A New Translation from ECTL* to SAT

Andrzej Zbrzezny
IMCS, Jan Długosz University in Częstochowa,
Al. Armii Krajowej 13/15, 42-200 Częstochowa, Poland.
a.zbrzezny@ajd.czest.pl

Abstract. In this paper we present a new translation from ECTL* to SAT and show that the proposed translation substantially increases the efficiency of verifying temporal properties using the Bounded Model Checking method. We have implemented our new translation and made experimental results, which demonstrate the efficiency of the method.

keywords: ECTL*, translation to SAT, Bounded Model Checking

1 Introduction

Bounded Model Checking (BMC) is a popular model checking technique for the verification of finite-state transition systems [2, 4]. BMC has been introduced as a technique complementary to symbolic BDD-based model checking. Given a finite transition system \( M \) and an ACTL* formula \( \alpha \), the BMC procedure tries to find a natural number \( k \) and a set of finite paths of length \( k \) in \( M \), the so called \( k \)-paths, such that an ECTL* formula, namely, the negation of the formula \( \alpha \) holds in \( M \). In such a case it means that \( M \) violates the property \( \alpha \). SAT-based BMC is performed by generating a propositional formula which is satisfiable if and only if such a number \( k \) and a set of \( k \)-paths exist. This formula is usually obtained by combining an encoding of the transition relation of the model and a translation of the negation of \( \alpha \). If for a given \( k \) this propositional formula is not satisfiable, then \( k \) is incremented until either the problem becomes intractable due to the complexity of solving the corresponding SAT instance or \( k \) reaches the size of the model. The main advantage of BMC comparing to BDDs is its space efficiency. Moreover, BMC produces counterexamples of minimal length, which eases their interpretation and understanding for debugging purposes.

In order to use the BMC method one needs to define a translation from a given temporal logic to the satisfiability problem (in short: to SAT). There are several translations proposed in literature for different kinds of temporal logics.

The first translation from LTL to SAT was introduced in [2] and another ones in [9] and [3]. The first translation from ECTL to SAT was introduced in [10] and then it was substantially improved in [12]. The first correct translation from ECTL* to SAT was introduced in [11]. A brief description of the idea of this translation is as follows: given an ECTL* formula \( \varphi \) and a natural number \( k \), first the the number of \( k \)-paths (paths of length \( k \)) sufficient for checking the formula \( \varphi \) is computed by using a function \( f_k \). Next, the set of \( f_k(\varphi) \) \( k \)-paths is created and this set is used to translate every
subformula of the formula \( \varphi \). This means that the translation uses all the \( f_k(\varphi) \) \( k \)-paths both for the formula \( \varphi \) and for each of its proper subformulae.

In this paper we introduce a new translation from ECTL* to SAT. The main idea of our translation consists in translating every subformula \( \psi \) of \( \varphi \) using only \( f_k(\psi) \) \( k \)-paths. More precisely, given a formula \( \varphi \) and a set \( A \) of \( k \)-paths such that \( |A| = f_k(\varphi) \) we divide the set \( A \) into subsets needed for translating the subformulae of \( \varphi \). In particular, in order to translate a temporal subformula \( \psi_1 O \psi_2 \), where \( O \in \{U,R\} \), on a given \( k \)-path \( \pi \), we use \( \pi \) and exactly \( f_k(\psi_1 O \psi_2) \) \( k \)-paths for translating the subformula \( \psi_1 O \psi_2 \), whereas the old translation uses all the \( f_k(\varphi) \) \( k \)-paths for translating the subformula in question. Generally, if \( \psi_1 O \psi_2 \) is a subformulae of \( \varphi \), then \( f_k(\psi_1 O \psi_2) < f_k(\varphi) \).

Another novelties provided in this paper are: a simpler definition of bounded semantics and the treatment of the so called loops, which results in significantly simpler propositional formulae being the results of the translation. In particular, unlike the method described in [11], our translation does not require that either all the \( k \)-paths used in the translation are loops or none is a loop.

We provide some experimental results which clearly confirm that the improvement in question substantially increases the efficiency of the bounded model checking of ECTL* formulae.

The rest of the paper is organised as follows. The next section recalls the syntax and (both the bounded and unbounded) semantics of ECTL*. In Section 3 our improved translation from ECTL* to SAT is defined. A sample experimental results are described in Section 4. Section 5 provides concluding remarks.

2 Syntax and Semantics of ECTL*

In this section we briefly recall the syntax and semantics of the logic ECTL*. For more thorough description one can see [1].

The Existential Computation Tree Logic ECTL* is a restriction of a propositional branching-time temporal logic CTL* that was introduced by Emerson and Halpern in [7] as a specification language for finite-state systems. The restriction consists in using only existential path quantifiers and allowing the negation to be applied to propositional variables only (instead of to arbitrary formulae).

2.1 Syntax of ECTL*

The language of ECTL* consists of two types of formulae: state formulae (interpreted at states) and path formulae (interpreted along paths). The syntax of ECTL* formulae is defined by the following rules, where \( AP \) is an arbitrary set of atomic propositions:

- \texttt{true} and \texttt{false} are state formulae,
- if \( p \in AP \), then \( p \) and \( \neg p \) are state formulae,
- if \( \alpha \) and \( \beta \) are state formulae, then \( \alpha \land \beta \) and \( \alpha \lor \beta \) are state formulae,
- if \( \varphi \) is a path formula, then \( E\varphi \) is a state formula,
- if \( \alpha \) is a state formula, then \( \alpha \) is also a path formula,
• if $\phi$ and $\psi$ are path formulae, then $\phi \land \psi, \phi \lor \psi, X\phi, \phi U \psi$ and $\phi R \psi$ are path formulae.

