CONVERTING A SUBSET OF LTL FORMULA TO BÜCHI AUTOMATA

Bilal Kanso and Ali Kansou

Department of Computer Science, Lebanese University, Beirut

ABSTRACT

The translation of LTL formula into equivalent Büchi automata plays an important role in many algorithms for LTL model checking, which consist in obtaining a Büchi automaton that is equivalent to the software system specification and another one that is equivalent to the negation of the property. The intersection of the two Büchi automata is empty if the model satisfies the property.

Generating the Büchi automaton corresponding to an LTL formula may, in the worst case, be exponential in the size of the formula, making the model checking effort exponential in the size of the original formula. There is no polynomial solution for checking emptiness of the intersection. That comes from the translation step of LTL formula into finite state models. This makes verification methods hard or even impossible to be implemented in practice. In this paper, we propose a subset of LTL formula which can be converted to Büchi automata whose the size is polynomial.

KEYWORDS

Linear Temporal Logic, Büchi automata, Model checking, Compositional modelling.

1. INTRODUCTION

Model checking becomes increasingly one of the most important tools to verify the correctness of computer-based control systems [1, 4, 12, 15]. It is a formal verification technique consisting in algorithmically verifying whether system properties such as the absence of deadlocks (described in some appropriate logical formalism such as temporal logic) are satisfied by the system (described as a suitable finite state model). The success of the model checking technique comes from the fact that it is completely automatic. Running a model checking on a given system model to verify a desired property leads automatically to fail state or successful state. In case the system model fails to satisfy the property, the model checking tool can offer a counterexample which can be used as an error trace provided for debugging purposes.

Model checking approaches vary according to the logic used to specify system properties [3, 12, 18, 21]. One of the most used logics is the Linear Temporal Logic (LTL) [11, 20]. The underlying idea consists in transforming the negation of the LTL expression into a Büchi automaton, and then computing the product between the Büchi automaton representing the system and the one representing the negation of the LTL expression. If the product is not empty, that means the property expressed by the negation of the LTL expression is not satisfied by the system, otherwise the property is well-satisfied. However, the decision problem for emptiness of the intersection is PSPACE-hard [2, 19]. That comes from the translation of LTL formula into Büchi automata. Indeed, the space complexity of this approach is linear in the size of Büchi automata and exponential in the length of the LTL formula: the Büchi automaton of a property (described as a LTL formula) is constructed in exponential space in the length of this property [22].
makes verification methods hard or even impossible to be implemented in practice and makes the scalability of the LTL model checking limited, which commonly referred to as the state explosion problem [8].

In this paper, we contribute to finding a subset of LTL properties that can be converted polynomially into Büchi automata. Finding such a subset of LTL logic will be viewed as one the most promising directions to bridge the gap between the increasing complexity of state models and actual model checking methods. We define a fragment that we call, Flat LTL Logic and we show how formula in this fragment can be transformed into Büchi automata whose state space size is linear. Due to the structure of flat LTL formula, our algorithm can be compositional in the sense that the final finite state model associated to a given formula is obtained by developing a sub-automaton for each sub-formula of the principal formula. Hence, the basic idea for developing the final automaton for a flat LTL formula $\varphi$ is that $\varphi$ can be recursively decomposed into a set of sub-formula, arriving at sub-formula that can be completely handled. Composition is then used for assembling different sub-automata and then forming larger ones. Such a composition can be seen as an operation taking sub-automata for sub-formula as well as the flat LTL operator to provide a new more complex automaton.

In order to guide the construction of the final automaton for a flat LTL formula $\varphi$ from the sub-automata associated to the sub-formula $\varphi_1$, $\varphi_2$, ..., $\varphi_n$ of $\varphi$, we build the finite syntax tree, $FST(\varphi)$ of the formula $\varphi$. The nodes of a finite syntax tree are labelled, either by flat LTL operators or by propositional operators. The leaves are labelled only by atomic propositions. Thus, the target Büchi automaton is obtained by exploring the tree in pre-order.

The rest of this article is organized as follows: Section 2 briefly describes Büchi automata. In Section 3, we describe our fragment of LTL logic and the reasons to choose it. In Section 4, we present for each formula in our fragment LTL, its equivalent Büchi automata and show the proof of this equivalence. Section 5 presents the finite syntax tree associated to a formula defined in our fragment LTL while Section 6 shows the final algorithm that generates to any formula in our fragment an equivalent Büchi automaton. Section 7 presents the conclusion and some future works.

2. AUTOMATA ON INFINITE WORDS

2.1. Büchi automata

Automata on infinite inputs were introduced by Büchi. A Büchi automaton is a non-deterministic finite-state automaton which takes infinite words as input [9, 10, 14]. A word is accepted if the automaton goes through some designated “good” states infinitely often while reading it. A Büchi automaton is a finite state automaton defined by a 5-tuple $A = (S, s_0, F, \Sigma, \delta)$ where:

- $S$ is a finite set of states,
- $s_0 \in S$ is the initial state,
- $\Sigma$ is a non-empty set of atomic propositions,
- $F \subseteq S$ is a finite set of accepting states,
- $\Delta : S \times \Sigma \rightarrow 2^S$ is a transition function.

In the following of this paper, the initial state of a Büchi automaton is pointed to by incoming arrows and the accepting states are marked by double circles.
A run of $A$ on $\sigma = \sigma(0)\sigma(1)\sigma(2) \cdots \in \Sigma^\omega$ is an infinite sequence of states $s_0s_1s_2 \cdots \in S^\omega$ starting with the initial state $s_0$ of $A$ such that $\forall i, i \geq 0, \ s_{i+1} \in \delta(s_i, \sigma(i))$. A run $s_0s_1s_2 \cdots$ is accepting by an automaton $A$ if $A$ goes through accepting states (i.e. $\in F$) infinitely often while reading it. The accepted language of a Büchi automaton $A$, denoted by $\mathcal{L}_\omega(A)$, is then defined by $\mathcal{L}_\omega(A) = \{ \sigma \in \Sigma^\omega \mid \text{there is an accepting run for } \sigma \text{ in } A \}$.

