Combining Symmetry Reduction and Under-Approximation for Symbolic Model Checking

Sharon Barner and Orna Grumberg
Computer Science Department
Technion - Israel Institute of Technology
Haifa 32000, Israel
{skeidar,orna}@cs.technion.ac.il
Contact author: Orna Grumberg, Phone number 972-4-8294327

Category A

Abstract

This work presents a collection of methods, integrating *symmetry reduction*, *under-approximation*, and *symbolic model checking* in order to reduce space and time for model checking. The main goal of this work is *falsification*. However, under certain conditions our methods provide *verification* as well.

We first present algorithms that perform on-the-fly model checking for temporal safety properties, using symmetry reduction. We then extend these algorithm for checking liveness properties as well.

Our methods are fully automatic. The user should supply some basic information about the symmetry in the verified system. However, the methods are *robust* and work correctly even if the information supplied by the user is incorrect. Moreover, the methods return correct results even in case the computation of the symmetry reduction has not been completed due to memory or time explosion.

We implemented our methods within IBM's model checker RuleBase, and compared the performance of our methods with that of RuleBase. In most cases, our algorithms outperformed RuleBase with respect to both time and space.

Keywords: symmetry reduction, under-approximation, symbolic model checking, hints

1 Introduction

This work presents a collection of methods, integrating *symmetry reduction*, *under-approximation*, and *symbolic model checking* in order to reduce space and time for model checking. The main goal of this work is *falsification*, that is, proving that a given system does not satisfy its specification. However, under certain conditions our methods provide also *verification*, i.e., they prove that the system satisfies its specification.

Our methods are fully automatic. The user should supply some basic information about the symmetry in the verified system. However, the methods are *robust* and work correctly even if the information supplied by the user is incorrect. Moreover, the methods return correct results even in case the computation of the symmetry reduction has not been completed due to memory or time explosion.

Temporal logic model checking [8] is a technique that accepts a finite state model of a system and a temporal logic specification and determines whether the system satisfies the specification. The main problem of model checking is its high memory requirements. *Symbolic model checking* [18], based on BDDs [5], can handle larger systems, but is still limited in its capacity. Thus, additional work is needed in order to make model checking feasible for larger systems.

This work exploits symmetry reduction in order to reduce the memory and time used by symbolic model checking. Symmetry reduction is based on the observation that many systems consist of several similar components. Exchanging the role of such components in the system does not change the system behavior. Thus, system states can be partitioned into equivalence classes called *orbits*, and the system can be verified by examining only representative states from each orbit.

Two main problems arise, however, when combining symbolic model checking with symmetry reduction. One is building the orbit relation and the other is choosing a representative for each orbit. [16] proves that the BDD for the orbit relation is exponential in the number of the BDD variables, and suggests choosing more than one representative for each orbit in order to obtain a smaller BDD for the orbit relation. Yet, this method does not solve the problem of choosing the representatives. The choice of representatives is significant since it strongly influences the size of the BDDs representing the symmetry-reduced model. [12] suggests to choose generic representatives. This approach involves compiling the symmetric program to a reduced model over the generic states, and such a compilation can only be applied to programs written with a special syntax in which symmetry is defined inside the program. [15] introduces an algorithm for explicit model checking which chooses as a representative for an orbit the first state from this orbit, discovered by the DFS. This method avoids choosing the representative in advance. Unfortunately, it is not applicable to symbolic model checking since performing DFS is very inefficient with BDDs.

We suggest a new approach that avoids building the orbit relation and chooses representatives on-the-fly while computing the reachable states. Unlike [15] the choice of the representatives is guided by BDD criteria. Reachability is performed using *under-approximation* which, at each step, explores only a subset of the reachable states. Some of the unexplored states are symmetric to the explored ones. By exploiting symmetry information, those states will never be explored. Thus, easier symbolic forward steps are obtained.

We first apply this approach for verifying properties of the form $AG(p)$[1], where p is a boolean formula. If we find a "bad" state that does not satisfy p, then we conclude that the checked system does not satisfy $AG(p)$. On the other hand, if no "bad" state is found we cannot conclude that the system satisfies $AG(p)$ since reachability with under-approximation does not necessarily explore every reachable state. We next present a special version of the previous algorithm in which the under-approximation is guided by *hints* [3]. Under certain conditions this algorithm can also verify the system.

The algorithms described above are based on reachability, and are often referred to as *on-the-fly* model checking. It is well known how to extend on-the-fly model checking for $AG(p)$ to verifying general *safety temporal properties.* This is done by building an automaton describing the property and running it together with the system. We specified conditions on the automaton that guarantee the correctness of the on-the-fly algorithm also when the automaton runs together with the symmetry-reduced model. The suggested conditions hold for the tableau construction used for symbolic LTL model checking [6], when restricted to LTL safety properties. They also hold for the satellite used in symbolic model checking of RCTL formulas [2]. By running the automaton together with the reduced model we save both space and time while verifying these type of formulas.

On-the-fly model checking cannot handle liveness properties. In order to handle such properties we developed two extensions combining symmetry reduction with classical (not on-the-fly) symbolic model checking. One is easy to perform and is mainly suitable for falsification. The other is more expensive but can handle verification as well.