In practice, many interesting temporal properties are formulated by using temporal operators $F$ (eventually) and $G$ (always) defined as follows: $F\psi \stackrel{df}{=} \text{true} U \psi$, and $G\psi \stackrel{df}{=} \text{false} R \psi$.

2.2 Semantics of ECTL*

The semantics of ECTL* formulae is determined with respect to a transition system (see [1]).

**Definition 1.** A transition system is a tuple $M = (S, \text{Act}, \rightarrow, I, AP, L)$, where $S$ is a nonempty finite set of states, $\text{Act}$ is a set of actions, $I \subseteq S$ is a set of initial states, $\rightarrow \subseteq S \times \text{Act} \times S$ is a transition relation, $AP$ is a set of atomic propositions, and $L : S \rightarrow 2^{AP}$ is a labelling function that assigns to each state a set of atomic propositions that are assumed to be true at that state. Transition systems are also called models.

For convenience, we shall write $s \xrightarrow{\sigma} s'$ instead of $(s, \sigma, s') \in \rightarrow$. Moreover, we shall write $s \rightarrow s'$ if $s \xrightarrow{\sigma} s'$ for some $\sigma \in \text{Act}$.

In the following it is assumed that a transition system has no terminal states, i.e., for every $s \in S$ there exist $s' \in S$ such that $s \rightarrow s'$. The set of all natural numbers is denoted by $\mathbb{N}$ and the set of all positive natural numbers by $\mathbb{N}_+$. A path in $M$ is an infinite sequence $\pi = (s_0, s_1, \ldots)$ of states such that $s_j \xrightarrow{\pi} s_{j+1}$ for each $j \in \mathbb{N}$. For a path $\pi = (s_0, s_1, \ldots)$, let $\pi(j) = s_j$ and $\pi^j = (s_j, s_{j+1}, \ldots)$, for each $j \in \mathbb{N}$. Note that if $\pi$ is a path in $M$ then the suffix $\pi^j$ is also a path in $M$. The set of all the paths in a model $M$ will be denoted by $P(M)$.

Let $M = (S, \text{Act}, \rightarrow, I, AP, L)$ be a model, $s$ be a state and $\pi$ be a path. For a state formula $\alpha$ over $AP$, the notation $M, s \models \alpha$ means that $\alpha$ holds at the state $s$ in the model $M$. Similarly, for a path formula $\varphi$ over $AP$, the notation $M, \pi \models \varphi$ means that $\varphi$ holds along the path $\pi$ in the model $M$.

**Definition 2.** Let $p$ be an atomic proposition, $\alpha, \beta$ be state formulae of ECTL*, and $\varphi, \psi$ be path formulae of ECTL*. The relation $\models$ is defined inductively as follows:

\[
\begin{align*}
M, s &\models \text{true}, \\
M, s &\not\models \text{false}, \\
M, s &\models p \quad \text{iff} \quad p \in L(s), \\
M, s &\models \neg p \quad \text{iff} \quad p \notin L(s), \\
M, s &\models \alpha \land \beta \quad \text{iff} \quad M, s \models \alpha \quad \text{and} \quad M, s \models \beta, \\
M, s &\models \alpha \lor \beta \quad \text{iff} \quad M, s \models \alpha \quad \text{or} \quad M, s \models \beta, \\
M, \pi &\models E \varphi \quad \text{iff} \quad \exists \pi \in P(M) \ (\pi(0) = s \quad \text{and} \quad M, \pi \models \varphi), \\
M, \pi &\models \alpha \quad \text{iff} \quad M, \pi(0) \models \alpha, \\
M, \pi &\models \varphi \land \psi \quad \text{iff} \quad M, \pi \models \varphi \quad \text{and} \quad M, \pi \models \psi, \\
M, \pi &\models \varphi \lor \psi \quad \text{iff} \quad M, \pi \models \varphi \quad \text{or} \quad M, \pi \models \psi, \\
M, \pi &\models X \varphi \quad \text{iff} \quad M, \pi^1 \models \varphi,
\end{align*}
\]
\( \mathcal{M}, \pi \models \varphi \mathbf{U} \psi \iff (\exists j \geq 0)(\mathcal{M}, \pi^j \models \psi \text{ and } (\forall 0 \leq i < j) \mathcal{M}, \pi^i \models \varphi) \),
\( \mathcal{M}, \pi \models \varphi \mathbf{R} \psi \iff (\exists j \geq 0)(\mathcal{M}, \pi^j \models \varphi \text{ and } (\forall 0 \leq i < j) \mathcal{M}, \pi^i \models \psi) \)

or \( (\forall j \geq 0) \mathcal{M}, \pi^j \models \psi \).

Note 1. From Definition 2 and the definitions of the operators \( \mathbf{F} \) and \( \mathbf{G} \) it follows that
\( \mathcal{M}, \pi \models \mathbf{F} \psi \iff (\exists j \geq 0)(\mathcal{M}, \pi^j \models \psi), \mathcal{M}, \pi \models \mathbf{G} \psi \iff (\forall j \geq 0)(\mathcal{M}, \pi^j \models \psi) \).

Definition 3. An ECTL\(^*\) state formula \( \alpha \) is valid in \( \mathcal{M} \), denoted by \( \mathcal{M} \models \alpha \), iff for each \( s \in I \), \( \mathcal{M}, s \models \alpha \), i.e., \( \alpha \) holds at every initial state of \( \mathcal{M} \).

