-universal, we assume that all occurrences of  $\mathcal{R}$  are also in the scope of some  $\Box$  for purely-universal formulas.

**Lemma 11.** *Every purely-universal formula  $\nu$  can be rewritten into a formula  $\nu_1 \vee \dots \vee \nu_n$ , where  $\nu_i$  is disjunction-free for all  $1 \leq i \leq n$ .*

Disjunction-free purely-universal formulas have a natural notion of “goal”.

**Definition 12.** *Let  $\nu$  be a disjunction-free and purely-universal formula. We define  $\mathbf{g}(\nu)$  inductively as follows:*

$$\begin{aligned}\mathbf{g}(\Box\varphi) &= \varphi & \mathbf{g}(\bigcirc\nu) &= \bigcirc\mathbf{g}(\nu) \\ \mathbf{g}(\nu_1 \wedge \nu_2) &= \mathbf{g}(\nu_1) \wedge \mathbf{g}(\nu_2) & \mathbf{g}(\Diamond\nu) &= \text{true}\end{aligned}$$

The reason for defining  $\mathbf{g}(\Diamond\nu)$  as **true** is that  $\Diamond\nu$  is an alternating formula and checking its validity can thus be temporarily suspended. Indeed, the definition satisfies the equivalence that we aimed for:

**Lemma 13.** *Let  $\nu$  be a disjunction-free and purely-universal formula. Then  $\mathbf{g}(\nu) \wedge \bigcirc\nu \equiv \nu$ .*

In Example 9  $\neg\mathbf{g}(\nu) \wedge \bigcirc\nu$  corresponds to  $\neg a \wedge \bigcirc\Box a$ , which is realized by the transition from state  $\varphi$  to state  $\Box a$  in Figure 4c.

Lemma 14 shows the general transformation scheme (applied left to right). It introduces non-determinism, but we show that it is not a cause of ambiguity as the languages of the two disjuncts are disjoint. An important difference to the known rule for  $\mathcal{U}$  is that the left-hand side of the  $\mathcal{U}$ -formula stays unchanged. This is favorable as it is the left-hand side that may introduce loops in the automaton.

**Lemma 14.** *Let  $\nu$  be a disjunction-free and purely-universal formula. Then*

$$1. \varphi\mathcal{U}(\nu \vee \psi) \equiv \nu \vee \gamma \quad \text{and} \quad 2. \mathcal{L}(\nu) \cap \mathcal{L}(\gamma) = \emptyset$$

where  $\gamma = \varphi\mathcal{U}((\varphi \wedge \neg \mathbf{g}(\nu) \wedge \bigcirc \nu) \vee (\psi \wedge \neg \nu))$ .

LTL formulas may become larger when applying this transformation. However, they are comparable to the LTL formulas produced by the standard disambiguation transformations in terms of the number of subformulas. If all occurrences of  $\bigcirc$  in  $\nu$  are in the scope of some  $\Box$ , then no subformulas are added. Otherwise,  $\mathbf{g}(\nu)$  and  $\bigcirc \nu$  may introduce new  $\bigcirc$ -subformulas.

## 6 Implementation and Experiments

The tool **Duggi** is an LTL-to-UBA translator based on the construction introduced in the foregoing sections.<sup>1</sup> It reads LTL formulas in a prefix syntax and produces (unambiguous) automata in the HOA format [2]. In the implementation we deviate from or extend the procedure described above in the following ways:

- We make use of the knowledge given by the VWAA-complement states in the translation steps to t-GBA  $\mathcal{G}_{\mathcal{A}}$  and the product  $\mathcal{G}_{\mathcal{A}} \otimes \mathcal{G}_{\mathcal{A}}$ . It allows an easy emptiness check: if  $s$  and  $\tilde{s}$  are present in some  $\mathcal{G}_{\mathcal{A}}$  or  $\mathcal{G}_{\mathcal{A}} \otimes \mathcal{G}_{\mathcal{A}}$  state, then it accepts the empty language and does not have to be further expanded.
- We have included the following optimization of the LTL-to-VWAA procedure: when translating a formula  $\Box \mu$ , where  $\mu$  is purely-eventual, we instead translate  $\Box \bigcirc \mu$ . This results in an equivalent state with fewer transitions. It is close to the idea of suspension as introduced in [3], but is not covered by it.
- Additionally, **Duggi** features an LTL rewriting procedure that uses many of the LTL simplification rules in the literature [33,16,3,29]. We have included the following rules that are not used by **SPOT**:

$$\text{I } (\Box \Diamond \varphi) \wedge (\Diamond \Box \psi) \mapsto \Box \Diamond (\varphi \wedge \Box \psi) \qquad \text{II } (\Diamond \Box \varphi) \vee (\Box \Diamond \psi) \mapsto \Diamond \Box (\varphi \vee \Diamond \psi)$$

These rewrite rules are more likely to produce formulas of the form  $\Diamond \Box \varphi$ , to which the heuristic of Section 5 can be applied. They stem from [29], where the reversed rules have been used to achieve a normal form.

**LTL benchmarks from the literature.** We now compare the UBA sizes for LTL formulas of the benchmark set LTLSTORE [24]. It collects formulas from various case studies and tool evaluation papers in different contexts. We include the negations of all formulas and filter out duplicates, leaving 1419 formulas.

Languages that are recognizable by *weak deterministic Büchi automata* (WDBA) can be efficiently minimized [26] and **ltl2tgba** applies this algorithm as follows: it computes the minimal deterministic Büchi automaton and the UBA

<sup>1</sup> **Duggi** and the **PRISM** implementation, together with all experimental data, are available at <https://www.tcs.inf.tu-dresden.de/ALGI/TR/FM19-UBA/>.

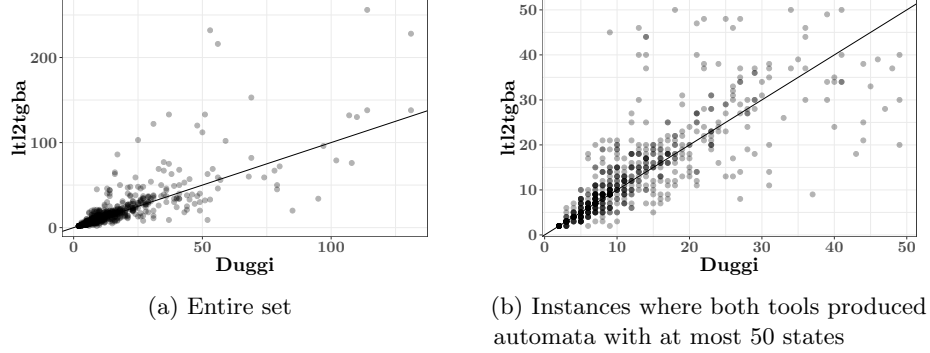


Fig. 5: Non-WDBA-recognizable fragment of LTLSTORE (948 formulas). Every point stands for a formula where the according automaton size for **Duggi** is the abscissa, the automaton size of **1t12tgba** the ordinate. Points above the line stand for formulas where **Duggi** performed better.

and returns the one with fewer states. Our formula set contains 472 formulas that are WDBA-recognizable and for which we could compute the minimal WDBA within the bounds of 30 minutes and 10 GB of memory using **1t12tgba**. Of these 472 formulas we found 11 for which the UBA generated by either **Duggi** or **1t12tgba** was smaller than the minimal WDBA, and only two where the difference was bigger than 3 states. On the other hand, the minimal WDBA were smaller than the UBA produced by **1t12tgba** (**Duggi**) for 164 (203) formulas. This supports the approach by **1t12tgba** to apply WDBA minimization when possible and in what follows we focus on the fragment of the LTLSTORE that does not fall into this class. In [14] it was noted that WDBA minimization often leads to smaller automata than the LTL-to-NBA translation of **1t12tgba**.

We consider the following configurations: **Duggi** is the standard configuration, **Duggi**<sub>\(R,H\)</sub> is **Duggi** without the new rewrite rules I and II (R) and/or without the heuristic introduced in Section 5 (H). For **SPOT**, **1t12tgba** is the standard configuration that produces UBA without WDBA-minimization, which is switched on in **1t12tgba**<sub>WDBA</sub>. We use simulation-based postprocessing as provided by **SPOT** in all **Duggi**-configurations (they are enabled by default in **1t12tgba**). We use **SPOT** with version 2.7.2. All computations, including the PMC experiments, were performed on a computer with two Intel E5-2680 8 cores at 2.70 GHz running Linux, with a time bound of 30 minutes and a memory bound of 10 GB.

Scatter plots comparing the number of states of UBA produced by **1t12tgba** and **Duggi** are shown in Figure 5. Table 2 gives cumulative results of different configurations on these formulas. All configurations of **Duggi** use more time than **1t12tgba**, but produce smaller automata on average. One reason why **Duggi** uses more time is the on-demand nature of algorithm, which rebuilds the intermediate t-GBA several times while disambiguating. The average number of disambiguation iterations per formula of **Duggi** on the entire LTLSTORE was 9.5.

