Online Enumeration of All Minimal Inductive Validity Cores

Jaroslav Bendík[1], Elaheh Ghassabani[2], Michael Whalen[2], and Ivana Černá[1]

[1] Faculty of Informatics, Masaryk University, Brno, Czech Republic
{xbendik,cerna}@fi.muni.cz
[2] Department of Computer Science & Engineering, University of Minnesota, MN, USA
{ghass013,mwwhalen}@umn.edu

Abstract. Symbolic model checkers can construct proofs of safety properties over complex models, but when a proof succeeds, the results do not generally provide much insight to the user. Minimal Inductive Validity Cores (MIVCs) trace a property to a minimal set of model elements necessary for constructing a proof, and can help to explain why a property is true of a model. In addition, the traceability information provided by MIVCs can be used to perform a variety of engineering analysis such as coverage analysis, robustness analysis, and vacuity detection. The more MIVCs are identified, the more precise analyses can be performed. However, a full enumeration of all MIVCs is in general intractable due to the large number of possible model element sets. The bottleneck of existing algorithms is that they are not guaranteed to emit minimal IVCs until the end of the computation, so returned results are not known to be minimal until all solutions are produced.
In this paper, we propose an algorithm that identifies MIVCs in an *online* manner (i.e., one by one) and can be terminated at any time. We benchmark our new algorithm against existing algorithms on a variety of examples, and demonstrate that our algorithm not only is better in intractable cases but also completes the enumeration of MIVCs faster than competing algorithms in many tractable cases.

Keywords: Inductive Validity Cores, SMT-based model checking, Inductive proofs, Traceability, Proof cores

1 Introduction

Symbolic model checking using induction-based techniques such as IC3/PDR [4], k-induction [20], and k-liveness [3] can be used to determine whether properties hold of complex finite or infinite-state systems. Such tools are popular both because they are highly automated (often requiring no user interaction other than the specification of the model and desired properties), and also because, in the event of a violation, the tool provides a counterexample demonstrating a situation in which the property fails to hold. These counterexamples can be used both to illustrate subtle errors in complex hardware and software designs [17, 15, 16] and to support automated test case generation [22, 23].

If a property is proved, however, most model checking tools do not provide additional information. This can lead to situations in which developers have an unwarranted level of confidence in the behavior of the system. Issues such as vacuity [13], incorrect environmental assumptions [21], and errors either in English language requirements or formalization [19] can all lead to failures of "proved" systems. Thus, even if proofs are established, one must approach verification with skepticism.

Recently, *proof cores* [1] have been proposed as a mechanism to determine which elements of a model are used when constructing a proof. This idea is formalized by Ghassabani et al. for inductive model checkers [7] as *Inductive Validity Cores* (IVCs). IVCs offer proof explanation as to why a property is satisfied by a model in a formal and human-understandable way. The idea lifts UNSAT cores [24] to the level of sequential model checking algorithms using induction. Informally, if a model is viewed as a conjunction of constraints, a minimal IVC (MIVC) is a set of constraints that is sufficient to construct a proof such that if any constraint is removed, the property is no longer valid. Depending on the model and property to be analyzed, there are many possible MIVCs, and there is often substantial diversity between the IVCs used for proof.

In previous work [7, 18, 9, 8] we have explored several different uses of IVCs, including:

Traceability: For functional properties that can be proven with inductive model checkers, inductive validity cores can provide accurate traceability matrices with no user effort. Given multiple IVCs, *rich traceability* matrices [18] can be automatically constructed that provide additional insight about *required* vs. *optional* design elements.

Vacuity detection: The idea of syntactic vacuity detection (checking whether all subformulae within a property are necessary for its validity) has been well studied [13]. IVCs allow a generalized notion of vacuity that can indicate weak or mis-specified properties even when a property is syntactically non-vacuous.

Coverage analysis: Closely related to vacuity detection is the idea of *coverage analysis*, e.g., are all atoms in the model necessary for at least one of the properties proven about the model? Several different notions of coverage have been proposed [2, 12], but these tend to be very expensive to compute.

Impact Analysis: Given a single (or for more accurate results, all) MIVCs, it is possible to determine which requirements may be falsified by changes to the model. This analysis allows for selective regression verification of tests and proofs: if there are alternate proof paths that do not require the modified portions of the model, then the requirement does not need to be re-verified.