2.2. Operations on Büchi automata

The basic idea of the construction of the union of two Büchi automata $A_1$ and $A_2$ is to add a new initial (nonaccept) state $s_{\text{new}}$ to the set of states union of $A_1$ and $A_2$. The transitions of the union of $A_1$ and $A_2$ are transitions of both $A_1$ and $A_2$ with the following transitions:

a) A transition from $s_{\text{new}}$ to a state $s$ labelled with a proposition $p$ if and only if there is transition from the initial state of $A_1$ to the state $s$ labelled with the proposition $p$;

b) A transition from $s_{\text{new}}$ to a state $s$ labelled with a proposition $p$ if and only if there is transition from the initial state of $A_2$ to the state $s$ labelled with the proposition $p$.

Definition 1 (Büchi automata union). Let $A_1 = (S_1, s_{10}, F_1, \Sigma, \delta_1)$ and $A_2 = (S_2, s_{20}, F_2, \Sigma, \delta_2)$ be two büchi automata. The union $A_1 \cup A_2$ of $A_1$ and $A_2$ is the büchi automaton $A = (S, s_0, F, \Sigma, \delta)$ defined as follows:

- $S = S_1 \cup S_2 \cup \{ s_0 \}$
- $s_0 \in S$ is the initial state
- $F = F_1 \cup F_2$
- the transition relation $\delta$ is defined as follows:

$$\delta(s, p) = \begin{cases} 
\delta_1(s, p) & \text{if } s \in S_1 \\
\delta_2(s, p) & \text{if } s \in S_2 \\
\delta_1(s_{10}, p) \cup \delta_2(s_{20}, p) & \text{if } s \text{ is the initial state } s_0
\end{cases}$$

The construction of the intersection automaton works a little differently from the finite state automata case. One needs to check whether both sets of accepting states are visited infinitely often. Consider two runs $r_1$ and $r_2$ and a word $\sigma$ where $r_1$ goes through an accept state after $\sigma(0), \sigma(2), \cdots$ and $r_2$ enters accept state after $\sigma(0) \sigma(3) \cdots$. Thus, there is no guarantee that $r_1$ and $r_2$ will enter accept states simultaneously. To overcome this problem, we need to identify the accept states of the intersection of the two automata. To do so, we create two copies of the intersected state space. In the first copy, we check for occurrence of the first acceptance set. In the second copy, we check for occurrence of the second acceptance set. When a run enters a final state in the first copy, we wait for that run also enters in an accept state in the second copy. When this is encountered, we switch back to the first copy and so on. We repeat jumping back and forth between the two copies whenever we find an accepting state.
In this section, we introduce our subset of LTL logic that we call Flat LTL Logic. This fragment will be used to express temporal properties and then translate them into Büchi automata in linear size. The syntax of our Flat LTL logic adds to usual boolean propositional operators \( \neg \) (negation) and \( \wedge \) (conjunction), some modal operators that describe how the behaviour changes with time.

- **Next**: \( X \varphi \) requires that the formula \( \varphi \) be true in the next state;
- **Until**: \( \varphi_1 U \varphi_2 \) requires that the formula \( \varphi_1 \) be true *until* the formula \( \varphi_2 \) is true, which is required to happen;
- **Eventually**: \( \Diamond \varphi \) requires that the formula \( \varphi \) be true at some point in the future (starting from the present) and it is equivalent to \( \Diamond \varphi \equiv \text{true} U \varphi \);
- **Release**: \( \varphi_1 R \varphi_2 \) requires that its second argument \( \varphi_2 \) always be true, a requirement that is released as soon as its first argument \( \varphi_1 \) becomes true. It is equivalent to \( \varphi_1 R \varphi_2 \equiv \neg (\neg \varphi_1 U \neg \varphi_2) \).

### 3.1. Our fragment LTL logic

**Definition 3 (Flat LTL formulae).** The set of Flat LTL formulae \( \mathcal{L}_f \) is given by the following grammar:

\[
\varphi ::= \Theta | \Theta U \varphi | \varphi R \Theta | X \varphi | \neg \Delta | \varphi_1 \wedge \varphi_2 | \varphi_1 \vee \varphi_2
\]

where \( \Theta \) is a propositional formula defined by the following grammar:

\[
\Theta ::= \text{true} | p | \neg \Theta | \Theta_1 \wedge \Theta_2
\]

and \( \Delta \) is the temporal formula defined by the following grammar:

\[
\Delta ::= \Delta U \Theta | \Theta R \Delta | X \varphi | \neg \Delta \quad \text{with } p \in \Sigma
\]
Example: the formula \( X(a \, U \, \neg(d \, R \, \neg(b \, U \, X \, c))) \) is not in \( \mathcal{L}_f \) since the sub-formula \( \neg(b \, U \, X \, c) \) should be of the form \( \Delta \, U \, \theta \) that is not the case. But, the formula \( X(a \, U \, \neg(d \, R \, \neg(b \, R \, X \, c))) \) is in \( \mathcal{L}_f \).

For the sake of brevity and the lack of space, we only discuss here why the fragment \( \theta \, U \, \phi \) is included within our LTL fragment to the detriment of both formula \( \phi_1 \, U \, \phi_2 \) and \( \phi_1 \, U \, \theta \). It is well-known the size of an Büchi automaton \( A \) that recognizes the complement language \( \mathcal{L}_\omega(A) \) of the language accepted \( \mathcal{L}_\omega(A) \) by an automaton \( A \) is exponential [13, 16]. Suppose we have separately built an automaton \( A_1 \) for \( \phi_1 \) and an automaton \( A_2 \) for \( \phi_2 \), and let us then try to compositionally obtain the resulting automaton \( A \) for \( \phi \). According to the until operator’s semantics, it is required that \( \phi \) holds at the current moment, if there is some future moment for which \( \phi_2 \) holds and \( \phi_1 \) holds at all moments until that future moment. That means constructing the automaton for \( \phi = \phi_1 \, U \, \phi_2 \) firstly requires constructing of the intersection of \( A_1 \) and \( A_2 \). As stated previously, computing \( A_2 \) is exponential and therefore, constructing the Büchi automaton associated to it, does not need to complement any Büchi automaton.

3.2 Flat Positive Normal Form (FPNF)

As LTL formula, flat LTL formula can be transformed into the so-called Flat Positive Normal form (FPNF). This form is characterized by the fact that negations only occur adjacent to atomic propositions. All negation symbols of the given LTL formula have to be pushed inwards over the temporal operators.