Table 2: Cumulative results on the LTLSTORE benchmark set.

	non-WDBA-recognizable				WDBA-recognizable			
	states	∅ states	time in s	timeouts	states	∅ states	time in s	timeouts
Duggi	16,169	20.702	38,932	167	6,866	16.308	5,958	51
Duggi <sub>∖R</sub>	15,450	20.196	37,803	183	6,857	16.287	5,978	51
Duggi <sub>∖RH</sub>	14,415	19.323	39,772	202	6,882	16.346	5,758	51
1t12tgba	19,547	24.618	6,089	154	9,250	20.240	3,965	15
1t12tgba <sub>WDBA</sub>	19,411	24.539	7,309	157	7,632	16.700	3,814	15

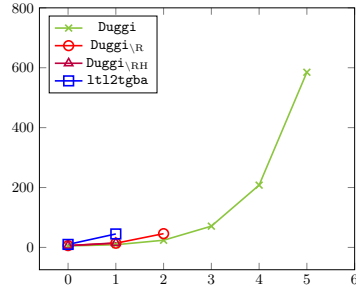
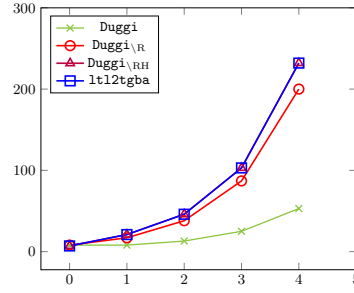
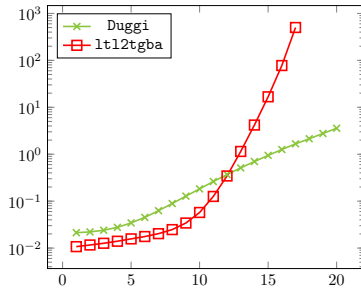
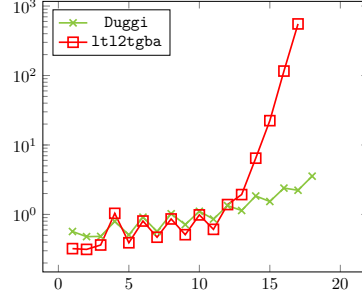
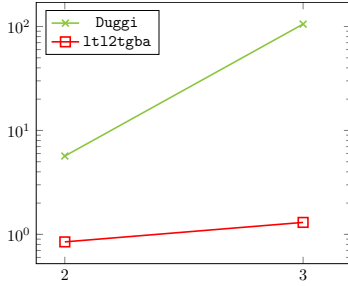
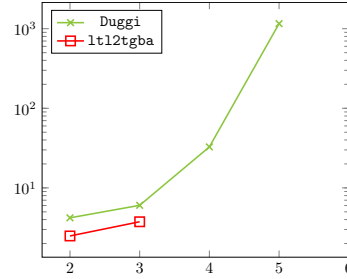
(a)  $\Phi_n = \bigwedge_{i \leq n} (\Diamond \Box p_{2i} \vee \Box \Diamond p_{2i+1})$ (b)  $\theta_n = (\bigwedge_{i \leq n} \Box \Diamond p_i) \rightarrow \Box (req \rightarrow \Diamond res)$ 

Fig. 6: UBA sizes for two sets of parametrized formulas.

**LTL rewrites and the purely-universal heuristic.** A formula that benefits from using the rewrite rules I and II is  $\Phi_n = \bigwedge_{i \leq n} \Diamond \Box p_{2i} \vee \Box \Diamond p_{2i+1}$ , which describes a *strong fairness* condition. Here **1t12tgba** applies the rule  $\Diamond \varphi \vee \Box \Diamond \psi \mapsto \Diamond (\varphi \vee \Box \Diamond \psi)$  which yields  $\bigwedge_{i \leq n} \Diamond (\Box p_{2i} \vee \Box \Diamond p_{2i+1})$ . Applying rule II yields the formula  $\Psi_n = \Diamond \Box (\bigwedge_{i \leq n} p_{2i} \vee \Diamond p_{2i+1})$ . Figure 6a shows that **Duggi** produces smaller automata for  $\Phi_n$ . Figure 6b shows the corresponding results for the parametrized formula  $\theta_n = (\bigwedge_{i \leq n} \Box \Diamond p_i) \rightarrow \Box (req \rightarrow \Diamond res)$  which is a request/response pattern under fairness conditions.

A property that profits from the “on-demand” disambiguation is: “*b* occurs *k* steps before *a*”. We express it with the formula  $\varphi_k^{\text{steps}} = \neg a \mathcal{U} (b \wedge \neg a \wedge \bigcirc \neg a \wedge \dots \wedge \bigcirc^{k-1} \neg a \wedge \bigcirc^k a)$ . Both **Duggi** and **1t12tgba** produce the minimal UBA, but **1t12tgba** produces an exponential-sized automaton in an intermediate step, because it does not realize that the original structure is already unambiguous. This leads to high run times for large *k* (see Figure 7a).

**Use case: probabilistic model checking.** Now we look at an important application of UBA, the analysis of Markov chains. We compare run times of an implementation of [5] for Markov chain model checking with UBA, using **PRISM** (version 4.4) and either **Duggi** or **1t12tgba** as automata generation backends. We take two models of the **PRISM** benchmark suite [25], the bounded retransmission protocol, and the cluster working protocol [21].

(a) Time in seconds needed for the translation of  $\varphi_k^{\text{steps}}$  into a UBA.(b) Time in seconds needed for model checking the BRP model with  $\varphi_k^{\text{steps}}$ .Fig. 7: Time consumption for translating and model checking  $\varphi_k^{\text{steps}}$  (which includes building the automaton).(a) Time consumption for  $\varphi_k$ .(b) Time consumption for  $\psi_k$ .Fig. 8: Model checking times for the cluster protocol with  $\varphi_k$  and  $\psi_k$ .

The bounded retransmission protocol (BRP) is a message transmission protocol, where a sender sends a message and receives an acknowledgment if the transmission was successful. We set the parameter  $N$  (the number of the message parts) to 16, and  $MAX$  (the number of maximal retries) to 128. We reuse  $\varphi_k^{\text{steps}}$ , which now means: “ $k$  steps before an acknowledgment there was a retransmit”, where we replace  $a$  by `ack_received` and  $b$  by `retransmit`. As expected, the faster automaton generation leads to lower model checking times when using Duggi (Figure 7b). The reason for the spikes in Figure 7b is that the probability of the property is zero in the BRP model for odd  $k$ . This makes the model checking (which uses the numeric procedure of [5]) easier. For bigger  $k$  the automaton generation uses a bigger share of the time, making this effect less pronounced.

As second model we analyse the cluster working model with the LTL properties presented in [20]. It consists of a workstation cluster with two sub-clusters that are connected by a backbone and have  $n = 16$  participants each. Let  $\text{fct}_i$  denote the number of functional working stations in sub-cluster  $i$ . We define  $\varphi_{\square\Diamond} = \square\Diamond(\text{fct}_1 = n)$ , which expresses that the first cluster stays functional

on the long run and  $\varphi_{\Diamond\Box} = \bigvee_{i \in \{0, \dots, k\}} \Diamond\Box(\text{fct}_2 = n - i)$ , which expresses the property that from some point, the second cluster contains at least  $n - k$  functional working stations. We check the two formula patterns  $\varphi_k = \varphi_{\Box\Diamond} \wedge \varphi_{\Diamond\Box}$  and  $\psi_k = \varphi_{\Box\Diamond} \vee \varphi_{\Diamond\Box}$ . We leave out a third property described in [20], which is WDBA-recognizable (see Appendix F for further details).

The results for  $\varphi_k$  are depicted in Figure 8a. Both tools have a time-out at  $k = 4$ , although, for smaller  $k$ , the time consumption of **Duggi** was bigger than **ltl2tgba**. Comparing the automata size, **Duggi** produces smaller automata for both  $k = 2$  and  $k = 3$ , e.g., 32 (**Duggi**) vs. 137 (**ltl2tgba**) states for  $k = 3$ . The results for  $\psi_k$  can be seen in Figure 8b. **Duggi** performed better than **ltl2tgba**, as **Duggi** reached the time-out at  $k = 6$  (vs.  $k = 4$  for **ltl2tgba**). However, if no time-out was reached, **ltl2tgba** consumed less time. Nevertheless, for  $k \leq 3$ , model checking time of both tools was below 7 s. Still, **Duggi** produced smaller automata, e.g., 25 (**Duggi**) vs. 59 (**ltl2tgba**) states for  $k = 3$ .

## 7 Conclusion

In this paper we have presented a novel LTL-to-UBA translation. In contrast to other LTL-to-UBA translations [12,6,14] we use alternating automata as an intermediate representation. To adapt the VWAA-to-NBA construction of [17] for the unambiguity setting, we introduced a notion of unambiguity for VWAA and a corresponding disambiguation procedure. This may be of independent interest when considering unambiguity for different types of alternating automata. We devise heuristics that exploit structural properties of purely-universal and alternating formulas for disambiguation. Furthermore, we identify LTL rewriting rules that benefit the construction of UBA.

Experimental analysis on a big LTL benchmark set shows that our tool **Duggi** produces smaller automata on average than the existing tools. In particular, formulas containing nested  $\Diamond$  and  $\Box$  benefit from our heuristics and rewrite rules. Such formulas occur often, for example when modelling fairness properties. Experiments on Markov chain model checking indicate that the positive properties of our approach carry over to this domain.