Design Optimization: Synthesis tools can benefit from MIVCs in the process of transforming an abstract behavior into a design implementation. A practical way of calculating all MIVCs allows synthesizers to find a minimum set of design elements (optimal implementation) for a certain behavior. Such optimizations can be performed at different levels of synthesis.

To be useful for these tasks, the generation process must be efficient and the generated IVC must be accurate and precise (that is, sound and minimal). In previous work, we have developed an efficient *offline* algorithm [8] for finding all minimal IVCs based on the MARCO algorithm for MUSes [14]. The algorithm is considered *offline* because it is not until all IVCs have been computed that one knows whether the solutions computed are, in fact, minimal. In cases in which models contain many IVCs, this approach can be impractically expensive or simply not terminate.

TODO: Jaroslav and Ivana: Make sure I'm speaking correctly here TODO: Elaheh: fill in experimental data TODO: ALL: talk about 'unknown' results when timeout occurs. TODO: ALL: do we need a running example? I'm not sure. If we do an illustration of it, the Brno team might be best at this.

In this paper, we consider three *online* algorithms for MIVC enumeration. With these algorithms, solutions are produced at a regular rate, and each solution produced is guaranteed to be minimal. Additionally, for models with a large number of IVCs, the proposed algorithms are considerably more efficient than the baseline MARCO algorithm. We demonstrate this via experiment, where the new

The rest of the paper is organized as follows...

2 Motivating Example

3 Preliminaries

A transition system (I, T) over a state space S consists of an initial state predicate $I : S \to \mathit{bool}$ and a transition step predicate $T : S \times S \to \mathit{bool}$. The notion of reachability for (I, T) is defined as the smallest predicate $R : S \to \mathit{bool}$ satisfying the following formulae:

$$\forall s \in S : \; I(s) \Rightarrow R(s)$$
$$\forall s, s' \in S : \; R(s) \wedge T(s, s') \Rightarrow R(s')$$

A safety property $P : S \to \mathit{bool}$ holds on a transition system (I, T) iff it holds on all reachable states, i.e., $\forall s \in S : \; R(s) \Rightarrow P(s)$. We denote this by $(I, T) \vdash P$. We assume the transiton step predicate T is equivalent to a conjunction of transition step predicates $T_1, \ldots, T_n$, called top level conjuncts. In such case, T can be identified with the set of its top level conjuncts $\{T_1, \ldots, T_n\}$. By further abuse of notation, we write $T \setminus \{T_i\}$ to denote removal of top level conjunct T_i from T, and $T \cup \{T_j\}$ to denote addition of top level conjunct T_j to T.

Definition 1. *A set of conjuncts $U \subseteq T$ is an* Inductive Validity Core (IVC) *for $(I, T) \vdash P$ iff $(I, U) \vdash P$. Moreover, U is a* Minimal IVC (MIVC) *for $(I, T) \vdash P$ iff $(I, U) \vdash P$ and $\forall T_i \in U : \; (I, U \setminus \{T_i\}) \nvdash P$.*

Note, that the minimality is with respect to the set inclusion and not wrt cardinality. There can be multiple MIVCs with different cardinalities.

TODO: Example

4 Existing Techniques

Let us first recall a naive enumeration algorithm that explicitly checks each subset of T for being an IVC and then finds the minimal IVCs using subset inclusion relation. The main disadvantage of this approach is the exponential number of subsets of T. We briefly describe existing techniques that can be used to find all MIVCs while checking only a a small portion of subsets of T for being IVCs. Most of the techniques were inspired by the MUS enumeration techniques [] proposed in the area of constraint processing and applied by Ghassabani et al. [].

Definition 2 (Inadequacy). *A set of conjuncts $U \subseteq T$ is an* inadequate *set for $(I, T) \vdash P$ iff $(I, U) \nvdash P$. Especially, $U \subseteq T$ is a* Maximal Inadequate Set (MIS) *for $(I, T) \vdash P$ iff U is inadequate and $\forall T_i \in (T \setminus U) : (I, U \cup \{T_i\}) \vdash P$.*

Inadequate sets are duals to inductive validity cores. Each $U \subseteq T$ is either inadequate set or an inductive validity core. In order to unify the notation, we use notation *inadequate* and *adequate*. Note that especially minimal inductive validity cores can be thus called minimal adequate sets.

The first property used to improve the naive enumeration algorithm is the *monotonicity* of adequacy with respect to the subset inclusion.

