

Domain-Type-Guided Refinement Selection Based on Sliced Path Prefixes

Dirk Beyer, Stefan Löwe, and Philipp Wendler

University of Passau, Germany



Technical Report, Number MIP-1501
Department of Computer Science and Mathematics
University of Passau, Germany
January 2015

arXiv:1502.00045v1 [cs.SE] 31 Jan 2015

Domain-Type-Guided Refinement Selection Based on Sliced Path Prefixes

Dirk Beyer, Stefan Löwe, and Philipp Wendler

University of Passau, Germany

Abstract—Abstraction is a successful technique in software verification, and interpolation on infeasible error paths is a successful approach to automatically detect the right level of abstraction in counterexample-guided abstraction refinement. Because the interpolants have a significant influence on the quality of the abstraction, and thus, the effectiveness of the verification, an algorithm for deriving the best possible interpolants is desirable. We present an analysis-independent technique that makes it possible to extract several alternative sequences of interpolants from one given infeasible error path, if there are several reasons for infeasibility in the error path. We take as input the given infeasible error path and apply a slicing technique to obtain a set of error paths that are more abstract than the original error path but still infeasible, each for a different reason. The (more abstract) constraints of the new paths can be passed to a standard interpolation engine, in order to obtain a set of interpolant sequences, one for each new path. The analysis can then choose from this set of interpolant sequences and select the most appropriate, instead of being bound to the single interpolant sequence that the interpolation engine would normally return. For example, we can select based on domain types of variables in the interpolants, prefer to avoid loop counters, or compare with templates for potential loop invariants, and thus control what kind of information occurs in the abstraction of the program. We implemented the new algorithm in the open-source verification framework CPACHECKER and show that our proof-technique-independent approach yields a significant improvement of the effectiveness and efficiency of the verification process.

I. INTRODUCTION

In the field of automatic software verification, abstraction is a well-understood and widely-used technique, enabling the successful verification of real-world, industrial programs (cf. [4, 12, 13]). Abstraction makes it possible to omit certain aspects of the concrete semantics that are not necessary to prove or disprove the program’s correctness. This may lead to a massive reduction of a program’s state space, such that verification becomes feasible within reasonable time and resource limits. For example, SLAM [5] uses predicate abstraction [17] for creating an abstract model of the software. One of the current research directions is to invent techniques to automatically find suitable abstractions. An ideal model is abstract enough to avoid state-space explosion and still contains enough detail to verify the property. Counterexample-guided abstraction refinement (CEGAR) [14] is an automatic technique that starts with a very coarse abstraction and iteratively refines an abstract model using *infeasible* error paths (witnesses of property violations). If the analysis does not find an error path in the abstract model, the analysis terminates and reports the

verdict TRUE (property holds). Because the abstract model over-approximates the program, the verdict applies for the actual program. If the analysis finds an error path, the path is checked for feasibility. If the found error path does not contain a contradiction and the error is indeed reachable according to the concrete program semantics, then the error path is feasible and a real error was found. The analysis terminates and reports the verdict FALSE (program violates property). If, however, the error path is actually infeasible, then a “spurious counterexample” was found, and the property violation is due to a too coarse abstract model. The (contradicting) constraints of the infeasible error path can then be passed to an interpolation engine, and the obtained interpolants identify information that is needed for refining the current abstraction, such that the same infeasible error path is excluded in subsequent CEGAR iterations. After refinement, the analysis proceeds with rebuilding a refined abstract model in the next CEGAR iteration. Several successful software verifiers (e.g., SLAM [5], BLAST [7], CPACHECKER [9], UFO [1]) make use of the CEGAR loop, which is illustrated in Figure 1.

Craig interpolation [15] is a technique that yields for two contradicting formulas an interpolant formula that contains less information than the first formula, but is still expressive enough to contradict the second formula. This can be extended to a sequence of formulas. In software verification, interpolation was first used for the domain of predicate abstraction [18], and later for value-analysis domains [11]. Independent of the analysis domain, interpolants for path constraints of infeasible error paths can be used to refine abstract models and to eliminate the infeasible error paths in subsequent CEGAR iterations. In this context, it is important to point out that the choice of interpolants is crucial for the performance of the analysis. Figure 2 gives an example: In this program, the analysis will typically find the shown error path, which is infeasible for