Previous works expect the user to provide a symmetry group that is also an invariance group [16]. In many cases two formulas checked on the same model require different invariance groups since each formula breaks differently the symmetry of the model. Thus, the user needs to supply different invariance group for different formulas. In other works [20, 9] the program is written in a special syntax which enables finding the invariance group according to this syntax. In these cases only formulas which do not break the symmetry of the model are allowed.

In contrast, we build the invariance group automatically, once the symmetry group is given. Supplying the symmetry group usually requires only a high level understanding of the system and therefore is easier than supplying the invariance group.

We implemented our methods within the enhanced model checking tool RuleBase [1], developed by the IBM Haifa Research Laboratories, and compared the performance of our methods with that of RuleBase. Our experiments show that our methods performed significantly better, with respect to both time and space, in checking liveness properties. For temporal safety properties they achieved better time requirements. However, their space requirements were worse for small examples and identical for larger ones.

The rest of the paper is organized as follows. Section 2 gives some basic definitions. Section 3 shows how to build the invariance group. Section 4 presents an algorithm for on-the-fly symbolic model checking with symmetry reduction and then introduces hints into this algorithm. Section 5 and 6 handle temporal safety properties and liveness properties, respectively and Section 7 presents our experimental results. We conclude

[1] $AG(p)$ means that p holds along every path, in every state on the path.

in Section 8 with directions for future research.

2 Preliminaries

Let AP be a set of atomic propositions. We model a system by a Kripke structure M over AP, $M = (S, S_0, R, L)$ where S is a finite set of states, S_0 is a set of initial states, $R \subseteq S \times S$ is a total transition relation, and $L : S \rightarrow 2^{AP}$ is a labeling function which labels each state with the set of atomic propositions true in that state.

As the specification language we use the branching time temporal logic CTL, defined over AP. The semantics of CTL is defined with respect to a Kripke structure. We write $M \models \varphi$ to denote that the formula φ is true in M. For a formal definition of CTL and its semantics see [7].

ACTL is the sub-logic of CTL in which all formulas contain only universal path quantifiers. A CTL formula is *boolean* if it contains only atomic propositions and boolean operators. A formula β is a *maximal boolean sub-formula* of a CTL formula φ if β is a boolean sub-formula of φ and for all sub-formulas β' of φ, if β is a sub-formula of β' then β' is not boolean.

The *bisimulation equivalence* and *simulation preorder* are relations over Kripke structures (see [7] for definitions) that have useful logical characterizations. We write $M \equiv_{bis} M'$ to denote that M and M' are bisimulation equivalent and $M \leq_{sim} M'$ to denote that M is smaller than M' by the simulation preorder. The following lemmas relate bisimulation and simulation with logics.

Lemma 2.1 *[4] For every CTL formula φ over atomic propositions AP and two Kripke structures M, M' over AP, if $M \equiv_{bis} M'$ then $M' \models \varphi \Leftrightarrow M \models \varphi$.*

Lemma 2.2 *[14] For every ACTL formula φ over AP and two Kripke structures M, M' over AP, if $M \leq_{sim} M'$ then $M' \models \varphi \Rightarrow M \models \varphi$.*

2.1 BDDs

A Binary Decision Diagram (BDD) [5] is a data structure for representing boolean functions. BDDs are defined over boolean variables, they are often (but not always) concise in their memory requirement, and most boolean operations can be performed efficiently on BDD representations. In [18] it was shown that BDDs can be very useful for representing Kripke structures and performing model checking symbolically. One of the most useful operations in model checking and in particular on-the-fly model checking is the *image computation*. Given a set of states S and a relation T, represented by the BDDs $S(\bar{v})$ and $T(\bar{v}, \bar{v}')$ respectively, the image computation finds the set of all states related by T to some state in S.

Definition 2.1 $Im_T(S(\bar{v})) = \exists \bar{v}(S(\bar{v}) \wedge T(\bar{v}, \bar{v}'))$

2.2 Partial search

While symbolic model checking can be very efficient, it can still suffer from explosion in the BDD size. One of the solutions is to perform partial search of the reachable state space while avoiding large BDDs [19]. Other methods perform partial search which is guided by the user [3] or by the checked specification [21]. In all methods the set of reachable states discovered in each step is an under approximation of the set of reachable states which would have been discovered in BFS. This property enables combining partial search with on-the-fly model checking.

2.3 The product model and the restricted model

We now define two special Kripke structures that will be used later.

Definition 2.2 *Let M, M' be two Kripke structures defined over the sets of atomic propositions AP and AP', respectively. The* product structure *of M and M' is a Kripke structure over $AP \cup AP'$ defined as follows. $M \times M' = (S_{M\times M'}, S^0_{M\times M'}, R_{M\times M'}, L_{M\times M'})$ where*

- $S_{M\times M'} = \{\ (s,s') \mid s \in S \wedge\ s' \in S' \wedge L(s) \cap AP' = L'(s') \cap AP\}$.
- $S^0_{M\times M'} = \{\ (s,s') \mid (s,s') \in S_{M\times M'}\ \wedge\ s \in S_0\ \wedge\ s' \in S'_0\ \}$.
- $\forall (s,s'),(t,t') \in S_{M\times M'}:\quad ((s,s'),(t,t')) \in R_{M\times M'} \Leftrightarrow (s,t) \in R \wedge (s',t') \in R'$.
- $\forall (s,s') \in S_{M\times M'}:\ L((s,s')) = L(s) \cup L'(s')$.