Lemma 1 (Monotonicity). *If a set of conjuncts $U \subseteq T$ is an adequate set for $(I, T) \vdash P$ than all its supersets are adequate for $(I, T) \vdash P$ as well:*

$$\forall U_1 \subseteq U_2 \subseteq T : (I, U_1) \vdash P \Rightarrow (I, U_2) \vdash P.$$

Symmetrically, if $U \subseteq T$ is an inadequate set for $(I, T) \vdash P$ than all its subsets are inadequate for $(I, T) \vdash P$ as well:

$$\forall U_1 \subseteq U_2 \subseteq T : (I, U_2) \nvdash P \Rightarrow (I, U_1) \nvdash P.$$

Proof. If $U_1 \subseteq U_2$ then reachable states of (I, U_2) form a subset of the reachable states of (I, U_1).

Monotonicity allows to determine status of multiple subsets of T while using only a single check for adequacy. For example, if a set $U \subseteq T$ is determined to be adequate, than all of its supersets are adequate and do not need to be explicitly checked. Let $\mathit{Sup}(U)$ and $\mathit{Sub}(U)$ denote the set of all supersets and subsets of U, respectively.

Every algorithm for computing MIVCs has to determine status (i.e adeqaute or inadequate) of every subset of T. In order to distinguish the subsets whose status is already known from those whose status is not known yet, we denote the former subsets as *explored* subsets and the latter as *unexplored* subsets. Moreover, we distinguish *maximal* resp. *minimal* unexplored subsets

- U_{max} is a *maximal unexplored subset* of T iff $U_{max} \subseteq T$, U_{max} is unexplored, and each of its proper supersets is explored.
- U_{min} is a *minimal unexplored subset* of T iff $U_{min} \subseteq T$, U_{max} is unexplored, and each of its proper subsets is explored.

Algorithm 1: Shrinking procedure

input : $(I, U) \vdash P$
output: MIVC for $(I, U) \vdash P$
1 **for** $T_i \in U$ **do**
2   **if** $(I, U \setminus \{T_i\}) \vdash P$ **then** $U \leftarrow U \setminus \{T_i\}$
3 **return** U

A straightforward way to find a (so far unexplored) MIVC of T is to find an unexplored adequate subset $U \subseteq T$ and turn U into an MIVC by a process called *shrinking*. Shrinking procedure iteratively attempts to remove elements from the set that is being shrunk, checking each new set for adequacy and keeping only changes that leave the set adequate (see Algorithm 1 for pseudocode). To find an unexplored adequate subset Ghassabani et al. [8] always choose maximal unexplored subset and test it for adequacy. The shrinking procedure itself can be extremely time demanding as each check for adequacy is in fact a model checking problem. Ghassabani et al. [] therefore substitute the shrinking procedure with an appproximative procedure which returns IVC which is not necessarilly minimal and identify MIVCs (using subset relation) at the end of the computation when status of all subsets is determined. Their experimental evaluations show that the approximative shrinking is much faster. However, it does not enumerate MIVCs online.

5 Algorithm

In this section, we propose a novel algorithm for online enumeration of all MIVCs. The algorithm is built out of two basic procedures: shrink and grow. The grow procedure is symmetric to the shrink one.

In our algorithm we maintain the sets *Explored* and *Unexplored*. Every time a set U of conjuncts is determined as adequate, the set U as well as all its supersets are moved to the set *Explored* as due to the monotonicity they are all adequate. Symetrically for an inadeaquate set of conjunncts and the set *Unexplored*.

5.1 Shrink Procedure

We can effectively use the set *Explored* for speeding up the shrinking procedure. When testing the set $U \setminus \{T_i\}$ (see line 2 in Algorithm 1) we first check whether $U \setminus \{T_i\}$ is explored. If so, the status of $U \setminus \{T_i\}$ is known and no test for adequacy is needed.

However, there is one more observation that can be exploited.

Observation 1. *Let U_1, U_2 be subsets of T such that U_1 is explored, U_2 is unexplored, and $U_1 \subset U_2$. Then U_1 is inadequate for $(I, T) \vdash P$.*
Symetrically, if U_1, U_2 are subsets of T such that U_2 is explored, U_1 is unexplored, and $U_1 \subset U_2$. Then U_2 is adequate for $(I, T) \vdash P$.