Our approach opens up many possibilities for optimization, for example by processing multiple source states at once, or in a certain order. This would let us decrease the number of disambiguation steps, and thus the run time. It would be interesting to investigate intermediate strategies in our framework that allow for a trade-off between automata sizes and computation times. Another promising direction is to identify more patterns on LTL or VWAA that allow special disambiguation transformations. As many interesting properties stem from the safety-/cosafety-class, a combination of our approach with the ideas of the UFA generation described in [28] seems to be beneficial. The application of simulation-based automata reductions to UBA is also an open question. Whereas bisimulation preserves unambiguity, it is unclear whether there exist simulation relations targeted specifically at shrinking unambiguous automata.

## References

1. Arnold, A.: Deterministic and non ambiguous rational  $\omega$ -languages. In: Automata on Infinite Words, Ecole de Printemps d'Informatique Théorique, Le Mont Dore, May 1984. Lecture Notes in Computer Science, vol. 192, pp. 18–27. Springer (1985)
2. Babiak, T., Blahoudek, F., Duret-Lutz, A., Klein, J., Křetínský, J., Müller, D., Parker, D., Strejček, J.: The Hanoi omega-automata format. In: Proceedings of the 27th International Conference on Computer Aided Verification (CAV). Lecture Notes in Computer Science, vol. 9206, pp. 479–486. Springer (2015)
3. Babiak, T., Křetínský, M., Řehák, V., Strejček, J.: LTL to Büchi automata translation: Fast and more deterministic. In: Proceedings of the 18th International Conference on Tools and Algorithms for the Construction and Analysis of Systems (TACAS). Lecture Notes in Computer Science, vol. 7214, pp. 95–109. Springer (2012)
4. Baier, C., Katoen, J.P.: Principles of Model Checking. MIT Press (2008)
5. Baier, C., Kiefer, S., Klein, J., Klüppelholz, S., Müller, D., Worrell, J.: Markov chains and unambiguous Büchi automata. In: Proceedings of the 28th International Conference on Computer Aided Verification (CAV) - Part I. Lecture Notes in Computer Science, vol. 9779, pp. 23–42. Springer (2016)
6. Benedikt, M., Lenhardt, R., Worrell, J.: LTL model checking of interval Markov chains. In: Proceedings of the 19th International Conference on Tools and Algorithms for the Construction and Analysis of Systems (TACAS). Lecture Notes in Computer Science, vol. 7795, pp. 32–46. Springer (2013)
7. Bousquet, N., Löding, C.: Equivalence and inclusion problem for strongly unambiguous Büchi automata. In: Proceedings of the 4th International Conference on Language and Automata Theory and Applications (LATA). Lecture Notes in Computer Science, vol. 6031, pp. 118–129. Springer (2010)
8. Carton, O., Michel, M.: Unambiguous Büchi automata. Theor. Comput. Sci. **297**(1-3), 37–81 (2003)
9. Clarke, E.M., Grumberg, O., Peled, D.A.: Model checking. MIT Press (2001)
10. Colcombet, T.: Unambiguity in automata theory. In: Proceedings of the 17th International Workshop on Descriptive Complexity of Formal Systems (DCFS). Lecture Notes in Computer Science, vol. 9118, pp. 3–18. Springer (2015)
11. Couvreur, J.: On-the-fly verification of linear temporal logic. In: Proceedings of the World Congress on Formal Methods in the Development of Computing Systems (FM). Lecture Notes in Computer Science, vol. 1708, pp. 253–271. Springer (1999)
12. Couvreur, J., Saheb, N., Sutre, G.: An optimal automata approach to LTL model checking of probabilistic systems. In: Proceedings of the 10th International Conference on Logic for Programming Artificial Intelligence and Reasoning (LPAR). Lecture Notes in Computer Science, vol. 2850, pp. 361–375. Springer (2003)
13. Duret-Lutz, A.: Manipulating LTL formulas using Spot 1.0. In: Proceedings of the 11th International Symposium on Automated Technology for Verification and Analysis (ATVA). Lecture Notes in Computer Science, vol. 8172, pp. 442–445. Springer (2013)
14. Duret-Lutz, A.: Contributions to LTL and  $\omega$ -Automata for Model Checking. Habilitation thesis, Université Pierre et Marie Curie (Paris 6) (Feb 2017)
15. Duret-Lutz, A., Lewkowicz, A., Fauchille, A., Michaud, T., Renault, E., Xu, L.: Spot 2.0 — a framework for LTL and  $\omega$ -automata manipulation. In: Proceedings of the 14th International Symposium on Automated Technology for Verification and Analysis (ATVA). Lecture Notes in Computer Science, vol. 9938, pp. 122–129. Springer (Oct 2016)



16. Etessami, K., Holzmann, G.: Optimizing Büchi automata. In: Proceedings of the 11th International Conference on Concurrency Theory (CONCUR). Lecture Notes in Computer Science, vol. 1877, pp. 153–167. Springer (2000)
17. Gastin, P., Oddoux, D.: Fast LTL to Büchi automata translation. In: Proceedings of the 13th International Conference on Computer Aided Verification (CAV). Lecture Notes in Computer Science, vol. 2102, pp. 53–65. Springer (2001)
18. Gerth, R., Peled, D., Vardi, M.Y., Wolper, P.: Simple on-the-fly automatic verification of linear temporal logic. In: Proceedings of the 15th IFIP WG6.1 International Symposium on Protocol Specification (PSTV). IFIP Conference Proceedings, vol. 38, pp. 3–18. Chapman & Hall (1995)
19. Grädel, E., Thomas, W., Wilke, T. (eds.): Automata, Logics, and Infinite Games: A Guide to Current Research, Lecture Notes in Computer Science, vol. 2500. Springer (2002)
20. Hahn, E.M., Li, G., Schewe, S., Turrini, A., Zhang, L.: Lazy Probabilistic Model Checking without Determinisation. In: 26th International Conference on Concurrency Theory (CONCUR 2015). Leibniz International Proceedings in Informatics (LIPIcs), vol. 42, pp. 354–367. Schloss Dagstuhl–Leibniz-Zentrum fuer Informatik, Dagstuhl, Germany (2015)
21. Haverkort, B.R., Hermanns, H., Katoen, J.P.: On the use of model checking techniques for dependability evaluation. In: 19th IEEE Symposium on Reliable Distributed Systems (SRDS). pp. 228–237. IEEE Computer Society (2000)
22. Isaak, D., Löding, C.: Efficient inclusion testing for simple classes of unambiguous  $\omega$ -automata. Information Processing Letters **112**(14–15), 578–582 (2012)
23. Karmarkar, H., Joglekar, M., Chakraborty, S.: Improved upper and lower bounds for Büchi disambiguation. In: Proceedings of the 11th International Symposium on Automated Technology for Verification and Analysis (ATVA). pp. 40–54 (2013)
24. Kretínský, J., Meggendorfer, T., Sickert, S.: LTL store: Repository of LTL formulae from literature and case studies. CoRR **abs/1807.03296** (2018), <http://arxiv.org/abs/1807.03296>
25. Kwiatkowska, M.Z., Norman, G., Parker, D.: The PRISM benchmark suite. In: Proceedings of the 9th International Conference on Quantitative Evaluation of Systems (QEST). pp. 203–204. IEEE Computer Society (2012)
26. Löding, C.: Efficient minimization of deterministic weak  $\omega$ -automata. Information Processing Letters **79**(3), 105–109 (2001)
27. Löding, C., Thomas, W.: Alternating automata and logics over infinite words. In: Theoretical Computer Science: Exploring New Frontiers of Theoretical Informatics IFIP TCS. pp. 521–535 (2000)
28. Mohri, M.: On the disambiguation of finite automata and functional transducers. International Journal of Foundations of Computer Science **24**(6), 847–862 (2013)
29. Müller, D., Sickert, S.: LTL to deterministic Emerson-Lei automata. In: Proceedings of the 8th International Symposium on Games, Automata, Logics and Formal Verification (GandALF). Electronic Proceedings in Theoretical Computer Science, vol. 256, pp. 180–194. Open Publishing Association (2017)
30. Muller, D.E., Saoudi, A., Schupp, P.E.: Weak alternating automata give a simple explanation of why most temporal and dynamic logics are decidable in exponential time. In: Proceedings of the Third Annual Symposium on Logic in Computer Science (LICS). pp. 422–427 (1988)
31. Muller, D.E., Schupp, P.E.: Alternating automata on infinite trees. Theoretical Computer Science **54**, 267–276 (1987)
32. Sistla, A.P., Clarke, E.M.: The complexity of propositional linear temporal logics. Journal of the ACM **32**(3), 733–749 (1985)

- 33. Somenzi, F., Bloem, R.: Efficient Büchi automata from LTL formulae. In: Proceedings of the 12th International Conference on Computer Aided Verification (CAV). Lecture Notes in Computer Science, vol. 1855, pp. 248–263. Springer (2000)
- 34. Stearns, R.E., Hunt, H.B.: On the equivalence and containment problem for unambiguous regular expressions, grammars, and automata. *SIAM Journal on Computing* pp. 598–611 (1985)
- 35. Vardi, M.Y.: Automatic verification of probabilistic concurrent finite-state programs. In: Proceedings of the 26th IEEE Symposium on Foundations of Computer Science (FOCS). pp. 327–338. IEEE Computer Society (1985)
- 36. Vardi, M.Y.: Nontraditional applications of automata theory. In: Proceedings of the International Conference on Theoretical Aspects of Computer Software (TACS). pp. 575–597 (1994)
- 37. Vardi, M.Y., Wolper, P.: An automata-theoretic approach to automatic program verification (preliminary report). In: Proceedings of the 1st Symposium on Logic in Computer Science (LICS). pp. 332–344. IEEE Computer Society Press (1986)