**Definition 4 (FPNF).** The set of Flat Positive Normal Form (FPNF) formulae \( \mathcal{L}_{FPNF} \) is given by the following grammar:

\[
\varphi := \text{true} \mid \text{false} \mid p \mid \neg p \mid \varphi_1 \, \land \, \varphi_2 \mid \varphi_1 \, \lor \, \varphi_2 \mid \theta \, U \, \varphi \mid \varphi \, R \, \theta \mid X\varphi
\]

Each formula \( \varphi \in \mathcal{L}_f \) can be transformed into a formula \( \varphi' \in \mathcal{L}_{FPNF} \). This is done by pushing negations inside, near to atomic propositions. To do this, we use the following transformation rules:

\[
\begin{align*}
\neg \text{true} & \rightarrow \text{false} \\
\neg \neg \varphi & \rightarrow \varphi \\
\neg X\varphi & \rightarrow X\neg\varphi \\
\neg (\varphi \, U \, \theta) & \rightarrow \neg\varphi \, \neg \theta \\
\neg (\varphi \, R \, \theta) & \rightarrow \varphi \, \neg \theta \, U \, \neg \varphi
\end{align*}
\]

This transformation is done in linear complexity as it is shown by the following theorem:

**Theorem 2.** For any flat LTL formula \( \varphi \in \mathcal{L}_f \), there exists an equivalent LTL formula \( \varphi' \in \mathcal{L}_{FPNF} \) in flat positive normal form with \( |\varphi'| = O(|\varphi|) \).

Example: the formula \( X(a \, U \, \neg(d \, R \, \neg(b \, R \, Xc))) \) is in \( \mathcal{L}_f \), but not in \( \mathcal{L}_{FPNF} \). It can be transformed into \( X(a \, U \, \neg(d \, U \, \neg(b \, U \, X \, \neg c))) \) which is in \( \mathcal{L}_{FPNF} \).
3.3 Semantics

The semantics of LTL formula is defined over infinite sequences $\sigma : \mathbb{N} \rightarrow 2\Sigma$. In other words, a model is an infinite sequence $A_0 A_1 \cdots$ of subsets of $\Sigma$. The function $\sigma$, called interpretation function, describes how the truth of atomic propositions changes as time progresses. For every sequence $\sigma$, we write $\sigma = (\sigma(0), \cdots, \sigma(n), \cdots)$. Thus, we have the following two notations:

- $\sigma(i)$ denotes the state at index $i$ and $\sigma(i:j)$ the part of $\sigma$ containing the sequence of states between $i$ and $j$;
- $\sigma(i\ldots) = A_i A_{i+1} A_{i+2} \cdots$ denotes the suffix of a sequence $\sigma = A_0 A_1 A_2 \cdots \in (2\Sigma)^{\omega}$ starting in the $(i+1)$st symbol $A_i$.

We also write $\sigma(i) \models \varphi$ to denote that "$\varphi$ is true at time instant $i$ in the model $\sigma$". This notion is defined inductively, according to the structure of $\varphi$.

The LTL formula are interpreted over infinite sequences of states $\sigma : \mathbb{N} \rightarrow 2\Sigma$ as follows:

1. $2\Sigma$ is the power set of the proposition set $\Sigma$.
2. $\omega$: is typically used to denote infinity.

4. Construction of Büchi Automata for Flat LTL Logic

Our algorithm is a compositional algorithm. It constructs for each sub-formula in our fragment LTL logic, an equivalent Büchi automaton and in a compositional way regroup all resulting Büchi automata in order to get the target Büchi automaton of the target flat LTL formula.

In the sequel, we firstly explain for each sub-formula in our fragment LTL logic how its equivalent Büchi automaton can be obtained.

4.1. Büchi automata for $\theta$ formula

The Büchi automaton associated to a propositional formula $\theta$ is obtained by creating two states $s_0$ and $s_1$ and two transitions $tr_1$ and $tr_2$. $s_0$ is the only initial state while $s_1$ is the only final state. $tr_1$ is the transition from $s_0$ to $s_1$ labelling with $\theta$ while the transition $tr_2$ is a loop labelled with $true$ over the state $s_2$.

**Definition 6 (\(\theta\) automaton).** Let $\theta$ be a propositional formula. The automaton $A_\theta = (S_\theta, s_0^\theta, F_\theta, \Sigma, \delta_\theta)$ associated to $\theta$ is defined as follows:

- $S_\theta = \{s_0, s_1\}$, $s_0^\theta = s_0$, $F_\theta = \{s_1\}$.
- The transition function $\delta$ is defined as follows:
  $\delta_\theta(s_0, \theta) = \{s_1\}$ and $\delta_\theta(s_1, true) = \{s_1\}$

**Figure 1** shows the Büchi automaton associated to the propositional formula $\theta = a \land \neg b$.

![Figure 1: Example of automaton associated to $\theta$](image-url)


4.2. Büchi automata for $\theta U \phi$ formula

The main idea behind the composition $\theta U \phi$ is to add a new initial (nonaccept) state $s_{\text{new}}$ to the set of states of the automaton $A_\phi$ associated to $\phi$ with the following transitions:

a) A loop over the added state $s_{\text{new}}$ labelled with the propositional formula $\theta$

b) Transitions $s_{\text{new}}$ to a state $s$ labelled with a proposition $p$ if and only if there a transition from the initial state $s^0$ of $A_\phi$ to the state $s$ labelled with the proposition $p$

All other transitions of $A_\phi$, as well as the accept states, remain unchanged. The state $s_{\text{new}}$ is the single initial state of the resulting automaton, is not accept, and, clearly, has no incoming transitions except the loop one.