Algorithm 2: Approximate grow

input : $(I,T) \vdash P$
input : inadequate $U \subset T$ for $(I,T) \vdash P$
input : set *Unexplored* of unexplored subsets of T
output: approximately maximal inadequate set for $(I,T) \vdash P$

$M \leftarrow$ a maximal $M \in Unexplored$ such that $M \supseteq U$
while $(I,M) \vdash P$ **do**
  $M_{IVC} \leftarrow$ `IVC_UC`$((I,M),P)$ `// gets approximately minimal IVC`
  $T_i \leftarrow$ choose $T_i \in (M_{IVC} \setminus U)$
  $M \leftarrow M \setminus \{T_i\}$
return M

Proof. If U_1 is adequate, then all of its supersets are necessarily adequate. Thus, if U_1 is determined to be adequate, then not just U_1 but also all of its supersets becomes explored. Since U_1 is explored and U_2 is unexplored, then U_1 is necessarily an inadequate subset of T.

In other words, we are quaranteed that whenever during the shrinking procedure we come across an explored set, this set is inadequate. Therefore as a further optimization in our algorithm we try to identify as many inadequate sets as possible before starting the shrinking procedure. The search for inadequate sets is done with the help of grow procedure.

5.2 Grow Procedure

Recall that if a set is determined to be inadequate then all of its subsets are necessarily also inadequate. Therefore, the larger set is determined to be inadequate, the more inadequate sets become explored. To identify inadequate sets as quickly as possible we search for maximal inadequate sets (MISes).

In order to find a MIS, we can find an inadequate set $U \subset T$ and use a process called *grow* which turns U to a MIS for $(I,T) \vdash P$. Grow procedure iteratively attempts to add elements from $T \setminus U$ to U, checking each new set for adequacy and keeping only changes that leave the set inadequate. Same as in the case of shrink procedure, we can use the set *Explored* to avoid checking sets whose status is already known. However, such grow procedure might still perform too many checks for adequacy and thus be very inefficient.

Instead, we propose to use a different approach. Algorithm 2 shows a procedure that, given an inadequate set U for $(I,T) \vdash P$, finds an *approximately* maximal inadequate set. It first finds some maximal unexplored set M such that $M \supseteq U$ and checks it for adequacy. If M is inadequate, then it is necessarily a MIS (this is a straightforward consequence of Observation 1.). Otherwise, if M is adequate then it is iteratively reduced until an inadequate set is found. In particular, whenever M is found to be adequate, the approximative procedure `IVC_UC` by Ghassabani et al. [**?**] is used to find an approximate MIVC M_{IVC} of

M which succinctly explains M's adequacy. In order to turn M into an inadequate set, it is reduced by one element from $M_{IVC} \setminus U$ and checked for adequacy. If M is still adequate then the approximate growing procedure continues with a next iteration. Otherwise, if M is inadequate, the procedure finishes.

Proposition 1. *Given an unexplored inadequate set U for $(I, T) \vdash P$ and a set of unexplored subsets of T, Algorithm 2 returns an* unexplored *inadequate subset M of T.*

Proof. Let us denote initial M as M_{init}. Since $M_{init} \supseteq U$ and M is recursively reduced only by elements that are not contained in U, then in every iteration holds that $U \subseteq M \subseteq M_{init}$. Since both U, M_{init} are unexplored, then M is necessarily also unexplored.

5.3 Complete Algorithm

In this section, we describe, how to combine the shrink and grow methods in order to form an efficient online MIVC enumeration algorithm.

Since knowledge of (approximately) maximal inadequate subsets might be used to speed up shrinking procedures, it might be tempting to first find all maximal inadequate subsets. However, there can be up to exponentially many such subsets with respect to the size of T. Thus, finding first all maximal inadequate subsets is in general intractable. Instead, we propose to alternate both shrinking and growing procedures. Note, that during shrinking, we might determine some subsets to be inadequate, and such subsets can be subsequently used as *seeds* for growing. Dually, we might determine some subsets to be adequate during the growing procedures, and such subsets can be used as *seeds* for shrinking procedures. Thus, both these procedures somehow complements each other.

Algorithm **??** shows our algorithm for online enumeration of all MIVCs. It iteratively... todo: description of the algorithm.