## A From LTL to VWAA

We follow the translation given in [17] but use all subformulas as states, as proposed in [3]. The main idea for the construction is to take advantage of the expansion laws for  $\mathcal{U}$  and  $\mathcal{R}$ :

$$\begin{aligned}\varphi\mathcal{U}\psi &= \psi \vee (\varphi \wedge \bigcirc(\varphi\mathcal{U}\psi)) \\ \varphi\mathcal{R}\psi &= \psi \wedge (\varphi \vee \bigcirc(\varphi\mathcal{R}\psi))\end{aligned}$$

For a better presentation, we give here the transition function in an unusual way, namely as a function:  $\delta : Q \rightarrow 2^{2^\Sigma \times 2^Q}$ . It is easy to transform a function  $\delta$  given in the above form into the representation  $\Delta : Q \times \Sigma \rightarrow \mathcal{B}^+$ :

$$\Delta(q, a) = \bigvee_{\substack{(\alpha, C) \in \delta(q) \\ a \in \alpha}} \bigwedge_{s \in C} s$$

The construction takes an LTL formula in which all negations appear in front of atomic propositions (*positive normal form*) as input. If all dual operators are included in the syntax ( $\mathcal{R}$  is the dual of  $\mathcal{U}$ ), LTL formulas can be transformed into an equivalent formula in positive normal form of the same length.

**Definition 15.** Let  $\varphi$  be an LTL formula in positive normal form over  $AP$ . We define the co-Büchi  $\omega$ -automaton  $\mathcal{A}_\varphi$  as a tuple  $(Q, \Sigma, \delta, \varphi, \text{Fin}(Q_f))$  where  $\Sigma = 2^{AP}$ ,  $Q$  is the set of subformulas of  $\varphi$ ,  $Q_f = \{\psi_1\mathcal{U}\psi_2 : \psi_1\mathcal{U}\psi_2 \in Q\}$  and  $\delta$  is defined as follows:

$$\begin{aligned}\delta(\text{true}) &= \{(\Sigma, \emptyset)\} & \delta(\text{false}) &= \emptyset \\ \delta(a) &= \{(\Sigma_a, \emptyset)\} & \delta(\neg a) &= \{(\Sigma_{\neg a}, \emptyset)\} \\ \delta(\varphi \wedge \psi) &= \delta(\varphi) \otimes \delta(\psi) & \delta(\varphi \vee \psi) &= \delta(\varphi) \cup \delta(\psi) \\ \delta(\bigcirc\varphi) &= \{(\Sigma, \{\varphi\})\} \\ \delta(\varphi\mathcal{U}\psi) &= \delta(\psi) \cup (\delta(\varphi) \otimes \{(\Sigma, \{\varphi\mathcal{U}\psi\})\}) \\ \delta(\varphi\mathcal{R}\psi) &= \delta(\psi) \otimes (\delta(\varphi) \cup \{(\Sigma, \{\varphi\mathcal{R}\psi\})\})\end{aligned}$$

where

$$\begin{aligned}M_1 \otimes M_2 &= \{(u \cap v, Q_1 \cup Q_2) : (u, Q_1) \in M_1, (v, Q_2) \in M_2\} \\ \Sigma_a &= \{m \in \Sigma : a \in m\} \\ \Sigma_{\neg a} &= \Sigma \setminus \Sigma_a\end{aligned}$$

The translation has been refined in [3] to produce smaller automata with less non-determinism. In our implementation, we use these optimizations, but these optimizations do not alter or simplify our disambiguation algorithm.

**Lemma 16 ([17,3]).** Let  $\varphi$  and  $\mathcal{A}_\varphi$  be as above. Then,  $\mathcal{L}(\varphi) = \mathcal{L}(\mathcal{A}_\varphi)$ .

## B Degeneralization

As some applications demand an automaton with a (non-generalized) Büchi acceptance on states, a degeneralization procedure, which also converts transition-based acceptance to state-based acceptance, is necessary. We use the degeneralization construction as presented in [17].

**Definition 17 (Degeneralization).** Let  $\mathcal{G} = (Q, \Sigma, \delta, Q_0, \text{Inf}(F_1) \wedge \dots \wedge \text{Inf}(F_n))$  be a t-GBA. Then we define  $\mathcal{N}_{\mathcal{G}}$  to be  $(Q \times \{0, \dots, n\}, \Sigma, \delta_{\mathcal{N}}, Q_0 \times \{0\}, Q \times \{n\})$  where

$$\begin{aligned} \delta_{\mathcal{N}}(\langle q, i \rangle, a) &= \{ \langle q', i' \rangle : q' \in \delta(q, a) \text{ and } i' = \text{next}(i, q \xrightarrow{a} q') \} \text{ and} \\ \text{next}(i, q \xrightarrow{a} q') &= \begin{cases} \max(\{i \leq j \leq n : \forall k \in \{i+1, \dots, j\}. q \xrightarrow{a} q' \in F_k\}) & \text{if } i \neq n \\ \max(\{0 \leq j \leq n : \forall k \in \{0, \dots, j\}. q \xrightarrow{a} q' \in F_k\}) & \text{if } i = n \end{cases} \end{aligned}$$

This construction creates copies of  $\mathcal{G}$  for every Büchi acceptance set  $\text{Inf}(F_i)$  and switches from copy  $i$  to copy  $i'$  if and only if the current transition satisfies all acceptance sets  $\text{Inf}(F_j)$  with  $i < j \leq i'$ . Therefore:

**Lemma 18 (Theorem 3 of [17]).** Let  $\mathcal{G}$  and  $\mathcal{N}_{\mathcal{G}}$  be as above. Then  $\mathcal{L}(\mathcal{G}) = \mathcal{L}(\mathcal{N}_{\mathcal{G}})$ .

Obviously, the number of states in  $\mathcal{B}_{\mathcal{G}}$  is  $|Q| \cdot n$ .

## C Proof for Section 3

The correctness proof of the original construction from VWAA to t-GBA in [17] features a similar proof, but here we want to show that we can find corresponding runs such that the layers of the VWAA are equal to the states of the t-GBA. To show this, we need the precondition that the VWAA is unambiguous.

The first part shows that every accepting run of  $\mathcal{G}_{\mathcal{A}}$  is matched by an accepting run of  $\mathcal{A}$ . It is complicated by the way that accepting transitions are defined in  $\mathcal{G}_{\mathcal{A}}$ : they require that a final state  $q_f$  must *have the option* to choose a non-looping transition infinitely often (see Definition 1). Thus, there might exist runs of  $\mathcal{G}_{\mathcal{A}}$  that represent runs of  $\mathcal{A}$  where  $q_f$  has this option infinitely often, but never takes the “good” transition. However, this situation cannot occur if  $\mathcal{A}$  is unambiguous as it would imply multiple accepting runs for the same word that differ on some layer.

In the following proofs we denote the successors of a node  $q \in V$  in a graph  $(V, E)$  by  $E(q)$ . As before, we use  $\delta(q, a)$  to denote the set of minimal models of  $\Delta(q, a)$ , and  $I$  to denote the set of minimal models of  $\iota$ .

**Lemma 5.** If  $\mathcal{A}$  is unambiguous, then for every accepting run  $r = Q_0 Q_1 \dots$  of  $\mathcal{G}_{\mathcal{A}}$  for  $w \in \Sigma^\omega$  there exists an accepting run  $\rho = (V, E)$  of  $\mathcal{A}$  for  $w$  such that  $Q_i = V(i)$  for all  $i \geq 0$ .

*Proof.* We show this lemma by induction over the number of elements in  $F$ .

Base case:  $F = \emptyset$ . We define  $V = \{(q, i) : q \in Q_i\}$ .  $Q_0 \in I$  must hold as  $r$  is a run of  $\mathcal{G}_A$  and thus the initial condition is satisfied. For every transition  $Q_i \xrightarrow{w[i]} Q_{i+1}$  we know that  $Q_{i+1} \in \bigotimes_{q \in Q_i} \delta(q, w[i])$  and we define the edges between  $V(i)$  and  $V(i+1)$  to match the corresponding successor sets. The result is a run of  $\mathcal{A}$  for  $w$  and it is accepting as  $F = \emptyset$ .

Now consider  $F = F' \cup \{q_f\}$ . Let  $\mathcal{A} = (Q, \Sigma, \Delta, \iota, \text{Fin}(F))$  and  $\mathcal{A}' = (Q, \Sigma, \Delta, \iota, \text{Fin}(F'))$  be  $\mathcal{A}$  where  $q_f$  is not marked final. Let  $w$  and  $r = Q_0 Q_1 \dots$  be an accepting run of  $\mathcal{G}_A$  for  $w$ . By Definition 1 the only difference between  $\mathcal{G}_{A'}$  and  $\mathcal{G}_A$  is that  $\mathcal{G}_A$  has the additional acceptance set  $T_{q_f}$ , which adds an obligation for runs to be accepting. This implies that any accepting run of  $\mathcal{G}_A$  is also an accepting run of  $\mathcal{G}_{A'}$ . So  $r$  is an accepting run of  $\mathcal{G}_{A'}$  and we make use of the induction hypothesis to get an accepting run  $\rho = (V, E)$  of  $\mathcal{A}'$  such that for all  $i$ :  $Q_i = V(i)$ . If we can show that  $\rho$  is also an accepting run of  $\mathcal{A}$  for  $w$  we are done.