2.3 Bounded Semantics of ECTL\(^*\)

In this section we define a bounded semantics of ECTL\(^*\). This is done in order to define the bounded model checking problem for ECTL\(^*\) and to translate it into the satisfiability problem.

From now on we assume that models are finite, i.e., the sets \( S, \text{Act} \) and \( \text{AP} \) are finite. To define the bounded semantics one needs to represent infinite paths in a model in a special way. To this aim, we define the notions of \( k \)-paths and loops.

Definition 4. Let \( \mathcal{M} \) be a model, \( k \in \mathbb{N} \), and \( 0 \leq l \leq k \). A \( k \)-path is a pair \((\pi, l)\), also denoted by \( \pi_l \), where \( \pi \) is a finite sequence \( \pi = (s_0, \ldots, s_k) \) of states such that \( s_j \rightarrow s_{j+1} \) for each \( 0 \leq j < k \). A \( k \)-path \( \pi_l \) is a loop if \( l < k \) and \( \pi(k) = \pi(l) \).

The set of all \( k \)-paths in a model \( \mathcal{M} \) will be denoted by \( \Pi_k(\mathcal{M}) \). Note that the set \( \Pi_k(\mathcal{M}) \) is finite since \( \mathcal{M} \) is finite. Although every \( k \)-path \( \pi_l \) is finite, it still can represent an infinite path if it is a loop. Namely, \( \pi_l \) represents the infinite path of the form \( u\omega^\infty \), where \( u = (\pi(0), \ldots, \pi(l-1)) \) and \( \omega = (\pi(l), \ldots, \pi(k-1)) \).

As in the definition of semantics one needs to define the satisfiability relation on suffixes of \( k \) paths, we denote by \( \pi_l^m \) the pair \((\pi_l, m)\), i.e. the \( k \)-path \( \pi_l \) together with the designated starting point \( m \), where \( 0 \leq m \leq k \).

Let \( s \) be a state and \( \pi_l \) a \( k \)-path. For a state formula \( \alpha \) over \( \text{AP} \), the notation \( \mathcal{M}, s \models_k \alpha \) means that \( \alpha \) \( k \)-holds at the state \( s \) in the model \( \mathcal{M} \). Similarly, for a path formula \( \varphi \) over \( \text{AP} \), the notation \( \mathcal{M}, \pi_l^m \models_k \varphi \), where \( 0 \leq m \leq k \), means that \( \varphi \) \( k \)-holds along the suffix \( \{\pi(m), \ldots, \pi(k)\} \) of \( \pi \). For convenience, in the following definition we shall write \( s \models_k \alpha \) instead of \( \mathcal{M}, s \models_k \alpha \) and \( \pi_l^m \models_k \varphi \) instead of \( \mathcal{M}, \pi_l^m \models_k \varphi \).

Definition 5 (Bounded semantics). Let \( p \) be an atomic proposition, \( \alpha, \beta \) be state formulae of ECTL\(^*\), and \( \varphi, \psi \) be path formulae of ECTL\(^*\). The relation \( \models_k \) is defined inductively as follows:

\begin{align*}
  s \models_k \text{true}, \\
  s \not\models_k \text{false}, \\
  s \models_k p \quad &\text{iff } p \in L(s), \\
  s \models_k \neg p \quad &\text{iff } p \notin L(s), \\
  s \models_k \alpha \land \beta \quad &\text{iff } s \models_k \alpha \text{ and } s \models_k \beta, \\
  s \models_k \alpha \lor \beta \quad &\text{iff } s \models_k \alpha \text{ or } s \models_k \beta,
\end{align*}
Let $\pi$ be a state formula. Then 
$$\pi \models \mathbf{F} \psi \quad \text{iff} \quad (\exists m \leq j \leq k) \; \pi_i \models \psi$$
$$\pi \models \mathbf{G} \psi \quad \text{iff} \quad (l < m \land \pi(k) \models \psi \land (\exists l < j < m) \; \pi_i \models \psi$$

Note 2. From Definition 5 and the definitions of the operators $\mathbf{F}$ and $\mathbf{G}$ it follows that

$$\pi_i \models \mathbf{F} \psi \quad \text{iff} \quad (\exists m \leq j \leq k) \; \pi_i \models \psi$$
$$\pi_i \models \mathbf{G} \psi \quad \text{iff} \quad (l < m \land \pi(k) \models \psi \land (\exists l < j < m) \; \pi_i \models \psi$$

Example 1. Let $\mathcal{M}_1 = \{a, b, c\}, \{\tau\}, \rightarrow, \{a\}, \{p\}, L$ be a transition system such that $\rightarrow = \{\{a, b, c\}, \{a, \tau, c\}, \{c, \tau, a\}\}, L(a) = L(c) = \{p\}$ and $L(b) = \emptyset$. Consider the path formula $\varphi = \mathbf{X} \mathbf{G} \psi$ and the sequence $\pi = \{a, b, c, a\}$. It is easy to check that $\mathcal{M}_1, \pi_0 \models \not\varphi$ but $\mathcal{M}_1, \pi_2 \models \varphi$.

Example 2. Let $\mathcal{M}_2 = \{a, b\}, \{\tau\}, \rightarrow, \{a\}, \{p, q\}, L$ be a transition system such that $\rightarrow = \{\{a, \tau, b\}, \{b, \tau, a\}\}, L(a) = \{p\}$ and $L(b) = \{q\}$. Notice that the only initial path in $\mathcal{M}_2$ is of the form $(ab)^\omega$.