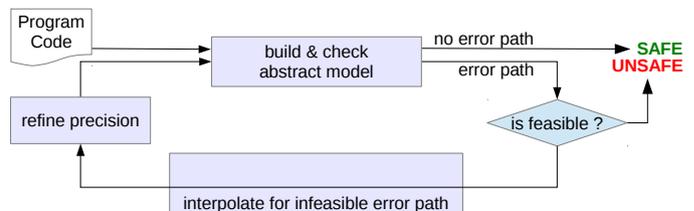


Fig. 1: Example of the CEGAR loop, using a single error path for interpolation

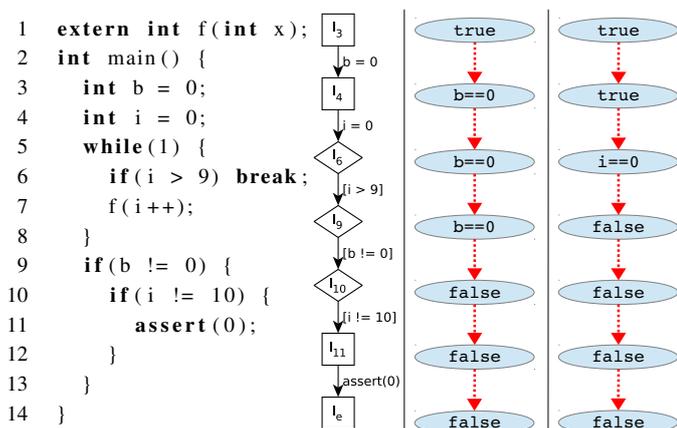


Fig. 2: Input program, with infeasible error path, and a “good” and a “bad” interpolant sequence

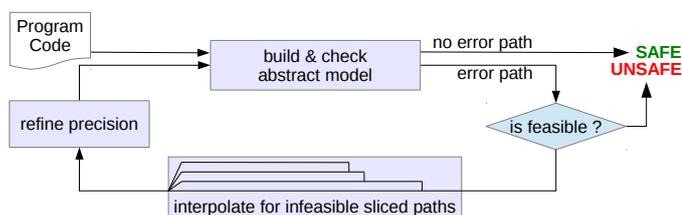


Fig. 3: Example of the CEGAR loop, using a set of paths for interpolation

two different reasons: both the value of i and the value of b can be used to find a contradiction. In general, it is now beneficial for the verifier to track the value of the boolean variable b , and not to track the value of the loop-counter variable i , because the latter has many more possible values, and tracking it would usually lead to an expensive unrolling of the loop. Instead, if only variable b is tracked, the verifier can conclude the safety of the program without unrolling the loop. Thus, we would like to get from the interpolation engine the left shown interpolant sequence (only with boolean variable) and not the right interpolant sequence (with loop-counter). However, interpolation engines typically do not allow to guide the interpolation process towards a “good”, or away from a “bad”, interpolant sequence. The interpolation engines inherently cannot do a better job here: they do not have access to information such as whether a specific variable is a loop counter and should be avoided in the interpolant. Instead, which interpolant is returned depends solely on the internal algorithms of the interpolation engine. This is especially true if the model checker in use does not provide its own implementation of an interpolation engine but rather makes calls to a library, e.g., a Satisfiability Modulo Theories (SMT) solver, which normally cannot be controlled on such a fine-grained level. In this case, the model checker is stuck to what the interpolation engine returns, be it good or bad for the verification process.

Therefore, we present an approach to *guide* the interpolation engine to produce interpolants that we would like to get, without

changing the interpolation engine. For this, we extract from an infeasible error path a set of infeasible sliced paths stemming all from the same infeasible error path. Each of these sliced paths can be used for interpolation, yielding different interpolant sequences that are all expressive enough to eliminate the original infeasible error path. As depicted in Figure 3, our approach fits well into CEGAR, because only the refinement component needs customization, and the new approach remains compatible with off-the-shelf interpolation engines.

Contributions. This paper makes the following key contributions:

- we introduce a domain- and analysis-independent method to extract infeasible *sliced* paths from infeasible error paths,
- we prove that interpolants for such a sliced path are also interpolants for the original infeasible error path,
- we explain that —and how— it is possible to obtain better interpolants (in comparison to the standard approach) from a set of infeasible sliced paths, and that refinement selection plays a significant role in CEGAR,
- we implement the presented concepts in the open-source framework for software verification CPACHECKER, and
- we show by experimental results that the novel approach to obtain better interpolants significantly improves the verification effectiveness and efficiency.

Related Work. The desire to control what interpolants an interpolation engine produces, and trying to make the verification process more efficient by finding good interpolants, is not new. Our goal is to contribute a technique that is independent from the abstract domain that a program analysis uses, and independent from specific properties of interpolation engines.

The first work in this direction was suggesting to make the interpolation configurable such that the user has a choice between strong and weak interpolants, by controlling the *interpolant strength* [16]. This approach is unfortunately not implemented in the standard interpolation engines; it requires to rewrite the algorithm that extracts interpolants from resolution proofs. The technique of interpolation abstractions [22], a generalization of term abstraction [2], can be used to guide solvers to pick good interpolants. This is achieved by extending the concrete interpolation problem by so called templates (e.g., terms, formulas, uninterpreted functions with free variables) to obtain a more abstract interpolation problem. An interpolant for the abstract interpolation problem is also a solution to the concrete interpolation problem. Because these interpolation abstractions form a lattice, suitable interpolants can be chosen using a cost function. Our approach is independent from the abstract domain and interpolation engine, and does not rely on SMT solving. For example, our technique is applicable to value and octagon domains.