So suppose that it is not accepting. Then there exists a rejecting path  $\pi$  through  $\rho$  that ultimately stabilizes on a state  $f \in F$ . But  $f$  can only be  $q_f$ , as any other rejecting path would contradict the fact that  $\rho$  is an accepting run of  $\mathcal{A}'$ . Hence there is some  $k$  such that for all  $j > k$ :  $q_f \in V(j)$  and  $((q_f, j), (q_f, j+1)) \in E$ .

As  $q_f \in F$  we know that there are infinitely many  $i$  such that  $(Q_i, w[i], Q_{i+1}) \in T_{q_f}$ .  $\mathcal{G}_A$  has only a finite amount of states, and  $\Sigma$  is also finite, so there must exist a triple  $(S, a, S')$  such that for infinitely many  $i > k$ :  $Q_i = S$ ,  $Q_{i+1} = S'$ ,  $w[i] = a$  and  $(S, a, S') \in T_{q_f}$ . Furthermore, for infinitely many of these  $i$  the edges of  $\rho$  between  $S$  and  $S'$  will be the same. We fix an edge relation between  $S$  and  $S'$  in  $\rho$  that occurs infinitely often, name the successors of  $q \in S$  by  $e(q)$  and set:

$$J = \{j \mid S = V(j), S' = V(j+1), w[j] = a \text{ and } E(q, j) = e(q) \text{ for all } q \in S\}$$

It follows that  $S' = \bigcup_{q \in S} e(q)$  and  $J$  is infinite. By the definition of  $T_{q_f}$  there exists a way for  $q_f$  to take a non-looping transition, i.e. there exists a successor configuration  $Y$ , such that  $Y \in \delta(q_f, a)$ ,  $q_f \notin Y$  and  $Y \subseteq S'$ .

Our aim is to show that we can choose  $Y$  instead of  $e(q_f)$  as successor set of  $q_f$  in  $S$  for all these edges, without losing the property that for all  $i$ :  $Q_i = V(i)$ . Clearly, we would not make the following layer bigger by choosing  $Y$ , as  $Y \subseteq S'$ . So our aim is to show that by choosing  $Y$  as successor set for  $q_f$ , the following layer is not strictly smaller. We show that the following holds:

$$S' = Y \cup \underbrace{\left( \bigcup_{q \in S \setminus \{q_f\}} e(q) \right)}_{S''}$$

This would mean that replacing  $e(q_f)$  by  $Y$  does not change the following layer in the run.

Suppose that this equality does not hold. Then  $S''$  is strictly contained in  $S'$ , as  $Y \subseteq S'$ . Both  $S''$  and  $S'$  are possible successor configurations of  $S$  and  $a$  as  $e(q) \in \delta(q, a)$  for every  $q \in S$  and both  $Y$  and  $e(q_f)$  are elements of  $\delta(q_f, a)$ .

But then for all  $j \in J$  we can construct an accepting run  $\rho_j$  of  $\mathcal{A}$  on  $w$  such that all pairs of runs in  $\{\rho_j \mid j \in J\}$  differ on some layer. We construct  $\rho_j = (V', E')$  by mimicking  $\rho$  up till position  $j$  and in position  $j$  we choose  $Y$  as successor configuration of  $q_f$ . For all following positions  $k > j$  such that  $k \in J$  we also choose  $Y$  as successor configuration of  $q_f$ , given that  $q_f \in V'(k)$ . All infinite paths of  $\rho_j$  that do not visit  $q_f$  infinitely often can be mapped to infinite paths in  $\rho$ , and thus in particular no  $f \in F'$  is visited infinitely often. Furthermore,  $q_f \notin E'(q_f, k)$  for all  $k \in J$  such that  $k \geq j$ , and thus no infinite path stabilizes on  $q_f$ . Therefore  $\rho_j$  is an accepting run of  $\mathcal{A}$  on  $w$ .

Two runs  $\rho_j, \rho_h$  with  $j, h \in J$  and  $j < h$  differ on depth  $j + 1$  as  $\rho_j$  chooses  $Y$  as successor set of  $q_f$  on level  $j$  and  $\rho_h$  chooses  $e(q_f)$ . This implies that the  $(j + 1)$ 'th layer of  $\rho_j$  is  $S''$  and the  $(j + 1)$ 'th is  $S'$  and hence  $\rho_j$  and  $\rho_h$  differ on layer  $j + 1$ , which contradicts the fact that  $\mathcal{A}$  is unambiguous.

Thus  $S' = S''$  and we get  $\rho' = (V', E')$  by mimicking  $\rho$  for all positions  $k \notin J$ , and setting  $E(q_f, j) = Y$  for all positions  $j \in J$ . Edges of the other states are not changed for any position. As  $S' = S''$  we can do this without changing the vertex set, i.e.  $V' = V$ . The new run is accepting as  $q_f \notin E'(q_f, j)$  for all positions  $j \in J$  and hence no infinite branch stabilizes at the state  $q_f$ . Furthermore, the property  $Q_i = V'(i)$  holds for all  $i \in \mathbb{N}$  as it already holds for  $\rho$  by induction hypothesis.

To find an accepting run of  $\mathcal{G}_{\mathcal{A}}$  that corresponds to an accepting run of  $\mathcal{A}$  we can directly use the definition of the transition relation of  $\mathcal{G}_{\mathcal{A}}$ :

**Lemma 6.** *For every accepting run  $\rho = (V, E)$  of  $\mathcal{A}$  for  $w \in \Sigma^\omega$  there exists an accepting run  $r = Q_0 Q_1 \dots$  of  $\mathcal{G}_{\mathcal{A}}$  for  $w$ , such that  $Q_i = V(i)$  for all  $i \geq 0$ .*

*Proof.* Let  $\mathcal{A}$  and  $\mathcal{G}_{\mathcal{A}}$  be as above and  $\rho = (V, E)$  be an accepting run of  $\mathcal{A}$  for  $w$ . We show that  $r = V(0)V(1)\dots$  is an accepting run of  $\mathcal{G}_{\mathcal{A}}$ . As  $\rho$  is a run of  $\mathcal{A}$  we get  $V(0) \in I$  and hence  $r$  satisfies the initial condition of a t-GBA run.

We show that for all  $i : V(i + 1) \in \bigotimes_{q \in V(i)} \delta(q, w[i])$ . As  $\rho$  is a run of  $\mathcal{A}$ , every  $q \in V(i)$  must have successors  $E(q, i) \subseteq V(i + 1)$  such that  $E(q, i) \in \delta(q, w[i])$ . Furthermore, each  $q' \in V(i + 1)$  must have a predecessor in  $V(i)$  which implies that  $\bigcup_{q \in V(i)} E(q, i) = V(i + 1)$ . By definition of  $\otimes$  we get

$$V(i + 1) \in \bigotimes_{q \in V(i)} \delta(q, w[i])$$

and thus  $(V(i), w[i], V(i + 1))$  is a transition of  $\mathcal{G}_{\mathcal{A}}$ .

Finally, we have to show that  $r$  is accepting. We show for all  $f \in F$  that for infinitely many positions  $i : (V(i), w[i], V(i + 1)) \in T_f$ . Suppose that this does not hold. Then there exists a  $k$  such that for all  $n \geq k : (V(n), w[n], V(n + 1)) \notin T_f$ . By the definition of  $T_f$  we know that for all  $n > k : f \in V(n)$  and there is no  $S \in \delta(f, w[n])$  such that  $f \notin S$  and  $S \subseteq V(n + 1)$ . This implies, however, that infinitely often  $f$  is in the set of successors of  $f$ . But then  $\rho$  is not accepting which contradicts the assumption.

**Lemma 7.**  $\mathcal{A}$  is unambiguous if and only if  $\mathcal{G}_{\mathcal{A}}$  is unambiguous.

*Proof.* Suppose that  $\mathcal{A}$  is unambiguous and  $\mathcal{G}_{\mathcal{A}}$  is not. Then there exists a word  $w \in \Sigma^\omega$  and two accepting runs  $r_1 = Q_0Q_1 \dots, r_2 = P_0P_1 \dots$  of  $\mathcal{G}_{\mathcal{A}}$  for  $w$  such that for some  $i \in \mathbb{N}$ :  $Q_i \neq P_i$ . Choose  $u \in \Sigma^*$  and  $v \in \Sigma^\omega$  such that  $w = uv$  and  $|u| = i$ . This contradicts the assumption that  $\mathcal{A}$  is unambiguous as  $V_1(i)$  and  $V_2(i)$  are both reachable via  $u$  and accept the word  $v$ .