It is easily seen that the state formula $\alpha = \mathbf{E} \mathbf{G} \mathbf{X} \psi$ holds in $\mathcal{M}_2$. In order to check that $\alpha$ holds in $\mathcal{M}_2$ we observe that $\mathcal{M}_2, a \models \alpha$ as $\mathcal{M}_2, \pi_0 \models \mathbf{G} \mathbf{X} \psi$, where $\pi = \{a, b, a\}$. This example shows that the second component of the condition for the bounded satisfiability of $\mathbf{X} \psi$ cannot be avoided.

It is easily seen that the state formula $\beta = \mathbf{E} \mathbf{G} \mathbf{F} \psi$ holds in $\mathcal{M}_2$. In order to check that $\beta$ holds in $\mathcal{M}_2$ we observe that $\mathcal{M}_2, a \models \beta$ as $\mathcal{M}_2, \pi_0 \models \mathbf{G} \mathbf{F} \psi \land \mathbf{G} \psi$, where $\pi = \{a, b, a\}$. This example shows that the second component of the condition for the bounded satisfiability of $\mathbf{F} \psi$ also cannot be avoided.

Definition 6. An ECTL* state formula $\alpha$ is $k$-valid in $\mathcal{M}$, denoted by $\mathcal{M} \models_k \alpha$, iff for each $s \in I, \mathcal{M}, s \models_k \alpha$, i.e., $\alpha$ $k$-holds at every initial state of $\mathcal{M}$.

The following theorem states that for a given model $\mathcal{M}$ and a formula $\alpha$ there exists a bound $k$ such that the model checking problem $\mathcal{M} \models \alpha$ can be reduced to the bounded model checking problem $\mathcal{M} \models_k \alpha$.

Theorem 1. Let $\mathcal{M}$ be a model and $\alpha$ be an ECTL* state formula. Then $\mathcal{M} \models \alpha$ iff for some $k \in \mathbb{N}$, $\mathcal{M} \models_k \alpha$.

The proof of Theorem 1 will be provided in the full version of the present paper.
3 Translation to SAT

First, we briefly describe how to translate the problem of \( k \)-validity of an ECTL\(^*\) formula to the problem of satisfiability of some propositional formula. Next, we describe the new translation from ECTL\(^*\) formulae to propositional formulae built over a set \( PV \) of propositional variables plus the constants \text{true} and \text{false}, with help of the propositional connectives \symbol{\neg}, \land, \lor and \rightarrow.

3.1 Boolean encoding

We begin with the encoding of the transitions of the given model \( M \). Since the set of states of \( M \) is finite, every element of \( S \) can be encoded as a bit vector of some length \( t \) depending on the number of states of \( M \). Then, each state of \( M \) can be represented by a valuation of a vector \( w = (w_{j_1}, \ldots, w_{j_r}) \) (called a symbolic state) of different propositional variables called propositional state variables.

The designated position \( l \) of any \( k \)-path can be encoded as a bit vector of the length \( t = \max(1, \lceil \log_2(k + 1) \rceil) \). Therefore, such a position \( l \) of the path of the number \( j \) can be represented by the valuation of a vector \( u = (u_{j_1}, \ldots, u_{j_i}) \) (called a symbolic number) of different propositional variables, called propositional natural variables, not appearing among propositional state variables. In the following we denote by \( \text{Var}(w) \) the set of propositional state variables appearing in \( w \) and by \( \text{Var}(u) \) the set of propositional natural variables appearing in \( u \).

From now on let \( SV \) be a set of propositional variables used as propositional state variables and \( NV \) be a set of propositional variables used as propositional natural variables. We assume that \( SV \cap NV = \emptyset \). Let \( PV = SV \cup NV \) and \( V : PV \to \{0, 1\} \) be a valuation of propositional variables (a valuation for short). Each valuation \( V \) induces the functions \( S : SV^r \to \{0, 1\}^r \) and \( J_t : NV^t \to \mathbb{N} \) defined in the following way:

\[
S(w_{j_1}, \ldots, w_{j_r}) = (V(w_{j_1}), \ldots, V(w_{j_r})), \quad J_t((u_{j_1}, \ldots, u_{j_i})) = \sum_{i=1}^{t} V(u_{j_i}) \cdot 2^{i-1}.
\]

A pair consisting of a finite sequence of length \( k \) of symbolic states and of a symbolic number is called a symbolic \( k \)-path. In order to translate an ECTL\(^*\) formula \( \varphi \) one needs to use not just one but exactly \( f_k(\varphi) \) symbolic \( k \)-paths, where \( f_k \) is a function defined below. In the following, \( F_k(\varphi) \) denotes the set \( \{ j \in \mathbb{N} \mid 1 \leq j \leq f_k(\varphi) \} \).

Now, for a given \( k \in \mathbb{N} \), let \( W_{r,k} \) be a fixed finite set of global state variables \( w_{m,n} \in PV^r \), where \((m, n) \in \{0, 0\} \cup \{0, \ldots, k\} \times F_k(\varphi) \), such that \( \text{Var}(w_{m,n}) \cap \text{Var}(w_{m',n'}) = \emptyset \), provided \((m, n) \neq (m', n') \). Moreover, let \( N_k \) be a fixed finite set of natural variables \( u_j \in PV^t \), where \( j \in F_k(\varphi) \), such that \( \text{Var}(u_j) \cap \text{Var}(u_{j'}) = \emptyset \), provided \( j \neq j' \). We define for each \( j \in F_k(\varphi) \) the \( j \)-th symbolic \( k \)-path as \( \pi_j = ((w_{0,j}, \ldots, w_{k,j}), u_j) \).