Definition 2.3 *Let M be a Kripke structure and A be a subset of S. The* restricted model *$M|_A = (S|_A, S_0|_A, R|_A, L|_A)$ is defined as follows:*

- $S|_A\ = A$.
- $S_0|_A = S_0 \cap A$.
- $\forall s,s' \in S|_A\ [(s,s') \in R|_A \Leftrightarrow (s,s') \in R]$.
- $\forall s \in S|_A\ [L|_A(s) = L(s)]$.

Lemma 2.3 *For every Kripke structure M and $A \subseteq S$, $\ M|_A \leq_{sim} M$.*

2.4 Symmetry

A permutation on a set A, $\sigma : A \to A$ is a one-to-one and onto function. For a set $A' \subseteq A$, $\sigma(A') = \{a | \exists a' \in A'\ \sigma(a') = a\}$. In this paper we use permutations over the set of states of a Kripke structure. Given a CTL formula β and a structure M, $\sigma(\beta)$ refers to applying σ to the set of states in M that satisfy β.

A permutation group G is a set of permutations together with the functional composition such that the identity permutation e is in G, for every permutation $\sigma \in G$, there is a permutation called σ^{-1} such that $\sigma\sigma^{-1}$ = e and for every $\sigma_1, \sigma_2 \in G$, $\sigma_1\sigma_2 \in G$. If there exists $\sigma \in G$ such that $\sigma(s) = s'$ we say that two states s,s' are symmetric.

Definition 2.4 *$\sigma_1, \sigma_2, \ldots, \sigma_k$ are* generators *of permutation group G (denoted $G = \langle \sigma_1, \sigma_2, \ldots, \sigma_k \rangle$) if G is the closure of the set $\{\sigma_1, \sigma_2, \ldots, \sigma_k\}$ under functional composition.*

Definition 2.5 *A permutation group G is a* symmetry group *of a Kripke structure M if every permutation $\sigma \in G$ preserves the transition relation. That is, $\forall s,s' \in S\ [(s,s') \in R \Leftrightarrow (\sigma(s),\sigma(s')) \in R]$.*

Definition 2.6 *A symmetry group G of a Kripke structure M is an* invariance group *for formula φ if for every maximal boolean sub-formula β of φ, every $\sigma \in G$ and $s \in S$ $[M,s \models \beta \Leftrightarrow M,\sigma(s) \models \beta]$.*

Given an invariance group G and a Kripke structure M we can partition S into equivalence classes. The equivalence class of s is $[s] = \{s' | \exists \sigma \in G,\ \sigma(s) = s'\}$. Each [s] is called an *orbit* and the relation OR = $\{(s,s') | s,s' \in S\ and\ [s] = [s']\}$ is called the *orbit relation*

For a Kripke structure $M = (S, S_0 R, L)$ and an invariance group G for φ the *quotient structure* $M_G = (S_G, S^0_G, R_G, L_G)$ is defined as follows:

- $S_G = \{\ [s] \mid s \in S\ \}$

- $S_G^0 = \{\ [s] \mid s \in S_0\ \}$
- $R_G = \{\ ([s], [s']) \mid (s, s') \in R\ \}$
- $L_G([s]) = L(s)$

In [11, 16] it has been proved that $M_G \equiv_{bis} M$. By Lemma 2.1 we therefore have that for every CTL formula ψ over the same AP as φ, $M_G \models \psi \Leftrightarrow M \models \psi$.

In order to build the quotient structure a representative is chosen from each orbit and a *representative function* $\xi : S \rightarrow Rep$, which maps each state s to its orbit representative, is defined. M_G is now defined by $S_G =$ Rep and $R_G = \{\ (s, s') \mid s, s' \in Rep \wedge \exists s'' \in S\ [R(s, s'') \wedge \xi(s'') = s']\}$.

In many cases, however, it is easier to choose more than one representative for each orbit. ξ is then a *representative relation* $\xi \subseteq Rep \times S$ which satisfies $(s, s') \in \xi \Leftrightarrow s \in Rep \wedge [s] = [s']$. In this case we define the structure $M_m = (S_m, S_m^0 R_m, L_m)$ (m for multiple representatives) where $S_m = Rep$, $S_m^0 = Rep \cap S_0$, $R_m = \xi^{-1}R$ and $L_m = L$. [16] shows that $M_m \equiv_{bis} M_G \equiv_{bis} M$.

3 Building the invariance group

In this section we show how to automatically compute the generators of an invariance group given the generators of a symmetry group.

Our method works as follows. Given a set of generators for a symmetry group G, an invariance group G_{inv} is defined by restricting the generators of G to those σ_i that satisfy $\sigma_i(\beta) = \beta$ for every $\beta \in AP$. As a result, the orbits of G_{inv} obtained by its generators, do not contain both states that satisfy β and states that do not. This implies that G_{inv} is an invariance group. The following lemma states the correctness of our approach.