To show the other direction, assume that  $\mathcal{A}$  is ambiguous. Then, there exists a word  $w \in \Sigma^\omega$  and two accepting runs  $\rho_1 = (V_1, E_1), \rho_2 = (V_2, E_2)$  of  $\mathcal{A}$  for  $w$  such that for some  $i \in \mathbb{N}$ :  $V_1(i) \neq V_2(i)$ . By Lemma 6, there exist two accepting runs  $r_1, r_2$  of  $\mathcal{G}_{\mathcal{A}}$  on  $w$  such that  $r_1(i) \neq r_2(i)$ , and hence  $\mathcal{G}_{\mathcal{A}}$  is ambiguous.

## D Unambiguity check on VWAA

We now turn to the complexity of deciding whether a VWAA is unambiguous. First, we define the union of two automata which we need to prove hardness.

**Definition 19.** Let  $\mathcal{A}_1 = (Q_1, \Sigma, \Delta_1, \iota_1, \text{Fin}(F_1))$  and  $\mathcal{A}_2 = (Q_2, \Sigma, \Delta_2, \iota_2, \text{Fin}(F_2))$  be two alternating co-Büchi automata over the same alphabet. The union automaton  $\mathcal{A}_1 \cup \mathcal{A}_2$  is defined as a tuple  $(Q_\cup, \Sigma, \Delta_\cup, \iota_\cup, \text{Fin}(F_\cup))$  where

- $Q_\cup = (Q_1 \times \{1\}) \cup (Q_2 \times \{2\})$ ,
- $\Delta_\cup((q, i), a) = \begin{cases} \Delta_1(q, a) \times 1 & \text{if } i = 1 \\ \Delta_2(q, a) \times 2 & \text{if } i = 2 \end{cases}$
- $\iota_\cup = (\iota_1 \times 1) \vee (\iota_2 \times 2)$ , and
- $F_\cup = (F_1 \times \{1\}) \cup (F_2 \times \{2\})$

and  $f \times i$ , for  $f \in \mathcal{B}^+(X)$  and  $i \in \mathbb{N}$ , is the positive Boolean formula over  $X \times \{i\}$  in which every variable  $x \in X$  that occurs in  $f$  is replaced by  $(x, i)$ .

As the definition of the union automaton merely creates an automaton consisting of two disjoint copies, it holds that  $\mathcal{L}(\mathcal{A}_1 \cup \mathcal{A}_2) = \mathcal{L}(\mathcal{A}_1) \cup \mathcal{L}(\mathcal{A}_2)$ .

**Lemma 20.** The problem of deciding whether a given VWAA  $\mathcal{A}$  is unambiguous is PSPACE-hard.

*Proof.* We reduce the PSPACE-hard problem of LTL satisfiability [32] to deciding unambiguity of VWAA.

Assume that we are given an LTL formula  $\varphi$ . We can construct a VWAA  $\mathcal{A}_\varphi$  that is equivalent to  $\varphi$  in time linear in the size of  $\varphi$  [36,30]. An accepting run in  $\mathcal{A}_\varphi$  corresponds to two distinct accepting runs in  $\mathcal{A}_\varphi \cup \mathcal{A}_\varphi$  and thus we have:

$$\mathcal{L}(\mathcal{A}_\varphi \cup \mathcal{A}_\varphi) = \emptyset \text{ iff } \mathcal{A}_\varphi \cup \mathcal{A}_\varphi \text{ is unambiguous.}$$

As  $\mathcal{L}(\mathcal{A}_\varphi \cup \mathcal{A}_\varphi) = \mathcal{L}(\mathcal{A}_\varphi) = \mathcal{L}(\varphi)$  holds,  $\mathcal{A}_\varphi \cup \mathcal{A}_\varphi$  is unambiguous if and only if  $\varphi$  is satisfiable.

**Lemma 21.** *The problem whether a given VWAA  $\mathcal{A}$  is unambiguous can be decided in PSPACE.*

*Proof.* By the theorem of Savitch, we know that NPSPACE=PSPACE. We give an NPSPACE (in the size of VWAA) algorithm for checking whether  $\mathcal{G}_{\mathcal{A}}$  is unambiguous, which is enough by Lemma 7.

The algorithm guesses a lasso in the self product of  $\mathcal{G}_{\mathcal{A}}$ : a path that reaches a state  $(C_1, C_2)$  such that  $C_1 \neq C_2$  and such that there exists an accepting loop starting in  $(C_1, C_2)$ . To guess a successor of a state  $(Q_1, Q_2)$  in  $\mathcal{G}_{\mathcal{A}} \otimes \mathcal{G}_{\mathcal{A}}$  for a symbol  $a \in \Sigma$  the algorithm first guesses a successor configuration of  $Q_1$  and  $a$  in  $\mathcal{A}$ . This can be done by guessing a set  $S_q \subseteq Q$  for each  $q \in Q_1$  and verifying that  $S_q$  is a minimal model of  $\Delta(q, a)$ . Then  $\bigcup_{q \in Q_1} S_q$  is chosen as successor configuration of  $Q_1$  and the procedure is repeated for  $Q_2$ . So non-deterministically guessing a successor of  $(Q_1, Q_2)$  in  $\mathcal{G}_{\mathcal{A}} \otimes \mathcal{G}_{\mathcal{A}}$  can be done in polynomial time.

The algorithm only needs to remember the current state in  $\mathcal{G}_{\mathcal{A}} \otimes \mathcal{G}_{\mathcal{A}}$  and, for the loop, which of the acceptance sets in  $\mathcal{T}$  have already been satisfied. Therefore, it requires at most space polynomial in the size of  $\varphi$ .

By combining Lemma 20 and Lemma 21 we get:

**Theorem 8.** *Deciding whether a VWAA is unambiguous is PSPACE-complete.*

## E Proofs for Section 5

**Lemma 11.** *Every purely-universal formula  $\nu$  can be rewritten into a formula  $\nu_1 \vee \dots \vee \nu_n$ , where  $\nu_i$  is disjunction-free for all  $1 \leq i \leq n$ .*

*Proof.* Let  $\nu$  be a purely-universal formula. As a first step we remove all occurrences of  $\mathcal{U}$  and  $\mathcal{R}$  in  $\nu$  that are not in the scope of some  $\Box$  by applying the rules  $\nu_1 \mathcal{U} \nu_2 \mapsto \nu_2 \vee (\nu_1 \wedge \Diamond \nu_2)$  and  $\varphi \mathcal{R} \nu' \mapsto \nu'$ . These transformation rules preserve equivalence if  $\nu_1, \nu_2$  and  $\nu'$  are purely-universal. Then, we proceed by induction on the structure of  $\nu$ . In the case that  $\nu = \Box \varphi$  it already has the desired structure. For the cases  $\vee, \wedge, \bigcirc$  and  $\Diamond$  we apply the induction hypothesis to the subformulas and then lift the disjunction over the corresponding operators.

**Lemma 13.** *Let  $\nu$  be a disjunction-free and purely-universal formula. Then  $\mathbf{g}(\nu) \wedge \bigcirc \nu \equiv \nu$ .*

*Proof.* Let  $\nu$  be a disjunction-free and purely-universal formula. We show the claim by induction on the structure of  $\nu$  ( $\nu_1, \nu_2$  and  $\nu'$  are assumed to be purely-universal).

- $\nu = \Box \varphi$ : By instantiation we get the  $\varphi \wedge \bigcirc \Box \varphi \equiv \Box \varphi$ .
- $\nu = \bigcirc \nu'$ : By induction hypothesis we have  $\mathbf{g}(\nu') \wedge \bigcirc \nu' \equiv \nu'$ , which implies  $\bigcirc \mathbf{g}(\nu') \wedge \bigcirc \bigcirc \nu' \equiv \bigcirc \nu'$ .
- $\nu = \nu_1 \wedge \nu_2$ : By induction hypothesis we have  $\mathbf{g}(\nu_1) \wedge \bigcirc \nu_1 \equiv \nu_1$  and  $\mathbf{g}(\nu_2) \wedge \bigcirc \nu_2 \equiv \nu_2$ . This implies  $\mathbf{g}(\nu_1) \wedge \mathbf{g}(\nu_2) \wedge \bigcirc \nu_1 \wedge \bigcirc \nu_2 \equiv \nu_1 \wedge \nu_2$ .



- $\nu = \Diamond \nu'$ : We have defined  $\mathbf{g}(\Diamond \nu') = \text{true}$ . As  $\Diamond \nu'$  is an alternating formula it satisfies  $\bigcirc \Diamond \nu' \equiv \Diamond \nu'$ , which proves the claim.

**Lemma 14.** *Let  $\nu$  be a disjunction-free and purely-universal formula. Then*

$$1. \varphi \mathcal{U}(\nu \vee \psi) \equiv \nu \vee \gamma \text{ and } 2. \mathcal{L}(\nu) \cap \mathcal{L}(\gamma) = \emptyset$$

where  $\gamma = \varphi \mathcal{U}((\varphi \wedge \neg \mathbf{g}(\nu) \wedge \bigcirc \nu) \vee (\psi \wedge \neg \nu))$ .

*Proof.* We show 1. by showing  $\varphi \mathcal{U}(\nu \vee \psi) \implies \nu \vee \gamma$  and  $\nu \vee \gamma \implies \varphi \mathcal{U}(\nu \vee \psi)$ .