**Definition 7 (\(\Theta U \phi\) automaton).** Let $\Theta$ be a propositional formula and $\phi$ be an LTL flat formula. Let $A_\phi = (S_\phi, s^0_\phi, F_\phi, \Sigma, \delta_\phi)$ be the automaton associated to $\phi$. The automaton $A_\psi = (S_\psi, s^0_\psi, F_\psi, \Sigma, \delta_\psi)$ associated to $\psi = \Theta U \phi$ is defined as follows:

- $S_\psi = \{s_{\text{new}}\} \cup S_\phi$
- $s^0_\psi = s_{\text{new}}$, $F_\psi = F_\phi$
- The transition function $\delta_\psi$ is defined as follows:

$$
\delta_\psi(s, p) = \begin{cases}
\delta_\phi(s, p) & \text{if } s \in S_\phi \text{ (} A_\phi \text{ transitions)} \\
\delta_\phi(s^0_\phi, p) & \text{if } s = s_{\text{new}} \text{ (Connection initial state to } A_\phi) \\
\{s_{\text{new}}\} & \text{if } s = s_{\text{new}} \text{ and } p = \Theta \text{ (Loop over the new initial state)}
\end{cases}
$$

**Example:** Figure 2 illustrates the composition definition of $\theta U \phi$. Figure 2a shows the Büchi automaton associated to $(\Diamond b) R c$. To construct the Büchi automaton associated to $(a U ((\Diamond b) R c))$, we add a new state $s_{\text{new}}$ that we consider as initial state. Then, for each transition outgoing from $s_{\text{new}}$ with label $l$ and goes to state $s$, we add a transition from $s_{\text{new}}$ to the state $s$ with a label $l$. Finally, we then add a loop labelled with the atomic proposition $a$ over the added state.

![Diagram](image)

**Figure 2:** Example of composition: $\theta U \phi$

**Theorem 3.** Let $\psi = \Theta U \phi$ be a flat LTL formula and $A_\phi$ be the büchi automaton equivalent to $\phi$. Let $A_\psi$ be the automaton built according to Definition 7. Then, $\text{Words}(\psi) = \mathcal{L}_\omega(A_\psi)$ (See Proof in Appendix).
Eventually operator $\Diamond \varphi$: the Büchi automaton construction of the formula $\Diamond \varphi$ is a particular case of the Büchi automaton construction of the formula $\theta U \varphi$ where $\theta$ is the true formula. Thus, the main idea behind the composition $\Diamond \varphi$ is to add a new initial (nonaccept) state $s_{\text{new}}$ to the automaton states set $A_\varphi$ associated to $\varphi$ with the same transitions defined for $\theta U \varphi$ where the loop over the added state $s_{\text{new}}$ is labelled with $true$ instead of the atomic formula $\theta$.

4.3. Büchi automata for $X \varphi$ formula

The main idea behind the composition $X \varphi$ consists in adding two new states $s_{\text{new}}$ (neither initial state or accept state) and $s_{\text{init}}$ (considered as the initial state) to the state set of the automaton $A_\varphi$ with the following transitions:

a) Add for any transition in $A_\varphi$ which starts from the initial state $s^0$ to a state $s$, a transition from $s_{\text{new}}$ to $s$;
b) Add a transition from the initial state $s_{\text{init}}$ to the $s_{\text{new}}$ labelled with $true$.

All other transitions of $A_\varphi$ remain unchanged and final states of $A_\varphi$ become accept ones of $A_\psi$ and initial state of $A_\psi$ become the state $s_{\text{init}}$.

Definition 8 ($X \varphi$ automaton). Let $\varphi$ be an Flat LTL formula. Let $A_\varphi = (S_\varphi, s_0, F_\varphi, \Sigma, \delta_\varphi)$ be the automaton equivalent to $\varphi$. The automaton $A_\psi = (S_\psi, s_0, F_\psi, \Sigma, \delta_\psi)$ equivalent to $\psi = X \varphi$ is defined as follows:

- $S_\psi = S_\varphi \cup \{s_{\text{new}}, s_{\text{init}}\}$
- $s_{\text{init}} = s_{\text{init}}, F_\psi = F_\varphi$
- The transition function $\delta$ is defined as follows:
  \[
  \delta_\psi(s, p) = \begin{cases} 
  \delta_\varphi(s, p) & \text{if } s \in S_\varphi \text{ (} A_\varphi \text{ transitions)} \\
  \{s_{\text{new}}\} & \text{if } s = s_{\text{init}} \text{ and } p = true \text{ (Connection } s_{\text{new}} \text{ state to initial state of } A_\varphi) 
  \end{cases}
  \]

Example: Figure 3 illustrates the definition of $X \varphi$. Figure 3a shows the Büchi automaton associated to the formula $(a U (X b R c))$. To construct the Büchi automaton equivalent to $X(a U (X b R c))$, we add a new state $s_{\text{new}}$ and for each transition $tr$ starting from the initial state $s^0_\varphi$ to a state $s$, a transition from $s_{\text{new}}$ to $s$ with the same label. Finally, we add the state $s_{\text{init}}$ that we consider as initial and we connect $s_{\text{init}}$ to $s_{\text{new}}$ with a transition labelled with the $true$ label.

![Figure 3: Example of composition: $X \varphi$](image-url)
**Theorem 4.** Let ψ = Xφ be a LTL formula and Aφ be the Büchi automaton equivalent to φ. Let Aψ be the LTL automaton built according to Definition 8. Then, \( \text{Words}(X\varphi) = \mathcal{L}_\omega(A_\psi) \) (See Proof in Appendix)

### 4.4. Büchi automata for φ R θ formula

The formula \( \varphi R \theta \) informally means that θ is true until φ becomes true, or θ is true forever. Thus, the construction of a Büchi automaton for \( \varphi R \theta \) can be done by constructing the Büchi automaton associated to the fact that θ is true until φ becomes true and the construction of a Büchi automaton associated to the fact that θ is true forever. Finally, make the union between the two constructed Büchi automata. Consequently, to build the Büchi automaton for \( \varphi R \theta \), we need to add two new states \( s_i \) and \( s_f \) to the set of states of the automaton \( A_\phi \). \( s_i \) becomes the single initial state of the resulting automaton and \( s_f \) is added to set of final states of the resulting automaton. The following transitions are added to the set of transitions of the resulting automaton:

- a) For any transition from the initial state \( s_0 \) of \( A_\phi \) to a state \( s \) labelled with a proposition \( p \), add a transition from the state \( s_i \) to \( s \) labelled with the proposition \( p \land \theta \);
- b) A loop over the added state \( s_i \) labelled with the propositional formula \( \theta \);
- c) A loop over the added state \( s_f \) labelled with the propositional formula \( \theta \);
- d) A transition from the state \( s_i \) to the state \( s_f \) labelled with the proposition \( \theta \).