Lemma 3.1 *Let $\sigma_1, \sigma_2, \ldots, \sigma_k$ be generators of a symmetry group G of a Kripke structure M and let φ be a formula over maximal atomic formulas in AP. Then IG = $\{\sigma_i | \forall\beta \in AP,\ \sigma_i(\beta) = \beta\}$ generates an invariance group G_{inv} of M for φ.*

The invariance group G_{inv} generated by IG may not be the largest invariance group for M. The largest invariance group can be obtained by restricting the largest symmetry group G as follows: $G_{inv} = \{\ \sigma \in G \mid \forall\beta \in AP,\ \sigma(\beta) = \beta\}$. However, the number of permutations in G may be exponential in the number of generators of G. Thus it is not practical to construct G_{inv} directly from G.

4 Symmetry with on-the-fly representatives

The symbolic algorithm **Symmetry_MC** presented in this section is aimed at avoiding the two main problems of symmetry reduction, namely building the orbit relation and choosing a representative for each orbit.

Let $M = (S, S_0, R, L)$ be a Kripke structure and $\sigma_1, \ldots, \sigma_k$ be a set of generators of a symmetry group G of M. Also let $\varphi = AG(p)$ where p is a boolean formula. The algorithm **Symmetry_MC**, presented in Figure 1, applies on-the-fly model checking for M and φ, using under-approximation and symmetry reduction.

The algorithm works in iterations. Starting with the set of initial states, at each iteration a subset `under` of the current set of states is chosen. The successors of `under` are computed. However, every state which is symmetric to (i.e., in the same orbit with) a previously reached state is removed. The states that are first found for each orbit are taken to be the orbit representatives. Note that an orbit may have more than one representative if several of its states are reached when the orbit is encountered for the first time. At any step, the set of representatives are checked to see if they include a state that violets p. If such a state is found (line 9) then the computation is stopped and a counterexample is produced. We then conclude that $M \not\models AG(p)$.

A useful optimization can be performed by deleting from memory the BDD for the set **`full_reach`** immediately after is has been used (after line 7). This may avoid memory explosion, since the BDD for **`full_reach`** is often quite large. In addition, the forward step in line 5 usually requires large memory utilization. By removing the BDD for **`full_reach`** before computing the forward step we decrease the memory usage in each iteration.

```
    **Symmetry_MC**(M, $\varphi$, $\sigma_1, \ldots \sigma_k$)
1       Calculate the generators of the invariance group of M
        IG = $\{\sigma_i |$ for all maximal boolean sub-formula $\beta$ of $\varphi$:  $\sigma_i(\beta) = \beta\}$
2       reach_rep = $S_0$, i=0
3       while $S_i \neq \emptyset$
4           choose under $\subseteq S_i$ (under is an under-approximation of $S_i$)
5           $S_{i+1} = Im_R$(under)
6           full_reach = $\sigma$_Step(reach_rep)
7           $S_{i+1} = S_{i+1}$ / full_reach
8           reach_rep = reach_rep $\cup S_{i+1}$
9           if $S_{i+1} \wedge \neg p \neq \emptyset$
10              generate a counter example and break.
11          i = i+1.
```

Figure 1: The algorithm **Symmetry_MC** performs on-the-fly model checking of φ on M, using symmetry reduction

The set of symmetric states that should be removed are computed using the procedure σ**_Step** (Figure 2) instead of using the orbit relation. σ**_Step** applies Im_{σ_i}(`sym_states`)[2]

```
    $\sigma$_Step(A, $\sigma_1, \sigma_2, \ldots, \sigma_k$)
        sym_states = A;
        old_sym_states = $\emptyset$
        while old_sym_states $\neq$ sym_states
            old_sym_states = sym_states
            for $i = 1 \ldots k$
                new_sym_states = $Im_{\sigma_i}$(sym_states)
                sym_states = sym_states $\cup$ new_sym_states
        return sym_states
```

Figure 2: The algorithm σ**_Steps** calculates the states belonging to the orbits of states in A

in order to obtain the states which are related by σ_i to states in **`sym_states`**. It repeatedly computes Im_{σ_i} for $i = 1, \ldots, k$, until a fixed point is reached. For a set of states A and a set of generators $IG = \{\ \sigma_1, \ldots, \sigma_k\ \}$, σ**_Step** returns the set of all states belonging to the orbits of states in A according to $G = \langle IG \rangle$.

[2] σ_i can be viewed as the binary relation $\bar{v} = \sigma(\bar{v}')$

By using **σ_Step** we exploit symmetry information without building the orbit relation. There are several reasons to expect that **σ_Step** will result in a BDD which is smaller than that of the orbit relation. First, it represents a set of states and not a relation. Thus, it depends on one set of BDD variables, while the orbit relation depends on two sets. Second, it is applied only to reachable states which are usually represented by smaller BDDs. Indeed, our experiments successfully applied **σ_Step** to designs for which building the orbit relation was impossible.

Computationally, **σ_Step** is quite heavy. To avoid this problem, in most of our experiments we stopped the computation of **σ_Step** before it got to a fixed point. In general there is a tradeoff between the amount of computation in **σ_Step** and the symmetry reduction obtained by **Symmetry_MC**.

4.1 Robustness of Symmetry_MC

We now discuss the robustness of the algorithm **Symmetry_MC** for falsification in the presence of an incomplete **σ_Step** and an incorrect set of generators. Consider first the case in which the computation of procedure **σ_Step** is stopped before a fixed point is reached. **σ_Step** then returns only a subset of the states in the orbits of states in A. In this case, less states are removed from S_{i+1} and as a result `reach_rep` contains more states. Thus, we might have more representatives for each orbit.