Furthermore, let \( w \) and \( w' \) be symbolic states from \( W_{r,k} \), \( u \) be a symbolic number from \( N_k \), and \( \pi_j \) be the \( j \)-th symbolic \( k \)-path. Given a model \( M \) we define the following propositional formulae:

\[
\begin{align*}
&- T_p(w), \text{ where } p \in AP, \text{ as a formula such that for each } V \in \{0, 1\}^{PV}, V \text{ satisfies } P_p(w) \iff p \in L(S(w)), \\
&- I(w) \text{ as a formula such that for each } V \in \{0, 1\}^{PV}, V \text{ satisfies } I(w) \iff S(w) \in I,
\end{align*}
\]
\( T(w, w') \) as a formula such that for each \( V \in \{0, 1\}^{PV} \), \( V \) satisfies \( T(w, w') \) iff \( S(w) \rightarrow S(w') \) in \( M \).

- \( H(w, w') \) as a formula such that for each \( V \in \{0, 1\}^{PV} \), \( V \) satisfies \( H(w, w') \) iff both \( S(w) \) and \( S(w') \) are states of \( M \) and \( S(w) = S(w') \).

- \( B_j^\sim(u) \) as a formula such that for each \( V \in \{0, 1\}^{PV} \), \( V \) satisfies \( B_j^\sim(u) \) iff \( j \sim J(u) \), where \( \sim \in \{<, \leq, =, \geq, >\} \).

- \( E_k^\bullet(u) \) as a formula \( H(w_{k,n}, w_{l,n}) \land B_l^\sim(u) \), which is satisfied by a valuation \( V \) iff \( S(w_{k,n}) = S(w_{l,n}) \) and \( l = J(u) \).

### 3.2 The new translation

In order to translate a formula \( \varphi \), one needs to know the number of \( k \)-paths sufficient to translate this formula. To this aim, the function \( f_k \), first introduced in [11], is used.

**Definition 7.** Let \( F_p \) be the set of all the ECTL\(^*\) path formulae built over a given set \( AP \) of atomic propositions. The function \( f_k : F_p \rightarrow \mathbb{N} \) is defined as follows:

\[
\begin{align*}
  f_k(\text{true}) &= 0, \\
  f_k(\text{false}) &= 0, \\
  f_k(\varphi) &= 0, \text{ if } \varphi \in AP, \\
  f_k(\varphi \lor \psi) &= \max\{f_k(\varphi), f_k(\psi)\}, \\
  f_k(\varphi \land \psi) &= f_k(\varphi) + f_k(\psi), \\
  f_k(\psi^\bullet) &= f_k(\psi) + 1, \\
  f_k(\psi^\wedge) &= f_k(\psi) + f_k(\psi), \\
  f_k(\psi^\mathcal{R}) &= (k + 1) \cdot f_k(\psi) + f_k(\varphi).
\end{align*}
\]

For every ECTL\(^*\) path formula \( \varphi \) the function \( f_k \) determines how many symbolic \( k \)-paths are needed for translating the formula \( \varphi \). Notice that the translation from [11] translates every subformula of the form \( \psi^\bullet \) using all the \( f_k(\psi) \) \( k \)-paths.

Our improvement consists in translating every subformula \( \psi \) of \( \varphi \) by using only \( f_k(\psi) \) \( k \)-paths. For this, given a formula \( \varphi \) and a set \( A \) of \( k \)-paths such that \( |A| = f_k(\varphi) \), we divide the set \( A \) into subsets needed for translating the subformulas of \( \varphi \). In particular, for translating a temporal subformula of the form \( \psi^\bullet \) we choose exactly one path for translating the operator \( \psi \) and use the remaining \( k \)-paths for translating \( \psi \).

In order to divide \( A \) into proper subsets we need some auxiliary functions. In the definitions of these functions we will use the relation \( \prec \) defined on the set of all finite subsets of \( \mathbb{N} \) in the following way: \( A \prec B \) iff for all natural numbers \( x \) and \( y \), \( x \in A \) and \( y \in B \) implies \( x < y \). Note that from the definition of \( \prec \) it follows that \( A \prec B \) iff either \( A = \emptyset \) or \( B = \emptyset \) or \( A \neq \emptyset, B \neq \emptyset \), \( A \cap B = \emptyset \) and \( \max(A) < \min(B) \).

In order to explain the purpose of the auxiliary functions defined above let us recall that each ECTL\(^*\) path formula \( \varphi \) will be translated by using a set of exactly \( f_k(\varphi) \) symbolic \( k \)-paths. In the following we will say that a set \( A \) of positive natural numbers is used to translate a formula \( \varphi \) instead of saying that the set of symbolic \( k \)-paths whose numbers are in \( A \) is used to translate the formula \( \varphi \). Moreover, saying that a set \( A \) is used to translate a formula \( \varphi \) assumes that \( |A| = f_k(\varphi) \).

Now, let \( A \) be a finite nonempty set of natural numbers, and \( k, p \) be natural numbers with \( p \leq |A| \). Then,

- \( g_l(A, p) \) denotes the subset \( B \) of \( A \) such that \( |B| = p \) and \( B \prec A \setminus B \),
- \( g_r(A, p) \) denotes the subset \( C \) of \( A \) such that \( |C| = p \) and \( A \setminus C \prec C \),
- \( g_c(A) \) denotes the set \( A \setminus \{\min(A)\} \),

- \( g_s(A) \) denotes the subset \( D \) of \( A \) such that \( |D| = p \) and \( D \prec A \setminus D \).