All other transitions of \( A_\phi \), as well as the accept states, remain unchanged.

**Definition 9** (φ R θ automaton). Let \( \theta \) be a propositional formula and \( \varphi \) be an LTL flat formula. Let \( A_\varphi = (S_\varphi, s_0^\varphi, F_\varphi, \Sigma, \delta_\varphi) \) be the automaton associated to \( \varphi \). The automaton \( A_\psi = (S_\psi, s_0^\psi, F_\psi, \Sigma, \delta_\psi) \) associated to \( \psi = \varphi R \theta \) is defined as follows:

- \( S_\psi = \{s_i, s_f\} \cup S_\varphi \)
- \( s_0^\psi = s_i, F_\psi = F_\varphi \cup \{s_f\} \)
- The transition function \( \delta \) is defined as follows:

\[
\delta_\psi(s, p) = \begin{cases} 
\delta_\varphi(s_0^\varphi, p) & \text{if } s = s_i \text{ and } p = \theta \land p' \text{ (Connection } s_i \text{ to initial state of } A_\varphi) \\
\{s_i, s_f\} & \text{if } s = s_i \text{ and } p = \theta \text{ (Loop over } s_i \text{ or connection } s_i \text{ to } s_f) \\
\{s_f\} & \text{if } s = s_f \text{ and } p = \theta \text{ (Loop over } s_f) 
\end{cases}
\]

**Example:** Figure 4 illustrates the composition definition of \( \varphi R \theta \). Figure 4a shows the Büchi automaton associated to the formula \( c U \diamond b \). To construct the Büchi automaton associated to the flat LTL formula \( (c U \diamond b) R a \), we add a state \( s \), that we consider as the only initial state and a state \( s_f \) that we consider as a final state. We add a loop labelled with the atomic proposition \( S_a \$ \) over the two added states. Finally, for each transition outgoing from the initial state of the automaton \( \varphi \) with label \( l \) and goes to state \( s \), we add a transition from the added state \( s_i \) to the state \( s \) with a label \( (l \land a) \). We also add a transition labelled with a from the state \( s_i \) to the state \( s_f \).
5. Finite Syntax Tree Of Flat LTL Formula

A flat LTL formula \( \varphi \) can be transformed into a tree containing all the information about the possible sub-formula of \( \varphi \). It will form the cornerstone of the construction of Büchi automata from flat LTL formula. We assume that our flat LTL formula are fully parenthesized and we show how to build the finite syntax tree, named \( \text{FST}(\varphi) \), algorithmically for a flat LTL formula \( \varphi \). This tree can be thought of as a data structure representing the sub-formula after a finite breaking up the formula into a list of tokens. We distinguish four kinds of tokens: left brackets "(" , right brackets ")", FLTL operators and propositional variables. FLTL operators represent the internal nodes of our tree while the propositional variables represent the leaf nodes. Our algorithm to build \( \text{FST}(\varphi) \) is inspired from [5] and uses a stack for operators and a stack for propositional variables, and consists of the following rules:

a) If the current token is a left bracket "(" (i.e. we are reading a new sub-formula), push it on the operator stack;

b) If the current token is a operator (i.e. in \{\'\land\', \'\lor\', \'\mathcal{F}\', \'\mathcal{U}\', \'\mathcal{X}\', \'\mathcal{G}\'\} with \( \mathcal{F} \) and \( \mathcal{G} \) are binary operators), push it on the operator stack;

c) If the current token is a propositional variable \( p \), create a tree with single node whose the value is \( p \) and push the created tree on the variable stack;

d) If the current token is a right bracket ")" (i.e. we have just finished reading a sub-formula), pop operators off the operator stack while this operator is not a left bracket. If the popped operator is an unary operator, pop one tree variable from variable stack and create new tree whose the root is the popped operator and it is only child is the popped tree. If the popped operator is a binary operator, pop two tree variables from variable stack and create new tree whose the root is the popped operator and its right child the first popped tree and its left child the second popped tree. If no left bracket is found during popping the variable stack, throw a mismatched bracket expression. Otherwise, pop found left bracket from the operator stack;

e) At the end of reading expression tokens, pop all operators off the operator stack and for each popped operator:
If the popped operator is an unary operator, pop one tree variable from variable stack and create new tree whose the root is the popped operator and it is only child is the popped tree. Then, push the created tree on the variable stack;

If the popped operator is a binary operator, pop two tree variables from variable stack and create new tree whose the root is the popped operator and its right child the first popped tree and its left child the second popped tree. Then, push the created tree on the variable stack;

If the popped operator is left or right bracket, throw an unbalanced left bracket.

Hence, our mechanism of creating $FST(\varphi)$ can be described by the algorithm illustrated in Figure 5:

```plaintext
Input: a positive flat LTL formula $\varphi$  
Output: the finite syntax tree $FST(\varphi)$;
operatorStack ← createStack();
operandStack ← createTreeStack();
l ← split($\varphi$);
for $e \in l$ do
    if isSpace($e$) then
        continue;
    else if leftBracket($e$) or unary($e$) or binary($e$) then
        push(operatorStack, e);
    else if variable($e$) then
        push(operandStack, createNode(e));
    else if rightBracket($e$) then
        while !emptyStack(operatorStack) do
            popped ← pop (operatorStack);
            if unary (popped) then
                push (operandStack, addRight (popped, pop (stackOperand), pop (stackOperand), e));
            else if binary (popped) then
                push(stackOperand, addRightLeft (popped, pop (stackOperand), pop (stackOperand), e));
            else
                break; // encountered a left bracket
        end
    if emptyStack (operatorStack) then
        throw Exception("Unbalanced right parentheses");
    else
        throw Exception(Unknown token);
end
while !emptyStack(operatorStack) do
    popped ← top (operatorStack);
    pop (operatorStack);
    if unary (popped) then
        push (operandStack, addRight (popped, pop (operandStack)));
    else if binary (popped) then
        push(stackOperand, addRightLeft (popped, pop (stackOperand), pop (stackOperand), e));
    else
        throw Exception("Unbalanced left parentheses");
end
if length (operandStack)=1 then
    return top (operandStack);
else
    throw Exception("Error in LTL expression");
```

Figure 5: Building syntax tree for a FLTL formula
**Example:** Figure 6a shows the finite syntax tree $FST(\varphi)$ generated for the FLTL expression:

$$\varphi = \Diamond a \rightarrow (b \rightarrow ((\neg f) U (d \land (\neg e)))) R c.$$

![Finite Syntax Tree](image)

**Figure 6: Example of finite syntax tree**

This finite syntax tree will be used to construct the Büchi automaton equivalent to a flat LTL formula $\varphi$ in flat positive normal form. Since our algorithm takes as input a flat positive LTL formula, any node in the finite syntax tree labelled with the negation operator $\neg$ is certainly located directly before a leaf. For technical reasons, we merge the nodes labelled with $\neg$ with the leaf which directly follows in the finite syntax tree. Figure 6b illustrates the finite syntax tree presented in Figure 6a after pushing negations to leaves.