Consider now the case in which the algorithm is given an incorrect set of generators. If a "bad" generator (a permutation which associates states that are not symmetric) is given, then **σ_Step** returns states which are not symmetric to any state in `reach_rep`. These states are removed from S_{i+1} and we might not add any representatives of their orbits to `reach_rep`. Thus, `reach_rep` represents an under-approximation of the reachable orbits. However, `reach_rep` does not contain a representative of an unreachable orbit. Thus, if there is a state $s \in$ `reach_rep` which does not satisfy p, this state is reachable in the original model and the counterexample generated by **Symmetry_MC** actually exists in the original model.

If a "good" generator (a permutation which associates pairs of symmetric states) is missing, then **σ_Step** returns less states and as a result there is more than one representative for each orbit. However, like in the previous case, `reach_rep` contains only reachable states and therefore **Symmetry_MC** generates only real counterexamples. The following lemma summarizes the discussion above.

Lemma 4.1 *Given any set of generators, the algorithm* ***Symmetry_MC*** *is sound for falsification.*

At termination of **Symmetry_MC**, if `reach_rep` contains at least one state from each reachable orbit then M_m, defined according to `reach_rep` ($S_m =$ `reach_rep`) is bisimilar to M (see Section 2.4). Thus, if $M_m \models AG(p)$ then $M \models AG(p)$ as well. Note that $M_m \models AG(p)$ can be checked on-the-fly.

Several BDD optimizations may be useful for procedure **σ_Step**. One is to simplify σ_i according to `sym_states` in each iteration. Another is to apply partitioned transition relation with early quantification in the computation of Im_{σ_i}. This is applicable since often a permutation is given as a conjunction of simpler expressions.

4.2 Symmetry reduction combined with hints

In this section we present a special case of the algorithm **Symmetry_MC** in which the under-approximation is guided by a sequence of hints given by the user [3].

Let M = (S,S$_0$,R,L) be a Kripke structure, $\sigma_1, \ldots, \sigma_k$ be a set of generators of a symmetry group G on M, and $h_1, \ldots, h_l$ be a sequence of hints where h_l = TRUE. Also, let $\varphi = AG(p)$ be a formula where p is a boolean formula. The algorithm **Hints_Sym**, presented in Figure 3, applies on-the-fly model checking for M and φ using hints and symmetry reduction.

If $\sigma_1, \ldots, \sigma_k$ contain no "bad" generator[3] then our hints guarantee that when $S_i = \emptyset$, `reach_rep` contains at least one state from each orbit. In this case, the algorithm **Hints_Sym** is suitable for verification as well as falsification.

A useful optimization is to compute the set `full_reach` only once for each hint and to use it in order to remove states in all steps in which this hint is used. This save computation time but may use more space since `full_reach` is in memory when Im_R is computed. Again there is a tradeoff here between time and space.

```
Hints_Sym(M, $\varphi$, $\sigma_1, \sigma_2, \ldots, \sigma_k$, $h_1, h_2, \ldots, h_l$)
    Calculate IG = $\{\sigma_i |$ for all maximal boolean sub-formula $\beta$ of $\varphi$: $\sigma_i(\beta) = \beta\}$
    reach_rep = $S_0$, i = 0, hint = $h_1$, j = 2
    while $S_i \neq \emptyset$
        under = $S_i \cap$ hint
        $S_{i+1} = Im_R$(under)
        full_reach = $\sigma$_Step(reach_rep)
        $S_{i+1} = S_{i+1}$/full_reach
        reach_rep = reach_rep $\cup S_{i+1}$
        if $S_{i+1} \wedge \neg p \neq \emptyset$
            generate counter example and break
        if $S_{i+1} = \emptyset \wedge j \leq l$
            hint = $hint_j$
            j = j+1
            $S_{i+1}$ = reach_rep
        i = i+1
    $\varphi$ is TRUE
```

Figure 3: The algorithm **Hints_Sym** applies on-the-fly model checking of φ on M, using hints and symmetry reduction

5 Extension for temporal safety properties

There are several known algorithms which use a construction A_φ for the evaluated formula φ and the product model $M \times A_\varphi$ in order to do model checking more efficiently. We now show that it is possible to combine symmetry reduction with these algorithms. We first specify the requirements on the construction A_φ so that it can be used with symmetry reduction.

Definition 5.1 *Given a logic $\mathcal{L}$ and a construction that associates with each $\varphi \in \mathcal{L}$ a structure A_φ, the construction A_φ is safe for symmetry reduction w.r.t $\mathcal{L}$ if it satisfies the following conditions:*

[3] In many cases, the nonexistence of bad generators can be easily determined by the program syntax. In other cases it is expensive but possible to check whether all generators are good.

1. $\exists\psi\forall\varphi \in \mathcal{L}$ $(M \models \varphi \Leftrightarrow M \times A_\varphi \models \psi)$.

2. For every invariance group G_{inv} of M for φ, every $\sigma \in G_{inv}$ and every $(s,t) \in S_{M\times A_\varphi}$, $\sigma((s,t)) = (\sigma(s),t)$ [4].

3. For every maximal boolean formula β of ψ and every $(s,t),(s',t) \in S_{M\times A_\varphi}$, $(s,t) \models \beta \Leftrightarrow (s',t) \models \beta$.