TODO: correctness and complexity

6 Symbolic Representation of Unexplored Subsets

Symbolic representation is based on a well known isomorphism between finite power sets and Boolean algebras. We encode $T = \{T_1, T_2, \ldots, T_n\}$ by using a set of Boolean variables $X = \{x_1, x_2, \ldots, x_n\}$. Each valuation of X then corresponds to a subset of T. This allows us to represent the set of unexplored subsets *Unexplored* using a Boolean formula $f_{Unexplored}$ such that each model of $f_{Unexplored}$ corresponds to an element of *Unexplored*.

Our algorithm uses two types of operations to manage *Unexplored*: it removes from *Unexplored* either supersets of some adequate $U \subseteq T$ or subsets of some inadequate $U \subseteq T$. To remove $U \subseteq T$ and all its supersets from *Unexplored* we add to $f_{Unexplored}$ one clause,

$$f_{Unexplored} \leftarrow f_{Unexplored} \ \wedge \bigvee_{i:T_i \in U} \neg x_i \ .$$

Algorithm 3: AllMIVC algorithm

Function `FindMIVCs`$((I,T) \vdash P)$:
 $Unex \leftarrow \mathcal{P}(T)$
 $shrinkingQueue \leftarrow$ empty queue
 while $Unex \neq \emptyset$ **do**
  $B \leftarrow$ a minimal $B \in Unex$
  if $(I,B) \vdash P$ **then**
   output B
   $Unex \leftarrow Unex \setminus Sup(B)$
  else
   $MIS \leftarrow$ `Grow`(B) `// side effect: fills` $shrinkingQueue$
   $Unex \leftarrow Unex \setminus Sub(MIS)$
  while $shrinkingQueue$ *is not empty* **do**
   $seed$`Dequeue`$(shrinkingQueue)$
   $MIVC \leftarrow$ `Shrink`$(seed)$
   remove from $shrinkingQueue$ each V such that $V \supset MIVC$
   output $MIVC$
   $Unex \leftarrow Unex \setminus Sup(MIVC)$

Function `Shrink`(U):
 for $T_i \in U$ **do**
  if $U \setminus \{T_i\} \in Unex$ **then**
   if $(I, U \setminus \{T_i\}) \vdash P$ **then**
    $U \leftarrow T \setminus T_i$
 return U

Function `Grow`(B):
 $M \leftarrow$ a maximal $M \in Unex$ such that $M \supseteq B$
 while $(I,M) \vdash P$ **do**
  $MIVCApprox \leftarrow$ `IVC_UC`$((I,M),P)$
  $T_i \leftarrow chooseT_i \in (MIVCApprox \cap (T \setminus M))$
  $M \leftarrow M \setminus \{T_i\}$
 return M

Symmetrically, to remove $U \subseteq T$ and all its subsets from *Unexplored* we add to $f_{Unexplored}$ one clause,

$$f_{Unexplored} \leftarrow f_{Unexplored} \ \wedge \bigvee_{i:T_i \notin U} x_i \ .$$

Example 1. Let us illustrate the symbolic representation on $T = \{T_1, T_2, T_3\}$. If all subsets of T are unexplored then $f_{Unexplored} = True$. If $\{T_1, T_3\}$ is classified as an MIVC and $\{T_1, T_2\}$ as a inadequate set, then $f_{Unexplored}$ is updated to $True \wedge (\neg x_1 \vee \neg x_3) \wedge (x_3)$.

In order to get an element of *Unexplored*, we can ask a SAT solver for a model of $f_{Unexplored}$. However, our algorithm requires specific elements of *Unexplored*: maximal unexplored subsets (these correspond to maximal models), minimal unexplored subsets (these correspond to minimal models), and elements that are supersets of some particular set. One of the SAT solvers that can be used to obtain models corresponding to these specific elements is the solver miniSAT[5]. In miniSAT, the user can fix values of chosen variables and select a default polarity of variables at decision points during solving. For example, in order to find a maximal unexplored superset of $U \subseteq T$, we set the default polarity of variables during solving to *True* and fix the truth assignment to the variables that correspond to elements in U to *True*.

7 Implementation

We have implemented the algorithm in an industrial model checker called `JKind` [6], which verifies safety properties of infinite-state synchronous systems. It accepts Lustre programs [11] as input. The translation of Lustre into a symbolic transition system in `JKind` is straightforward and is similar to what is described in [10]. Verification is supported by multiple "proof engines" that execute in parallel, including K-induction, property directed reachability (PDR), and lemma generation engines that attempt to prove multiple properties in parallel. To implement the engines, `JKind` emits SMT problems using the theories of linear integer and real arithmetic. `JKind` supports the `Z3`, `Yices`, `MathSAT`, `SMTInterpol`, and `CVC4` SMT solvers as back-ends. When a property is proved and IVC generation is enabled, an additional parallel engine executes the `IVC_UC` algorithm [7] to generate an (approximately) minimal IVC. To implement our method, we have extended `JKind` with a new engine that implements Algorithm 3 on top of `Z3`. We use the `JKind` IVC generation engine to implement the `IVC_UC` procedure in Algorithm 3.