**6. From Finite Syntax Tree To Büchi Automata**

Our algorithm to build Büchi automata from flat LTL formula is compositional in the sense that the final Büchi automaton is obtained by developing a sub-automaton for each sub-formula of the principal formula. Hence, the basic idea for developing the final automaton for a flat LTL formula $\varphi$ is to explore $FST(\varphi)$ in a pre-order traversal. That is to say, we visit the root node first, then recursively do a pre-order traversal of the left sub-tree, followed by a recursive pre-order traversal of the right sub-tree. The algorithm, illustrated in Figure 7, allows us to build a Büchi automaton from a finite syntax tree of a positive flat LTL formula $T=FST(\varphi)$ and uses the following five functions:

a) **CreateBuchiProp(\theta):** takes as input a propositional formula $\theta$ and returns the automaton as defined in Definition 6 (Section 4);

b) **CreateBuchiUnary(op, BA):** takes as input an unary LTL formula ($i.e.$ $\text{op} \in \{X, \Diamond\}$) and a Büchi automaton $BA$ and returns a Büchi automaton defined according to definitions of $\Diamond$ and $X$ given in Section 4;

c) **CreateBuchiBinary(op, BA_l, BA_r):** that takes as input $\land$ or $\lor$ operator and two Büchi automata $BA_l$ and $BA_r$, and returns a Büchi automaton defined according to definitions of $\land$ and $\lor$ given in Section 2;

b) **BuchiUntil(\theta, BA):** that takes as input a propositional formula $\theta$ and a Büchi automaton $BA$ and returns the automaton as defined in Definition 7 (Section 4);

e) **BuchiRelease(\theta, BA):** that takes as input a propositional formula $\theta$ and a Büchi automaton $BA$ and returns the automaton as defined in Definition 9 (Section 4).
This paper presents a compositional algorithm for generating Büchi automata from a fragment of LTL logic. We firstly proposed the grammar of this fragment and then built for each formula $\varphi$, its equivalent automata. We secondly showed how to compositionally build from Büchi automata associated to each sub-formula, the Büchi automaton of the target formula. We thirdly showed the complexity and the correctness of our Büchi automata generation method.

**Future work**: several research lines can be continued from the present work. First, some temporal operators such as always, precedes or since are not considered in this paper, as an immediate perspective, we will study how to include these operators in our LTL fragment. Second, Dwyer's presents a translational semantics for his pattern properties. Indeed, for each pattern property, he associates an equivalent LTL formula. It will be interesting to study whether the translational semantics given by Dwyer is covered by our fragment. This will be done by comparing Büchi automata generated by LTL2BA proposed by Gastin with the Büchi automata generated by our algorithm.
REFERENCES


APPENDIX

Proof Theorem 1

Proof. It is well-known that Büchi automata are closed under union (see Definition 1) and intersection (see Definition 2) [9, 14, 10]. That means:

\[ \mathcal{L}(A_{\varphi_1} \lor A_{\varphi_2}) = \mathcal{L}(A_{\varphi_1}) \cup \mathcal{L}(A_{\varphi_2}) \quad (1) \quad \text{and} \quad \mathcal{L}(A_{\varphi_1} \land A_{\varphi_2}) = \mathcal{L}(A_{\varphi_1}) \cap \mathcal{L}(A_{\varphi_2}) \quad (2) \]

By hypothesis, \( \text{Words}(\varphi_1) = \mathcal{L}(A_{\varphi_1}) \) and \( \text{Words}(\varphi_2) = \mathcal{L}(A_{\varphi_2}) \), then

\[ \text{Words}(\varphi_1 \lor \varphi_2) = \mathcal{L}(A_{\varphi_1}) \cup \mathcal{L}(A_{\varphi_2}) \quad (3) \quad \text{and} \quad \text{Words}(\varphi_1 \land \varphi_2) = \mathcal{L}(A_{\varphi_1}) \cap \mathcal{L}(A_{\varphi_2}) \quad (4) \]

Hence, (1) + (3) leads to \( \text{Words}(\varphi_1 \lor \varphi_2) = \mathcal{L}(A_{\varphi_1 \lor \varphi_2}) \) and (2) + (4) leads to \( \text{Words}(\varphi_1 \land \varphi_2) = \mathcal{L}(A_{\varphi_1 \land \varphi_2}) \).

Proof Theorem 2

Proof. Each transformation rule provides a formula \( \varphi' \) whose the size \( |\varphi'| \) is the size of the origin formula \( \varphi \) plus one or minus one. Then, \( |\varphi'| = |\varphi| + k \) which leads to \( |\varphi'| = \mathcal{O}(|\varphi|) \).
Proof Theorem 3: $\Theta \cup \varphi$

Proof. Let first extend the transition function $\delta : s \times \Sigma \longrightarrow 2^S$ to the function $\delta^* : S \times \Sigma^* \longrightarrow 2^S$ as follows: $\delta^*(s, A) = \delta(s, A)$ and $\delta^*(s, A_0A_1A_2\ldots A_n) = \bigcup_{a \in \delta(s, A_0)} \delta^*(s^a, A_1A_2\ldots A_n)$.

Let us first prove that $\text{Words}(\psi) \subseteq \mathcal{L}_\omega(A_\psi)$. For this, let $\sigma \in \text{Words}(\psi)$, then $\sigma = A_0A_1A_2\ldots \in (2^S)^\omega$ such that $\sigma \models \Theta \cup \varphi$. By the definition of "until" operator, there exists a $j \geq 0$ such that:

$$A_iA_{i+1}A_{i+2}\ldots \models \Theta \text{ for all } 0 \leq i < j \text{ and } A_jA_{j+1}A_{j+2}\ldots \models \varphi$$

Distinguish between $j = 0$ and $j > 0$.