- \( g_d(A) \) denotes the subset \( E \) of \( A \) such that \( |E| = p \) and \( E \prec A \setminus E \).
• if $k$ divides $|A| - p$, then $h^U_k(A, p)$ denotes the sequence $(B_0, \ldots, B_k)$ of subsets of $A$ such that $\bigcup_{j=0}^k B_j = A$, $|B_0| = \ldots = |B_{k-1}|$, $|B_k| = p$, and $B_i \prec B_j$ for every $0 \leq i < j \leq k$,

• if $k + 1$ divides $|A| - p$, then $h^R_k(A, p)$ denotes the sequence $(B_0, \ldots, B_{k+1})$ of subsets of $A$ such that $\bigcup_{j=0}^{k+1} B_j = A$, $|B_0| = \ldots = |B_k|$, $|B_{k+1}| = p$, and $B_i \prec B_j$ for every $0 \leq i < j \leq k$.

If $h^U_k(A, p) = (B_0, \ldots, B_k)$, then $h^U_k(A, p)(j)$ denotes the set $B_j$, for every $0 \leq j \leq k$. Similarly, if $h^R_k(A, p) = (B_0, \ldots, B_{k+1})$, then $h^R_k(A, p)(j)$ denotes the set $B_j$, for every $0 \leq j \leq k + 1$.

The functions $g_l$ and $g_r$ are used in the translation of the formulae with the main connective being either conjunction or disjunction: for a given ECTL* formula $\varphi \land \psi$, if a set $A$ is to be used to translate this formula, then the set $g_l(A, f_k(\varphi))$ is used to translate the subformula $\varphi$ and the set $g_r(A, f_k(\psi))$ is used to translate the subformula $\psi$; for a given ECTL* formula $\varphi \lor \psi$, if a set $A$ is to be used to translate this formula, then the set $g_l(A, f_k(\varphi))$ is used to translate the subformula $\varphi$ and the set $g_r(A, f_k(\psi))$ is used to translate the subformula $\psi$.

The function $g_s$ is used in the translation of the formulae with the main connective $E$: for a given ECTL* formula of the form $E \varphi$, if a set $A$ is to be used to translate this formula, then the symbolic path $\pi_{\text{min}(A)}$ is used to translate the operator $E$ and the set $g_s(A)$ is used to translate the subformula $\varphi$ at the symbolic path $\pi_{\text{min}(A)}$.

The function $h^U_k$ is used in the translation of subformulae of the form $\varphi \text{ U } \psi$: if a set $A$ is to be used to translate the subformula $\varphi \text{ U } \psi$ at the symbolic $k$-path $\pi_n$ (with starting point $m$), then for every $i$ such that $m \leq j < i$, the set $h^U_k(A, f_k(\psi))(k)$ is used to translate the formula $\psi$ at the symbolic path $\pi_n$ with starting point $i$; moreover, for every $j$ such that $m \leq j < i$, the set $h^U_k(A, f_k(\psi))(j)$ is used to translate the formula $\varphi$ at the symbolic path $\pi_n$ with starting point $j$. Notice that if $k$ does not divide $|A| - p$, then $h^U_k(A, p)$ is undefined. However, for every set $A$ such that $|A| = f_k(\varphi \text{ U } \psi)$, it is clear from the definition of $f_k$ that $k$ divides $|A| - f_k(\psi)$.

The function $h^R_k$ is used in the translation of subformulae of the form $\varphi \text{ R } \psi$: if a set $A$ is to be used to translate the subformula $\varphi \text{ R } \psi$ at a symbolic $k$-path $\pi_n$ (with starting point $m$), then for every $i$ such that $m \leq i \leq k$, the set $h^R_k(A, f_k(\psi))(k)$ is used to translate the formula $\varphi$ at the symbolic paths $\pi_n$ with starting point $i$; moreover, for every $j$ such that $m \leq j \leq i$, the set $h^R_k(A, f_k(\varphi))(j)$ is used to translate the formula $\psi$ at the symbolic path $\pi_n$ with starting point $j$. Notice that if $k + 1$ does not divide $|A| - 1$, then $h^R_k(A, p)$ is undefined. However, for every set $A$ such that $|A| = f_k(\varphi \text{ R } \psi)$, it is clear from the definition of $f_k$ that $k + 1$ divides $|A| - f_k(\varphi)$.

Now we are ready to define our translation. Let $\alpha$ be an ECTL* state formula and $A \subseteq \mathbb{N}_+$ be a set of numbers of symbolic $k$-paths such that $|A| = f_k(\alpha)$, $\varphi$ be an ECTL* formula and $B \subseteq \mathbb{N}_+$ be a set of numbers of symbolic $k$-paths such that $|B| = f_k(\varphi)$, $n \in \mathbb{N}$, and $0 \leq m \leq k$. By $\langle \alpha \rangle_k^{[m,n,A]}$ we denote the translation of an ECTL* state formula $\alpha$ at the symbolic state $w_{m,n}$ by using the set $A$, and by $\langle \varphi \rangle_k^{[m,n,B]}$ we denote the translation of an ECTL* path formula $\varphi$ on the symbolic $k$-path $\pi_n$ with starting point $m$ by using the set $B$. 