TODO: ALL: how do timeouts fit into the general framework?

8 Experiment

9 Conclusion

References

1. Cadence JasperGold Formal Verification Platform. https://www.cadence.com/.
2. H. Chockler, O. Kupferman, and M. Vardi. Coverage metrics for formal verification. *Correct hardware design and verification methods*, pages 111–125, 2003.
3. K. Claessen and N. Srensson. A liveness checking algorithm that counts. In *FMCAD*, pages 52–59. IEEE, 2012.
4. N. Een et al. Efficient implementation of property directed reachability. In *FMCAD'11*, 2011.

5. N. Eén and N. Sörensson. An extensible sat-solver. In E. Giunchiglia and A. Tacchella, editors, *Theory and Applications of Satisfiability Testing*, pages 502–518, Berlin, Heidelberg, 2004. Springer Berlin Heidelberg.
6. A. Gacek, J. Backes, M. Whalen, L. Wagner, and E. Ghassabani. The jkind model checker. *arXiv preprint arXiv:1712.01222*, 2017.
7. E. Ghassabani et al. Efficient generation of inductive validity cores for safety properties. In *FSE'16*, 2016.
8. E. Ghassabani, A. Gacek, and M. W. Whalen. Efficient generation of all minimal inductive validity cores. In *FMCAD2017: International Conference on Formal Methods in Computer-Aided Design*, 2017.
9. E. Ghassabani, A. Gacek, M. W. Whalen, M. Heimdahl, and W. Lucas. Proof-based coverage metrics for formal verification. In *ASE2017: 32nd IEEE/ACM International Conference on Automated Software Engineering*, 2017.
10. G. Hagen and C. Tinelli. Scaling up the formal verification of lustre programs with smt-based techniques. In *FMCAD'08*, 2008.
11. N. Halbwachs et al. The synchronous dataflow programming language Lustre. *Proceedings of the IEEE*, 1991.
12. O. Kupferman, W. Li, and S. Seshia. A theory of mutations with applications to vacuity, coverage, and fault tolerance. In *Proceedings of the 2008 Int'l Conf. on Formal Methods in Computer-Aided Design*, page 25, 2008.
13. O. Kupferman and M. Y. Vardi. Vacuity detection in temporal model checking. *STTT*, 2003.
14. M. Liffiton et al. Fast, flexible MUS enumeration. *Constraints*, 2016.
15. K. L. McMillan. A methodology for hardware verification using compositional model checking. Technical Report 1999-01, Cadence Berkeley Labs, Berkeley, CA 94704, 1999.
16. S. P. Miller, M. W. Whalen, and D. D. Cofer. Software model checking takes off. *Commun. ACM*, 53(2):58–64, 2010.
17. A. Murugesan et al. Compositional verification of a medical device system. In *HILT'13*, 2013.
18. A. Murugesan et al. Complete traceability for requirements in satisfaction arguments. In *RE'16 (RE@Next! Track)*, 2016.
19. L. Pike. A note on inconsistent axioms in rushby's "systematic formal verification for fault-tolerant time-triggered algorithms". *TSE*, 2006.
20. M. Sheeran et al. Checking safety properties using induction and a SAT-solver. In *FMCAD'02*, 2000.
21. M. Whalen et al. Integration of formal analysis into a model-based software development process. In *FMICS*, 2007.
22. M. Whalen, G. Gay, D. You, M. Heimdahl, and M. Staats. Observable modified condition/decision coverage. In *Proceedings of the 2013 Int'l Conf. on Software Engineering*. ACM, May 2013.
23. D. You, S. Rayadurgam, M. Whalen, and M. Heimdahl. Efficient observability-based test generation by dynamic symbolic execution. In *26th International Symposium on Software Reliability Engineering (ISSRE 2015)*, November 2015.
24. L. Zhang and S. Malik. Extracting small unsatisfiable cores from unsatisfiable boolean formula. In *SAT'03*.