If $j = 0$, then $A_0A_1A_2\ldots \models \varphi$. Hence, $\sigma \in \text{Words}(\varphi)$. According to the hypothesis that $\text{Words}(\varphi) = \mathcal{L}_\omega(A_\varphi)$, we conclude $\sigma \in \mathcal{L}_\omega(A_\varphi)$. According to Rule 1 of Definition 7, adding $s_{\text{new}}$ to $S_{A_\varphi}$ does not affect the accepted language of $A_\varphi$. Hence, there exists a run $s_{\text{new}}s_1s_2s_3\ldots$ for $\sigma$ such that $s_k \in \delta^*(s_{\text{new}}, \sigma(0 : k - 1))$ for all $k \geq 1$. By construction of $A(\Theta \cup \varphi)$, it is obvious that $s_{\text{new}}s_1s_2\ldots$ is also a run for $\sigma$ in $A(\Theta \cup \varphi)$. Consequently, $\sigma \in \mathcal{L}_\omega(A_\psi)$.

If $j > 0$, then $A_0A_1\ldots A_{j-1} \models \Theta$ and $A_jA_{j+1}\ldots \models \varphi$. $A_jA_{j+1}\ldots \models \varphi$ yields that $A_jA_{j+1}\ldots \in \text{Words}(\varphi)$ which in turn yields that $A_jA_{j+1}\ldots \in \mathcal{L}_\omega(A_\varphi)$ (by hypothesis). Then, there exists a run $s_js_{j+1}s_{j+2}\ldots$ for $A_jA_{j+1}A_{j+2}\ldots$ in $A_\varphi$ such that $s_k \in \delta^*(s_j, A_{k-1})$ for all $k \geq j$ and $s_j$ is the initial state of $A_\varphi$. According to Rule 1 of Definition 7, adding $s_{\text{new}}$ to $S_{A_\varphi}$ does not affect the accepted language of $A_\varphi$. Then:

$$s_{\text{new}}s_{j+1}s_{j+2}\ldots \text{ is a run for } A_jA_{j+1}A_{j+2}\ldots \text{ in } A_\psi \quad (1)$$

By construction of $A_\psi$, there is a loop over $s_{\text{new}}$ (Rule 3 of Definition 7) labeled with $\Theta$ and $s_{\text{new}}$ is the only initial state. Since, $A_iA_{i+1}A_{i+2}\ldots \models \Theta$, then:

$$s_{\text{new}}s_{\text{new}}s_{\text{new}}\text{ is a run for } A_iA_{i+1}A_{i+2}\ldots \quad (2)$$

(1) + (2) leads to:

$$s_{\text{new}}s_{j+1}s_{j+2}\ldots \text{ is a run for } A_iA_{i+1}A_{i+2}\ldots A_jA_{j+1}A_{j+2}\ldots$$

Hence, $A_iA_{i+1}\ldots A_jA_{j+1}\ldots \in \mathcal{L}_\omega(A_\psi)$. Finally, $\text{Words}(\psi) \subseteq \mathcal{L}_\omega(A_\psi)$. 

50
Now, let us prove that $L_\omega(A_\psi) \subseteq \text{Words}(\psi)$. For this, let $\sigma = A_0 A_1 A_2 \ldots \in L_\omega(A_\psi)$, then there exists an accepting run $s_0 s_1 s_2 s_3 \ldots$ for $\sigma$ in $A_\psi$ such that $\forall i, i \geq 1, s_i \in \delta^*(s_0, \sigma(i : 1 :))$. We distinguish two cases:

Case 1: if the run $s_0 s_1 s_2 \ldots$ in $A_\psi$ does not go thought the loop over the initial state $s_{\text{new}}$. That means $s_0 = s_{\text{new}}$ and $\forall i \geq 1, s_i = s_{\text{new}}$. In this case, $s_0 s_1 s_2 \ldots$ is also a run of $\sigma$ in $A_\varphi$ according to Rule 2 of Definition 7. Then, $\sigma \in L_\omega(A_\varphi)$. Since $L_\omega(A_\varphi) \subseteq \text{Words}(A_\varphi)$ (by hypothesis), then $\sigma \models \varphi$. $\varphi_1 \cup \varphi_2 \equiv \varphi_2 \lor \chi(\varphi_1 \cup \varphi_2)$, we can conclude that $\sigma \models \Theta \cup \varphi$.

Case 2: if the run $s_0 s_1 s_2 \ldots$ in $A_\psi$ goes through the loop over the initial state $s_{\text{new}}$. In this case:

- there exists a $j, j \geq 0$ such that $\forall i, 0 \leq i < j, \exists s_i$ such that $s_i = s_{\text{new}}, A_i = \Theta$ and $s_i \in \delta^*(s_{i-1}, \sigma(0 : i))$ (1)
- and $\forall k, k \geq j, \exists s_k$ such that $s_k \in \delta^*(s_{k-1}, A_k)$ (2)

From (1), we can conclude that:

$$\forall i, 0 \leq i < j, A_i A_{i+1} A_{i+2} A_{i+3} \ldots \models \Theta \quad (3)$$

From (2), we can conclude that $s_j s_{j+1} s_{j+2} \ldots$ is a run for $A_j A_{j+1} A_{j+2} \ldots$ in $A_\varphi$ since $s_j = s_{\text{new}}$. By hypothesis, since $L_\omega(A_\varphi) \subseteq \text{Words}(A_\varphi)$, then:

$$A_j A_{j+1} A_{j+2} \ldots \models \varphi \quad (4)$$

From (3) + (4), we can conclude that $\sigma \models \text{Words}(\psi)$.