The second condition requires that σ is defined only on s and leaves t unchanged. The third condition requires that the truth of all β in ψ depend only on t.

Lemma 5.1 *For every construction A_φ which is safe for symmetry reduction w.r.t $\mathcal{L}$, if G is an invariance group of structure M for formula $\varphi \in \mathcal{L}$ then G is an invariance group of structure $M \times A_\varphi$ for formula ψ.*

Corollary 5.2 *For every construction A_φ which is safe for symmetry reduction w.r.t $\mathcal{L}$ and for every $\varphi \in \mathcal{L}$ and $\psi \in CTL$, the quotient structure $(M \times A_\varphi)_G$, built for $M \times A_\varphi$ and an invariance group G of M, satisfies $(M \times A_\varphi)_G \models \psi \Leftrightarrow M \models \varphi$.*

Note that using safe construction enables us to find the generators of the invariance group of M according to φ and then to evaluate formula ψ on $M \times A_\varphi$ with symmetry reduction that use the same generators. There are several A_φ constructions which are safe for symmetry reduction w.r.t logic $\mathcal{L}$. One example is the tableau construction in [6] when restricted to LTL safety properties. In this case the tableau includes no fairness constraints and it fulfills the requirements of Definition 5.1. Another safe construction is the satellite for RCTL formulas defined in [2]. By combining safe construction with symmetry reduction we make symmetry reduction applicable with a new set of algorithms, like symbolic on-the fly model checking and symbolic LTL model checking, for which it was not applicable until now. We implemented our algorithms using the construction introduced in [2], which enabled us to check RCTL formulas on-the-fly while using a symmetry reduction.

6 Extensions for liveness formulas

We now describe two possible extensions that combine classical (not on-the-fly) symbolic model checking with symmetry reduction. These extensions are useful for checking liveness properties, and other properties which cannot be checked on-the-fly.

6.1 Liveness restricted to representatives

The purpose of this extension is to falsify ACTL formulas with respect to a structure M, while avoiding the construction of its quotient model M_G. The idea is to get a set of representatives Rep and to construct the restricted model $M|_{Rep}$ (see Definition 2.3). Since $M|_{Rep} \leq_{sim} M$, we have that for every ACTL formula φ, if $M|_{Rep} \not\models \varphi$ then $M \not\models \varphi$. Thus, φ can be checked on the smaller model $M|_{Rep}$.

Note that in principle this idea works correctly with any set of representatives, even such that does not include a representative for each orbit. There are however advantages to choosing as Rep the set `reach_rep` which results from the algorithm **Symmetry_MC**. First, since `reach_rep` includes only reachable states, its BDD is usually smaller than the BDD of an arbitrarily chosen set of states. Second, by construction, the states in `reach_rep` are connected by transitions while an arbitrary set

[4]since s and $\sigma(s)$ agree on AP, $[(s,t) \in S_{M\times A_\varphi} \Leftrightarrow (\sigma(s),t) \in S_{M\times A_\varphi}]$.

of representatives *Rep* might not be connected. Thus, $M|_{reach\_rep}$ often includes more behaviors than $M|_{Rep}$. As a result, it is more likely that a bad behavior, if exists, will be found in $M|_{reach\_rep}$. Third, the states in `reach_rep` represent many other states in the system, thus if the system includes a bad behavior, it is more likely that `reach_rep` will reflect it.

Following the discussion above we suggest the Algorithm **Live_Rep** that works as follows: it first runs **Symmetry_MC** to obtain `reach_rep` and then performs classical symbolic model checking on $M|_{reach\_rep}$.

6.2 Liveness with the representative relation

We now present another possibility for handling liveness properties. It is applicable only if no bad generators exist. This method is more expensive computationally, but is suitable for verification of liveness properties. Similarly to the previous section we first compute `reach_rep` using the algorithm **Symmetry_MC**. However, now we apply the procedure **Create_ξ**, presented below, in order to compute the representative relation $\xi \subseteq$ `reach_rep` $\times S$ (see definition in Section 2.4). Next we construct a new structure $M' = (S', S'_0, R', L')$ where $S' =$ `reach_rep`, $S'_0 = S_0 \cap S'$, $R' = \xi^{-1}R$ and $L' = L$. Finally, we run classical symbolic model checking on φ and M'.
Lemma 6.1 *If S' contains at least one representative for each orbit then $M \equiv_{bis} M'$. Otherwise, $M' \leq_{sim} M$.*
If `reach_rep` is the result of the algorithm **Hints_Sym**, then `reach_rep` indeed contains at least one representative for each orbit, and M' is bisimilar to M. Thus, M' can be used for verifying full CTL.

Figure 4 presents the BDD-based procedure **Create_ξ** for building the representative relation ξ for a given set of representatives Rep and a set of generators $\sigma_1, \ldots, \sigma_k$ of an invariance group G of M for φ.