- $\varphi \mathcal{U}(\nu \vee \psi) \implies \nu \vee \gamma$ : Take a word  $w$  that satisfies  $\varphi \mathcal{U}(\nu \vee \psi)$ . If  $w \models \nu$  we are done. We know that either  $w \models \varphi \mathcal{U} \nu$  or  $w \models \varphi \mathcal{U} \psi$ .  
In the first case there is a *least*  $i$  such that  $w[i..] \models \nu$  and for all  $j < i$ :  $w[j..] \models \varphi$ . We get  $w[(i-1)..] \models \varphi \wedge \neg \mathbf{g}(\nu) \wedge \bigcirc \nu$  by Lemma 13 and thus  $w \models \gamma$ .  
In the second case there is an  $i$  such that  $w[i..] \models \psi$  and for all  $j < i$ :  $w[j..] \models \varphi$ . We can assume that  $w[i..] \not\models \nu$ , as the first case applies otherwise, which implies  $w[i..] \models \psi \wedge \neg \nu$  and thus proves  $w \models \gamma$ .
- $\nu \vee \gamma \implies \varphi \mathcal{U}(\nu \vee \psi)$ : We have:

$$\gamma \implies \varphi \mathcal{U}((\varphi \wedge \bigcirc \nu) \vee \psi) \implies \varphi \mathcal{U}(\nu \vee \psi)$$

which proves this case.

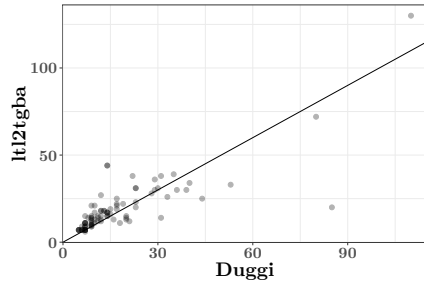
To show 2. we show that  $\mathcal{L}(\nu) \cap \mathcal{L}(\gamma) = \emptyset$ . By Lemma 13 we have  $\neg \nu \equiv \Diamond \neg \mathbf{g}(\nu)$  or  $\neg \nu \equiv \text{true}$ . If  $\neg \nu \equiv \text{true}$  we get  $\mathcal{L}(\nu) = \emptyset$ , which proves the claim. Otherwise, we have  $\neg \nu \equiv \Diamond \neg \mathbf{g}(\nu)$ . Both  $\varphi \wedge \neg \mathbf{g}(\nu) \wedge \bigcirc \nu$  and  $\psi \wedge \neg \nu$  imply  $\Diamond \neg \mathbf{g}(\nu)$ . As  $\Diamond \neg \mathbf{g}(\nu)$  is prefix invariant we get  $\gamma \implies \neg \nu$  and thus  $\mathcal{L}(\nu) \cap \mathcal{L}(\gamma) = \emptyset$ .

## F Implementation and Experiments

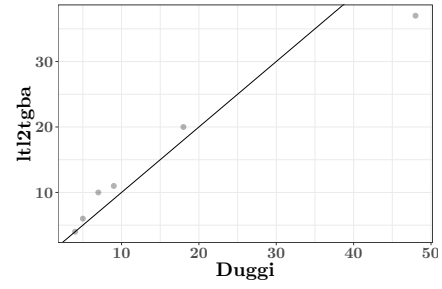
### F.1 LTL to UBA

In this section we give some more details on the experiments with the 1542 benchmark formulas of the LTLSTORE [24]. Those 1542 formulas include every formula in [24] and its negation, as well as duplicates occurring over several files. If one removes the duplicated formulas, the benchmark formula set amounts to 1419 formulas (including negations). The odd number comes from the fact, that  $p_0$  and  $\neg p_0$  are counted as duplicates, as they differ only on the structure of the literal. We use here the version of June 29, 2018, commit AD8B5CD7C9C30D1E65DBDA676FDF41821C3A8ADB.

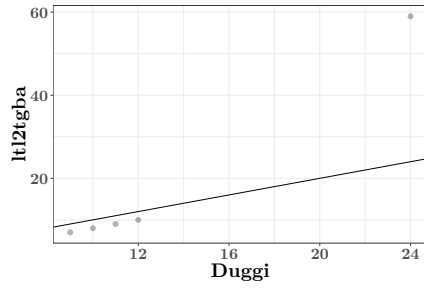
The LTLSTORE is grouped in sets of formulas, that contain either a parametrized family of formulas or a set of formulas used for a case study or tool evaluation. The following figures show scatter plots for the individual formula sets, whose names are given in the caption. We removed the plots where **Duggi** and **ltl2tgba** showed similar behavior concerning automata sizes and time-outs to not enlarge the appendix unnecessarily. You find a description how to generate the missing plots at <https://www.tcs.inf.tu-dresden.de/ALGI/TR/FM19-UBA/>.



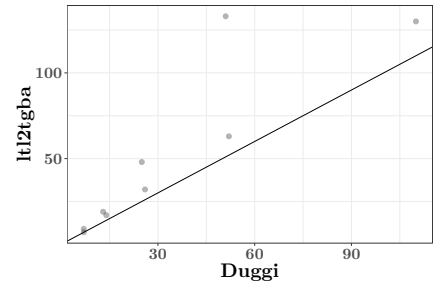
(a) acacia, timeouts: Duggi: 26, lt12tgba: 32



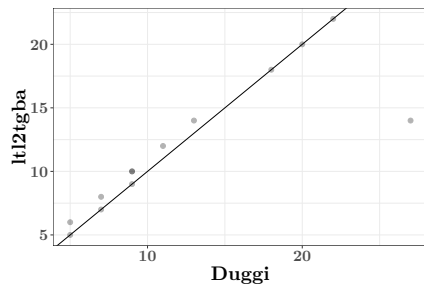
(b) chained, timeouts: Duggi: 2, lt12tgba: 6



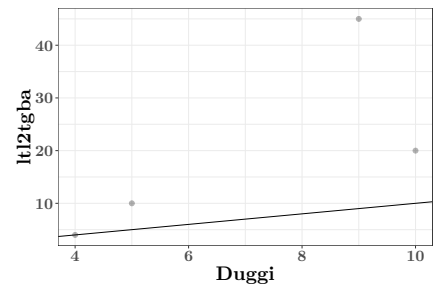
(a) cluster, timeouts: Duggi: 1, lt12tgba: 3



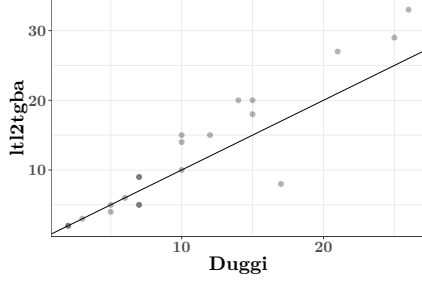
(b) detector\_1-5,10,20, timeouts: Duggi: 4, lt12tgba: 5



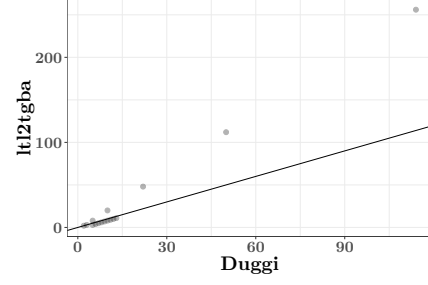
(a) family\_f, timeouts: Duggi: 5, lt12tgba: 3



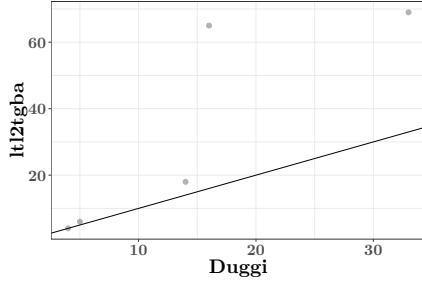
(b) fgff, timeouts: Duggi: 1, lt12tgba: 6



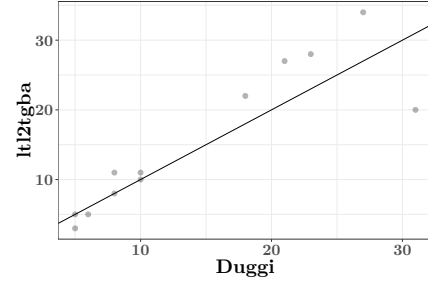
(a) further, timeouts: Duggi: 4, lt12tgba: 7



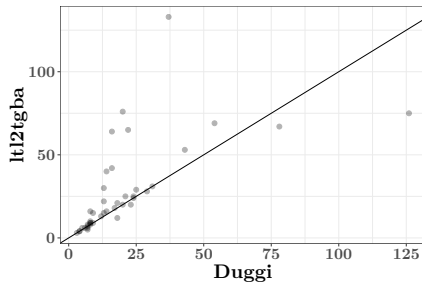
(b) gf\_and, timeouts: Duggi: 4, lt12tgba: 4



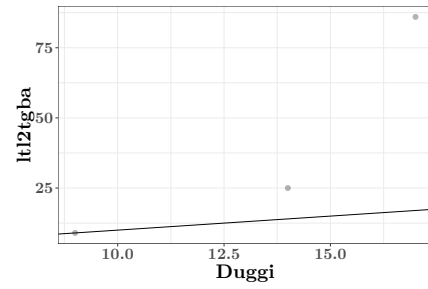
(a) gr1, timeouts: Duggi: 7, lt12tgba: 9