---

A. Zbrzezny
Translation of ECTL\(^*\) formulae

\[
(\text{true})_{k}^{[m,n,A]} := \text{true} \\
(\text{false})_{k}^{[m,n,A]} := \text{false} \\
\langle p \rangle_{k}^{[m,n,A]} := P_{p}(w_{m,n}) \\
\langle \neg p \rangle_{k}^{[m,n,A]} := \neg P_{p}(w_{m,n}) \\
\langle \alpha \cdot \beta \rangle_{k}^{[m,n,A]} := (\langle \alpha \rangle_{k}^{[m,n,g_r(A,f_{k}(\alpha))]} \cdot \langle \beta \rangle_{k}^{[m,n,g_r(A,f_{k}(\beta))]}), \text{ where } \phi \in \{\wedge, \vee\} \\
\langle E \varphi \rangle_{k}^{[m,n,A]} := H(w_{m,n},w_{0,\min(A)}) \land \bigvee_{i=0}^{k} B_{i}^{\varphi}(u_{\min(A)}) \land \lbrack \varphi \rbrack_{k}^{[0,\min(A),g_r(A)]} \\
\langle [A] \rangle_{k}^{[m,n,A]} := \langle (\alpha)_{k}^{[m,n,A]} \rangle \\
\langle \alpha \cdot \gamma \rangle_{k}^{[m,n,A]} := [\varphi]_{k}^{[m,n,g_r(A,f_{k}(\alpha))]} \cdot [\psi]_{k}^{[m,n,g_r(A,f_{k}(\gamma))]}, \text{ where } \phi \in \{\wedge, \vee\} \\
(X \varphi)_{k}^{[m,n,A]} := \bigvee_{i=0}^{k-1} \left( L_{k}^{\varphi}(\pi_{n}) \land \lbrack \varphi \rbrack_{k}^{[i+1,n,A]} \right), \text{ if } m = k \\
\langle \varphi U \psi \rangle_{k}^{[m,n,A]} := \bigvee_{j=m}^{k} \left( [\psi]_{k}^{[j,n,h_{k}^{L}(A,f_{k}(\psi))(k)]} \land \bigwedge_{i=m}^{j-1} [\varphi]_{k}^{[i,n,h_{k}^{L}(A,f_{k}(\psi))(i)]} \right) \\
\vee \left( \bigvee_{i=0}^{j-1} \left( B_{i}^{\varphi}(u_{n}) \land \lbrack \varphi \rbrack_{k}^{[i,n,h_{k}^{L}(A,f_{k}(\psi))(i)]} \right) \right) \\
\land \left( \bigwedge_{i=m}^{k} [\varphi]_{k}^{[i,n,h_{k}^{L}(A,f_{k}(\psi))(i)]} \right) \\
\langle \varphi R \psi \rangle_{k}^{[m,n,A]} := \bigvee_{j=m}^{k} \left( [\psi]_{k}^{[j,n,h_{k}^{R}(A,f_{k}(\psi))(k)]} \land \bigwedge_{i=m}^{j-1} [\varphi]_{k}^{[i,n,h_{k}^{R}(A,f_{k}(\psi))(i)]} \right) \\
\vee \left( \bigvee_{i=0}^{j-1} \left( B_{i}^{\varphi}(u_{n}) \land \lbrack \varphi \rbrack_{k}^{[i,n,h_{k}^{R}(A,f_{k}(\psi))(i)]} \right) \right) \\
\land \left( \bigwedge_{i=m}^{k} [\varphi]_{k}^{[i,n,h_{k}^{R}(A,f_{k}(\psi))(i)]} \right) \\
\vee \left( \bigvee_{i=m}^{j-1} \left( B_{i}^{\varphi}(u_{n}) \land \lbrack \varphi \rbrack_{k}^{[i,n,h_{k}^{R}(A,f_{k}(\psi))(i)]} \right) \right) \\
\land \left( \bigwedge_{i=m}^{k} [\varphi]_{k}^{[i,n,h_{k}^{R}(A,f_{k}(\psi))(i)]} \right) \
\right)
\]
Note 3. From the definitions of the operators $F$ and $G$ it follows that

$$[F \psi]_{m,n,A}^k := \bigvee_{i=m}^k \left[ \psi \right]_{k}^{i,n,k_h^U(A,f_k(\psi))(k)}$$

$$\forall \left( L_{k}(\pi_n) \land \bigvee_{i=0}^{m-1} \left( B_{k}^{>}(u_n) \land \left[ \psi \right]_{k}^{i,n,k_h^U(A,f_k(\psi))(k)} \right) \right)$$

$$[G \psi]_{m,n,A}^k := \bigvee_{l=0}^{k-1} \left( L_{k}^l(\pi_n) \right) \land \bigwedge_{j=0}^{m-1} \left( B_{k}^{>}(u_n) \rightarrow \left[ \psi \right]_{k}^{j,n,k_h^R(A,f_k(\varphi))(j)} \right)$$

$$\land \bigwedge_{j=m}^k \left[ \psi \right]_{k}^{j,n,k_h^R(A,f_k(\varphi))(j)}$$

3.3 Bounded Model Checking of ECTL\(^{*}\) properties

First, define the unfolding of the transition relation $[M]^A_k$ for a finite set $A$ of symbolic $k$-paths, where $A = \{ j \in \mathbb{N} | 1 \leq j \leq n \}$, for some $n \in \mathbb{N}$.

$$[M]^A_k := \bigwedge_{j \in A} \bigvee_{i=0}^{k-1} T(w_{i,j}, w_{i+1,j})$$

Now let us recall how to verify by the BMC method a given ECTL\(^{*}\) state formula $\alpha$. For this one has to check the satisfiability of the following conjunction:

$$[M]^\alpha_k := \mathcal{I}(w_{0,0}) \land [M]^F_k(\alpha) \land \left( \alpha \right)_{k}^{0,0,0,F_k(\alpha)}$$

starting with $k = 0$. If for a given $k$ the formula $[M]^\alpha_k$ is not satisfiable, then $k$ is increased and the resulting formula is to be checked by a SAT-solver again. The method described relies on the following Theorem 2.