**Proof Theorem 4: X\varphi**

**Proof.** Let us first prove that $\text{Words}(\psi) \subseteq L_\omega(A_\psi)$. For this, let $\sigma \in \text{Words}(\psi)$ then $\sigma = A_0 A_1 A_2 \ldots \in (2^L)^w$ such that $\sigma \models X\varphi$. By the definition of "next" operator $A_1 A_2 A_3 \ldots \models \varphi$, then there exists a run $s_0 s_1 s_2 s_3 \ldots$ for $\sigma(1 :)$ such that $s_k \in \delta^*(s_{k-1}, A_k) \models (2^L)^w$ for all $k \geq 1$. According to Rule 1 of Definition 7, adding $s_{\text{new}}$ to $S_{\text{new}}$ does not affect the accepted language of $A_\psi$. Then $s_{\text{new}} s_1 s_2 s_3 \ldots$ for $\sigma(1 :) \models A_\varphi$. By construction of $A_\psi$, $s_{\text{init}} s_{\text{new}} s_1 s_2 \ldots$ is a run for $A_0 A_1 A_2 \ldots$ in $A_\psi$. Hence, we can conclude that $s_{\text{init}} s_{\text{new}} s_1 s_2 \ldots$ is also a run for $A_0 A_1 A_2 \ldots$ in $A_\psi$. Consequently, $\sigma \in L_\omega(A_\psi)$.

Now, let us prove that $L_\omega(A_\psi) \subseteq \text{Words}(\psi)$. For this, let $\sigma = A_0 A_1 A_2 \ldots \in L_\omega(A_\psi)$, then there exists an accepting run $s_0 s_1 s_2 s_3 \ldots$ for $\sigma$ in $A_\psi$ such that $\forall i, i \geq 1, s_i \in \delta^*(s_0, \sigma(i : 1 :))$ and $s_{\text{init}} = s_0$ and $s_1 = s_{\text{new}}$. By construction of $A_\psi$, it is not difficult to see that $s_1 s_2 s_3 \ldots$ is a run for $\sigma(1 :) \models A_\psi$. By construction, we can conclude that $\sigma(1 :) \models \text{Words}(\varphi)$ which means that $\sigma(1 :) \models \varphi$. According to the semantics of "next" operator, $\sigma \models X\varphi$, consequently, $\sigma \models \text{Words}(\psi)$.
Proof Theorem 5: $\varphi \text{ R } \varphi$

Proof. Let denote by $\mathbb{G}$ the Büchi automata language recognizing the fact that $\Theta$ should be true forever and by $\mathbb{G}$ the set of words equivalent to the formula “$\Theta$ should be true forever”. It is straightforward to see that $\mathbb{G} \equiv \mathbb{G}$ (1).

By the construction given in Definition 9, we can conclude that $L(A_{\varphi \text{ R } \Theta})$ is equivalent to $L(A_{\Theta \cup \varphi}) \cup \mathbb{G}$ (3). By Theorem 3, we have that $\text{Words}(A_{\Theta \cup \varphi}) \equiv L(A_{\Theta \cup \varphi})$ (2).

By Theorem 1, (1) + (2) leads to $\text{Words}(A_{\Theta \cup \varphi}) \cup \mathbb{G} \equiv L(A_{\Theta \cup \varphi}) \cup \mathbb{G}$. But, the semantics of $\varphi \text{ R } \Theta$ is $\text{Words}(A_{\Theta \cup \varphi}) \cup \mathbb{G}$. Then, we can conclude that $\text{Words}(\varphi \text{ R } \Theta) \equiv L(A_{\varphi \text{ R } \Theta})$.

Proof Theorem 6

Proof. Let $|\varphi|$ denote the length of flat LTL formulae $\varphi$ in terms of the number of operators in $\varphi$. It is obvious that $|\Theta| = 0$. Let us first compute the complexity of converting a $\text{FST}(\varphi)$ to a Büchi automaton. For this, let $T(m)$ be the number of states of the Büchi automaton generated by Algorithm 2 for a flat LTL formula $\psi$ with $m = |\varphi|$. We can easily see that Algorithm 2 satisfies the following recurrence:

$$
T(m) = \begin{cases} 
2 & \text{if } \text{Root(}\text{FST}(\varphi)\text{)} = \Theta \\
2 + T(m - 1) & \text{if } \text{Root(}\text{FST}(\varphi)\text{)} = X \\
1 + T(m - 1) & \text{if } \text{Root(}\text{FST}(\varphi)\text{)} = U \\
2 + T(m - 1) & \text{if } \text{Root(}\text{FST}(\varphi)\text{)} = R \\
1 + 2T(m/2) & \text{if } \text{Root(}\text{FST}(\varphi)\text{)} = V \\
e + 2T(m/2) & \text{if } \text{Root(}\text{FST}(\varphi)\text{)} = \land 
\end{cases}
$$

By master theorem, we can conclude our algorithm lies in $O(|\varphi|)$.

Proof Theorem 7

Proof. By structural induction on the structure of $\psi$:

Basic case: $\text{FST}(\psi)$ is a single node. Then, its root is an atomic formula which means that $\psi = \Theta$. It is obvious that $\text{Words}(\Theta) = L_\omega(A_\Theta)$

General case: many cases have to be considered:

- If the root of $\text{FST}(\psi)$ is equal to $X$, then $\psi = X\varphi$. By induction hypothesis, we have that $\text{Words}(\varphi) = L_\omega(A_\varphi)$. By Theorem 4, we can conclude that $\text{Words}(\psi) = L_\omega(A_\psi)$.
- If the root of $\text{FST}(\psi)$ is equal to $U$, then $\psi = \Theta \cup \varphi$. By induction hypothesis, we have that $\text{Words}(\varphi) = L_\omega(A_\varphi)$. By Theorem 3, we can conclude that $\text{Words}(\psi) = L_\omega(A_\psi)$.
- If the root of $\text{FST}(\psi)$ is equal to $R$, then $\psi = \varphi \text{ R } \Theta$. By induction hypothesis, we have that $\text{Words}(\varphi) = L_\omega(A_\varphi)$. By Theorem 5, we can conclude that $\text{Words}(\psi) = L_\omega(A_\psi)$.
- If the root of $\text{FST}(\psi)$ is equal to $V$ (resp. $\land$), then $\psi = \varphi_1 \land \varphi_2$ (resp. $\psi = \varphi_1 \lor \varphi_2$). By induction hypothesis, we have that $\text{Words}(\varphi_1) = L_\omega(A_{\varphi_1})$ and $\text{Words}(\varphi_2) = L_\omega(A_{\varphi_2})$. By Theorem 1, we can conclude that $\text{Words}(\psi) = L_\omega(A_\psi)$. 

52