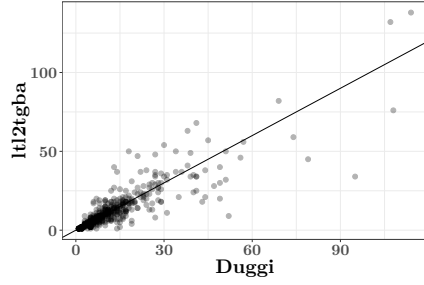
(b) libe\_router, timeouts: Duggi: 2, lt12tgba: 0



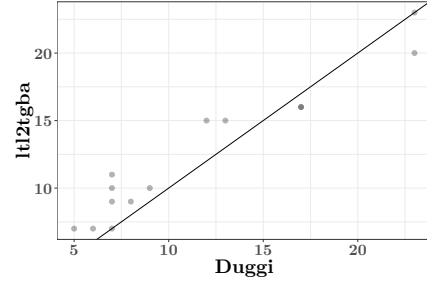
(a) lily2, timeouts: Duggi: 2, lt12tgba: 4



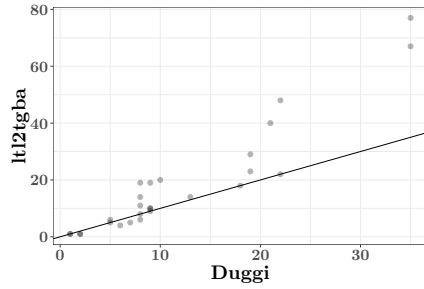
(b) load\_balancer\_1-5, timeouts: Duggi: 6, lt12tgba: 5



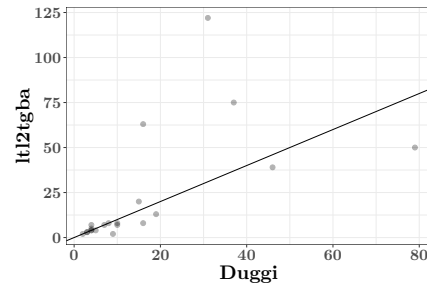
(a) lpar19\_all, timeouts: Duggi: 54, ltl2tgba: 2



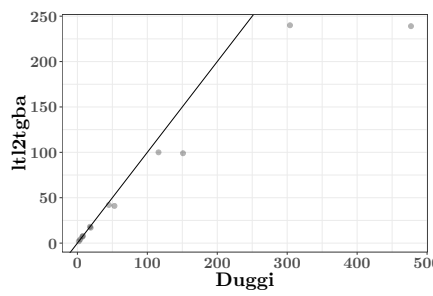
(b) LtINfBa, timeouts: Duggi: 0, ltl2tgba: 0



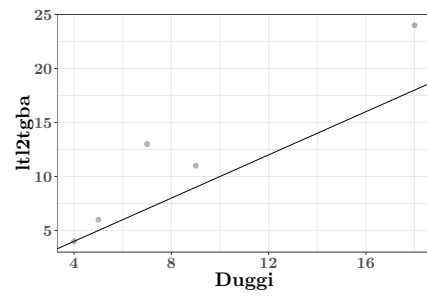
(a) mutual, timeouts: Duggi: 0, ltl2tgba: 1



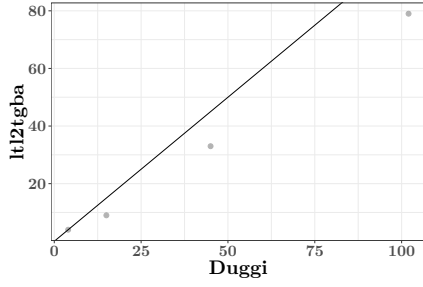
(b) rabinizer3, timeouts: Duggi: 0, ltl2tgba: 3



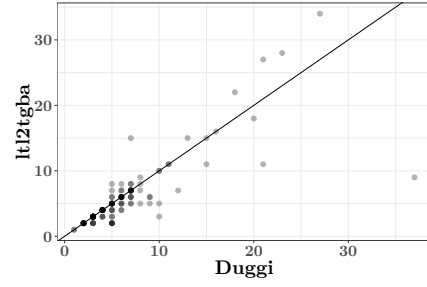
(a) Q, timeouts: Duggi: 8, ltl2tgba: 6



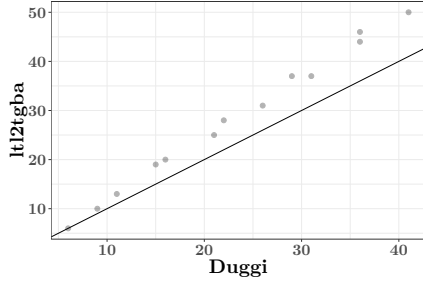
(b) R, timeouts: Duggi: 10, ltl2tgba: 15



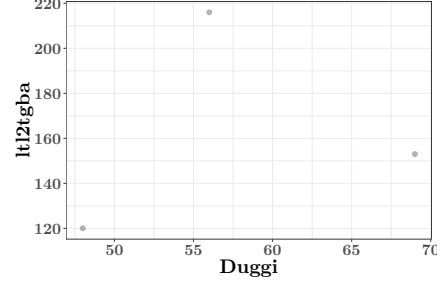
(a) round\_robin\_arbiter\_1-5, timeouts: Duggi: 6, ltl2tgba: 4



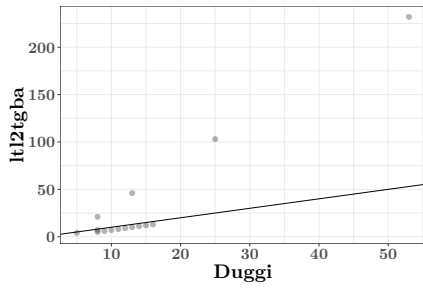
(b) real, timeouts: Duggi: 0, ltl2tgba: 0



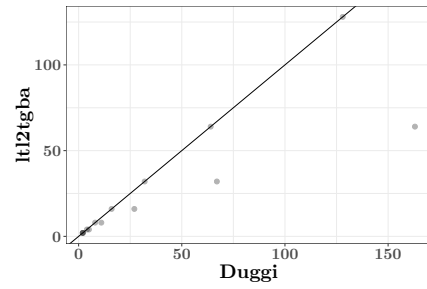
(a) scc, timeouts: Duggi: 2, ltl2tgba: 3



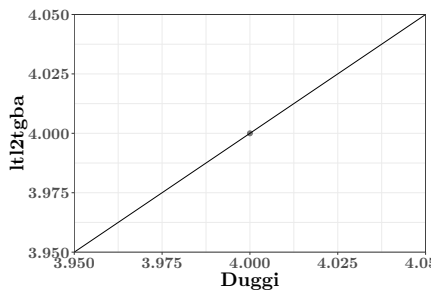
(b) suspension, timeouts: Duggi: 5, ltl2tgba: 4



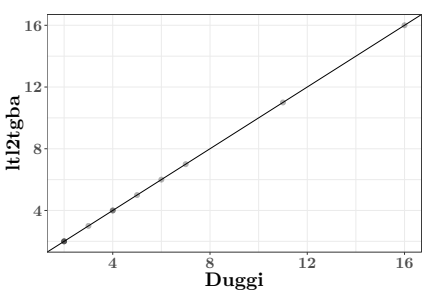
(a) theta, timeouts: Duggi: 5, ltl2tgba: 5



(b) U\_left, timeouts: Duggi: 5, ltl2tgba: 2



(a) lift, timeouts: Duggi: 30, lt12tgba: 15



(b) U\_right, timeouts: Duggi: 8, lt12tgba: 0

## F.2 Markov chain analysis

Here, we analyse shortly the missing formula for the cluster workstation protocol described in Section 6. The formula is

$$\varphi_k^{\mathcal{U}} = \text{left} = n \mathcal{U} \left( \text{left} = n - 1 \mathcal{U} \left( \dots \mathcal{U} (\text{left} = n - k \mathcal{U} \text{right} \neq n) \right) \right)$$

and describes: “The first  $k$  failures occur on the first/left cluster.”

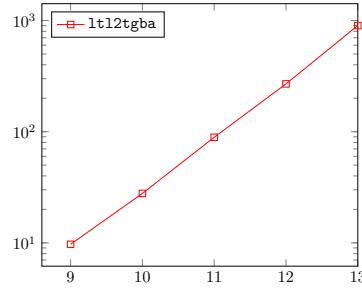


Fig. 22: Model checking times for the cluster workstation protocol for the formula  $\varphi_k^{\mathcal{U}} = \text{left} = n \mathcal{U} \left( \text{left} = n - 1 \mathcal{U} \left( \dots \mathcal{U} (\text{left} = n - k \mathcal{U} \text{right} \neq n) \right) \right)$

For  $\varphi_k^{\mathcal{U}}$ , the results (depicted in Figure 22) shows that in this case **1t12tgba** performed significantly better, as **Duggi** was not able to finish the model checking procedure for any  $k$ . In particular the automata generation took too long. However, the language described by  $\varphi_k^{\mathcal{U}}$  is WDBA recognizable, which for this particular formula, can be recognized already syntactically. Also, the WDBA generated by **1t12tgba** is smaller than the UBA generated (without WDBA minimization), e.g., for  $k = 9$  the (complete) WDBA has 11 states, whereas the (complete) UBA has 57 states.