**Theorem 2.** Let $M$ be a model and $\alpha$ be an ECTL\(^{*}\) state formula. Then for every $k \in \mathbb{N}$, $M \models \alpha$ if, and only if, the propositional formula $[M]^\alpha_k$ is satisfiable.

The proof of Theorem 2 will be provided in the full version of the present paper.

4 Experimental results

In order to compare the new translation with the old one we have implemented both the old translation (as described in [11]) and the new one as the standalone programs written in the programming language C++. Our experiments were performed on a computer equipped with Intel Xeon 2 GHz processor, 4GB of RAM and the operating system Ubuntu Linux Server with the kernel 2.6.38. The both tools and the benchmarks can be found at the webpage http://ajd.czest.pl/~modelchecking/software.html, together with an instruction how to repeat the experiments.
As the benchmark we used the well-known dining philosophers problem ([5], [8]). We have modelled this problem by means of communicating finite automata. The system consists of \( n \) automata each of which models a philosopher, together with \( n \) automata each of which models a fork, together with one automaton which models the lackey. The latter automaton is used to coordinate the philosophers’ access to the dining-room. In fact, this automaton ensures that no deadlock is possible. The global system is obtained as the parallel composition of the components, which are shown in Figure 1.

\[
\begin{align*}
0 & \xrightarrow{\text{in}_j} 1 & \xrightarrow{\text{get}_j} 2 & \xrightarrow{\text{get}_{(j+1) \mod n}} 3 \\
3 & \xrightarrow{\text{put}_j} 4 & \xrightarrow{\text{put}_{(j+1) \mod n}} 3 \\
0 & \xrightarrow{\text{get}_j, \text{get}_{(j-1) \mod n}} 1 \\
0 & \xleftarrow{\text{put}_j, \text{put}_{(j-1) \mod n}} 1 \\
0 & \xleftarrow{\text{in}_1, \ldots, \text{in}_n} 1 & \ldots & n-2 & \xrightarrow{\text{in}_1, \ldots, \text{in}_n} n-1 \\
0 & \xrightarrow{\text{out}_1, \ldots, \text{out}_n} 1 & \ldots & 0
\end{align*}
\]

**Fig. 1.** The automata for the \( j \)-th Philosopher, the \( j \)-th Fork and the Lackey

Let \( AP = \{ p^i_j \mid 0 \leq j < n \land 0 \leq i \leq 5 \} \) and assume that the variable \( p^i_j \) is true only at the state \( s_j \) of the automaton for the philosopher of number \( i \). We have tested the following ECTL* state formula over \( AP \):

\[
\varphi_n = \bigwedge_{j=0}^{n-1} \text{EF} \left( p^2_j \land \text{E} \left( \text{F} p^3_j \land \text{G} \neg p^4_j \right) \right).
\]

in order to compare experimental results both for the old and new translations. This formula expresses the following property:

*For each philosopher there exists a path such that eventually this philosopher has got his left fork and from such a situation there exists a path on which eventually the philosopher has got his right fork and then he is eating forever.*

It turned out that the formula \( \varphi_n \) is valid in the considered model for every \( n \geq 4 \). Moreover, notice that \( LL = f_k(\varphi_n) = 2 \cdot n \).

Table 1 shows the experimental results of verification of the formula \( \varphi_n \) for different numbers of philosophers using the old and the new translation, respectively. The data in the table are organised as follows: the first column indicates which of the two translations applies to the given row, the second column shows the number of philosophers, the third the length of the witness, and the fourth the value of \( f_k(\varphi_n) \). Then, the fifth and sixth columns contains the time and memory consumed by BMC to generate
the set of clauses, while the next two columns give the time and memory consumed by the state of the art SAT solver MiniSAT [6]. The last two columns are self-explanatory.

<table>
<thead>
<tr>
<th>Translation</th>
<th>n</th>
<th>k</th>
<th>LL</th>
<th>sec</th>
<th>MB</th>
<th>sec</th>
<th>MB</th>
<th>sec</th>
<th>MB</th>
</tr>
</thead>
<tbody>
<tr>
<td>old</td>
<td>15</td>
<td>7</td>
<td>30</td>
<td>1073.4</td>
<td>206.1</td>
<td>22244.9</td>
<td>4596.0</td>
<td>23318.3</td>
<td>4596.0</td>
</tr>
<tr>
<td>new</td>
<td>15</td>
<td>7</td>
<td>30</td>
<td>11.3</td>
<td>43.1</td>
<td>224.0</td>
<td>65.3</td>
<td>224.0</td>
<td></td>
</tr>
<tr>
<td>new</td>
<td>35</td>
<td>7</td>
<td>70</td>
<td>126.7</td>
<td>340.5</td>
<td>3037.5</td>
<td>3256.0</td>
<td>3164.2</td>
<td></td>
</tr>
<tr>
<td>new</td>
<td>40</td>
<td>7</td>
<td>80</td>
<td>186.6</td>
<td>491.5</td>
<td>8329.7</td>
<td>5411.0</td>
<td>8516.3</td>
<td>5411.0</td>
</tr>
</tbody>
</table>

Table 1: The results of the old and the new translation of $\varphi_n$.

5 Conclusions and Future Work

We have shown that our new translation from ECTL* to SAT results in reducing the effort of the SAT solver while testing satisfiability of the formula to be verified. The experiments confirm that the improvement in question leads to a reduction of the size of the CNF formulae submitted to the SAT solver, and therefore to a significant reduction both in the time and memory required by the SAT solver to return an answer. Additional experiments are necessary to compare the effectiveness of our new translation with the standalone translation from ECTL to SAT described in [12] and with various translations for LTL.

References