Suppose that each σ_i is represented by a BDD $\sigma_i(\bar{v}, \bar{v}')$ and Rep is represented by a BDD Rep($\bar{v}$). **Create_ξ** returns the BDD $\xi(\bar{v}, \bar{v}')$ for the relation $\xi \subseteq Rep \times S$. Line

```
   Create_ξ($\sigma_1, \sigma_2 \ldots \sigma_k$, Rep)
1      $\xi(\bar{v}, \bar{v}') = \text{Rep}(\bar{v}) \wedge (\bar{v} = \bar{v}')$
2      old_ξ$(\bar{v}, \bar{v}') = \phi$
3      while old_ξ$(\bar{v}, \bar{v}') \neq \xi(\bar{v}, \bar{v}')$
4          old_ξ$(\bar{v}, \bar{v}') = \xi(\bar{v}, \bar{v}')$
5          for i = 1 ... k
6              new$(\bar{v}, \bar{v}'') = \exists \bar{v}'(\xi(\bar{v}, \bar{v}') \wedge \sigma_i(\bar{v}'', \bar{v}'))$
7              $\xi(\bar{v}, \bar{v}') = \xi(\bar{v}, \bar{v}') \cup$ new$(\bar{v}, \bar{v}')$
8      return $\xi(\bar{v}, \bar{v}')$
```

Figure 4: The algorithm **Create_ξ** for computing $\xi \subseteq Rep \times S$

6 is implemented with the operator `compose_odd` [17] which computes $\exists \bar{v}'(\xi(\bar{v}, \bar{v}') \wedge \sigma_i(\bar{v}'', \bar{v}'))$ using only two sets of BDD variables instead of three.

7 Experimental results

We implemented the algorithms **Hints_Sym**, **Live_Rep**, and **Create_ξ** in the IBM's model checker RuleBase [1]. We ran it on a number of examples which contain symmetry. For each example we tuned our algorithms according to the evaluated formula, the difficulty level of computing the reachable states and the difficulty level of building the transition relation. In most cases, our algorithms outperformed RuleBase with respect to both time and space. In the tables below time is measured by seconds, memory (mem) in bytes, and the transition relation size (TR size) in number of BDD nodes.

The Futurebus example: We ran the algorithm Live_Rep in order to check liveness properties on the Futurebus cache-coherence protocol with a single bus and a single cache line for each processor. We checked the property "along every path infinitely often some processor is in exclusive write". This property fails since our model does not include fairness constraints. The table in Figure 5 presents the results of evaluating the property for a different number of processors. For comparison we ran also the RuleBase classical symbolic model checking algorithm. Both algorithms applied dynamic BDD reordering. The BDD order is very important since the best BDD order for the classical algorithm is different from the best BDD order of our algorithm. In order to obtain a fair comparison between these algorithms we ran each algorithm twice. In the first run the algorithm reordered the BDD without time limit in order to find a good BDD order. The initial order of the second run was the BDD order which was found by the first run.

The most difficult step in the Futurebus example is building the transition relation. By restricting the transition relation to the representatives which were chosen on-the-fly, the transition relation became smaller and as a result the evaluation became easier. In this case we chose to complete the calculation of σ**_Step** in order to obtain the maximal reduction in the size of the transition relation. Figure 5 shows that both time and space were reduced dramatically using **Live_Rep**. We can also observe that the larger the number the processors was the better the results were. This is to be expected, as the increase in the number of the reachable representatives is smaller than the increase in the number of reachable states.

#	#	classic algorithm			Live_Rep		
of processors	vars	time	mem	TR size	time	mem	TR size
5	45	132	43M	144069	101	41M	122769
6	54	607	118M	260625	265	56M	219572
7	63	2852	277M	418701	704	76M	379428
8	72	8470	589M	839055	3313	101M	457781
9	81	81,171	709M	1935394	4571	106M	819871
10	90	-	> 1G	-	4909	120M	642083

Figure 5: **Hints_Sym** on Future bus example

The Arbiter example: We ran algorithm **Hints_Sym** on an arbiter example with n processes. We checked the arbiter w.r.t. RCTL formulas which were translated to safe A_φ and ψ. For comparison we ran RuleBase on-the-fly model checking and on-the-fly model checking with hints (without symmetry). All algorithms used dynamic BDD reordering and partitioned transition relation [10]. In this case we calculated σ**_Step** only when we changed hints and stopped σ**_Step** before the fixed point has been reached. The table in Figure 6 presents the results of the three algorithms on arbiter with 6,8 and 10 processes. For each case we checked one property that passed and one that failed. We notice that **Hints_Sym** reduced time but not necessarily space. This can be explained by the fact that σ**_Step** produced large intermediate BDDs but resulted in a significant reduction in S_i, thus reduced the computation time of the image steps.

# of processors	status	# vars	on-the-fly		on-the-fly + hints		Hints_Sym	
			time	mem	time	mem	time	mem
6	passed	65	53	40M	39	40M	42	40M
6	failed	65	213	52M	64	41M	51	87M
8	passed	84	581	64M	255	49M	179	87M
8	failed	84	745	71M	524	71M	292	83M
10	passed	105	1470	94M	598	67M	358	92M
10	failed	105	1106	93M	740	73M	520	91M

Figure 6: **Hints_Sym** compared to other on-the-fly algorithms

Comparing Create_ξ and Orbit_To_ξ: [16] presents an algorithm for computing ξ by building the orbit relation and then choosing the representatives. We refer to this algorithm by **Orbit_To_ξ**. We compare this algorithm with **Create_ξ**. Both algorithms find the representative relation $\xi \subseteq Rep \times S$ for the set of representatives Rep chosen according to the lexicographic order. The results in Figure 7 show that **Create_ξ** gave better results in both time and space, We believe that this is due to the fact that it saves less information while building ξ.

num of generators	num of vars	orbit_to_ξ		Create_ξ	
		time	mem	time	mem
3	16	0.26	26M	0.23	26M
4	20	30.4	33M	1.2	28M
5	24	1017	114M	18	42M
6	28	-	>1.5G	735	132M
5	32	-	>1.5G	29083	1.2G

Figure 7: **Create_ξ** compared to **Orbit_To_ξ**

8 Directions for future research

Our liveness algorithms can easily be extend to models with fairness constraints. This will be done by requiring that the generators of the invariance group satisfy $\sigma(\beta) = \beta$ for maximal boolean formulas in the fairness constraints as well as in the checked formula.

It is possible to use the algorithm **Symmetry_MC** with other partial search algorithms which choose the next set of states to be explored according to criteria different than hints, such as High-Density Reachability presented in [19]. σ**_Step** can be used in other search algorithms like Prioritized Traversal [13]. There, the BDDs in the priority queue can be reduced by eliminating states which are in the orbits of states that were already explored.

Acknowledgment: We thank Cindy Eisner for many helpful discussions. Somesh Jha is thanked for his help with the examples.

References

[1] I. Beer, S. Ben-David, C. Eisner, and A. Landver. RuleBase: An industry-oriented formal verification tool. In *Design Automation Conference*, pages 655–660, June 1996.

[2] I. Beer, S. Ben-David, and A. Landver. On-the-fly model checking of RCTL formulas. In A. J. Hu and M. Y. Vardi, editors, *Proceedings of the 10th International conference on Computer-Aided Verification*, volume 1427 of *LNCS*, pages 184–194. Springer-Verlag, June 1998.

[3] R. Bloem, K. Ravi, and F. Somenzi. Symbolic guided search for CTL model checking. In *Design Automation Conference*, pages 29–34, June 2000.

[4] M. Browne, E. Clarke, and O. Grumberg. Characterizing finite Kripke structures in propositional temporal logic. *Theoretical Comput. Sci.*, 59:115–131, 1988.

[5] R. E. Bryant. Graph-based algorithms for boolean function manipulation. *IEEE transactions on Computers*, C-35(8):677–691, 1986.

[6] E. Clarke, O. Grumberg, and H. Hamaguchi. Another look at LTL model checking. *Formal Methods in System Design*, 10(1), 1997.

[7] E. Clarke, O. Grumberg, and D. Peled. *Model Checking*. MIT press, December 1999.

[8] E. M. Clarke, E. A. Emerson, and A. P. Sistla. Automatic verification of finite-state concurrent systems using temporal logic specifications. *ACM Trans. Prog. Lang. Syst.*, 8(2):244–263, 1986.

[9] C.N. Ip and D.L. Dill. Better verification through symmetry. In D. Agnew, L. Claesen, and R. Camposano, editors, *Computer Hardware Description Languages and their Applications*, pages 87–100, Ottawa, Canada, 1993. Elsevier Science Publishers B.V., Amsterdam, Netherland.

[10] D. Geist and I. Beer. Efficient model checking by automated ordering of transition relation. In David L. Dill, editor, *Proceedings of the sixth International Conference on Computer-Aided Verification CAV*, volume 818, pages 299–310. Springer-Verlag, June 1994.

[11] E. A. Emerson and A. P. Sistla. Symmetry and model checking. In C. Courcoubetis, editor, *Proceedings of the 5th International conference on Computer-Aided Verification*, volume 697 of *LNCS*. Springer-Verlag, June 1993.

[12] E. A. Emerson and R. J. Trefler. From asymmetry to full symmetry: New techniques for symmetry reduction in model checking. In *Conference on Correct Hardware Design and Verification Methods*, pages 142–156, 1999.

[13] R. Fraer, G. Kamhi, B. Ziv, M. Y. Vardi, and L. Fix. Prioritized traversal: Efficient reachability analysis for verification and falsification. In E. A. Emerson and A. P. Sistla, editors, *Proceedings of the 12th International conference on Computer-Aided Verification*, volume 1855 of *LNCS*, pages 389–402. Springer-Verlag, July 2000.

[14] O. Grumberg and D. Long. Model checking and modular verification. *ACM Trans. on Programming Languages and Systems*, 16(3):843–871, 1994.

[15] V. Gyuris and A. P. Sistla. On-the-fly model checking under fairness that exploits symmetry. *Formal Methods in System Design: An International Journal*, 15(3):217–238, November 1999.

[16] S. Jha. *Symmetry and Induction in Model Checking*. PhD thesis, CMU, 1996.

[17] S. Katz. Coverage of model checking. Master's thesis, Technion, haifa, Israel, 2001.

[18] K. L. McMillan. *Symbolic Model Checking: An Approach to the State Explosion Problem*. Kluwer Academic Publishers, 1993.

[19] K. Ravi and F. Somenzi. High-density reachability analysis. In *Proc. Intl. Conf. on Computer-Aided Design*, pages 154–158, November 1995.

[20] A. P. Sistla, V. Gyuris, and E. A. Emerson. SMC: a symmetry-based model checker for verification of safety and liveness properties. *Software Engineering and Methodology*, 9(2):133–166, 2000.

[21] C. H. Yang and D. L. Dill. Validation with guided search of the state space. In *Design Automation Conference*, pages 599–604, June 1998.