Path slicing [20] is a technique that was introduced to reduce the burden of the interpolation engine: Before the constraints of the path are given to the interpolation engine, the constraints are weakened by removing facts that are not important for the infeasibility of the error path, i.e., a more abstract error path

is constructed. We also make the error path more abstract, but in different directions to obtain different interpolant sequences, from which we can choose the ones that yield the best abstract model. While path slicing is interested in reducing the run time of the *interpolation engine* (by omitting some facts), we are interested in reducing the run time of the *verification engine* (by spending more time on interpolation but creating a better abstract model).

II. BACKGROUND

Our approach is based on several existing concepts, and in this section we remind the reader of some basic definitions and our previous work in this field [11].

Programs, Control-Flow Automata, States, Paths, Precisions. We restrict the presentation to a simple imperative programming language, where all operations are either assignments or assume operations, and all variables range over integers¹. A program is represented by a control-flow automaton (CFA). A CFA $A = (L, l_0, G)$ consists of a set L of program locations, which model the program counter, an initial program location $l_0 \in L$, which models the program entry, and a set $G \subseteq L \times Ops \times L$ of control-flow edges, which model the operations that are executed when control flows from one program location to the next. The set of program variables that occur in operations from Ops is denoted by X . A *verification problem* $P = (A, l_e)$ consists of a CFA A , representing the program, and a target program location $l_e \in L$, which represents the specification, i.e., “the program must not reach location l_e ”.

A *concrete data state* of a program is a variable assignment $cd : X \rightarrow \mathbb{Z}$, which assigns to each program variable an integer value; the set of integer values is denoted as \mathbb{Z} . A *concrete state* of a program is a pair (l, cd) , where $l \in L$ is a program location and cd is a concrete data state. The set of all concrete states of a program is denoted by \mathcal{C} , a subset $r \subseteq \mathcal{C}$ is called *region*. Each edge $g \in G$ defines a labeled transition relation $\xrightarrow{g} \subseteq \mathcal{C} \times \{g\} \times \mathcal{C}$. The complete transition relation \rightarrow is the union over all control-flow edges: $\rightarrow = \bigcup_{g \in G} \xrightarrow{g}$. We write $c \xrightarrow{g} c'$ if $(c, g, c') \in \rightarrow$, and $c \rightarrow c'$ if there exists a g with $c \xrightarrow{g} c'$.

An *abstract data state* represents a region of concrete data states, formally defined as abstract variable assignment. An *abstract variable assignment* is a partial function $v : X \dashrightarrow \mathbb{Z}$ or \perp , where v maps variables in its definition range to integer values, and \perp is used to represent no variable assignment (i.e., no value is possible, similar to the predicate *false* in logic). The special abstract variable assignment $\top = \{\}$ does not map any variable to a value and is used as initial abstract variable assignment in a program analysis. Variables that do not occur in the definition range of an abstract variable assignment are either omitted by purpose for abstraction in the analysis, or the analysis is not able to determine a concrete value (e.g., resulting from an uninitialized variable declaration or from an external function call). For two partial functions f

and f' , we write $f(x) = y$ for the predicate $(x, y) \in f$, and $f(x) = f'(x)$ for the predicate $\exists c : (f(x) = c) \wedge (f'(x) = c)$. We denote the *definition range* for a partial function f as $\text{def}(f) = \{x \mid \exists y : f(x) = y\}$, and the *restriction* of a partial function f to a new definition range Y as $f|_Y = f \cap (Y \times \mathbb{Z})$. An abstract variable assignment v represents the set $\llbracket v \rrbracket$ of all concrete data states cd for which v is valid, formally: $\llbracket \perp \rrbracket = \{\}$ and for all $v \neq \perp$, $\llbracket v \rrbracket = \{cd \mid \forall x \in X : v(x) = cd(x)\}$. The abstract variable assignment \perp is called *contradicting*. The *implication* for abstract variable assignments is defined as follows: v implies v' (written $v \Rightarrow v'$) if $v = \perp$, or for all variables $x \in \text{def}(v')$ we have $v(x) = v'(x)$. The *conjunction* for abstract variable assignments v and v' is defined as:

$$v \wedge v' = \begin{cases} \perp & \text{if } v = \perp \text{ or } v' = \perp \\ & \text{or } \exists x \in \text{def}(v) \cap \text{def}(v') : v(x) \neq v'(x) \\ v \cup v' & \text{otherwise} \end{cases}$$

The *semantics of an operation* $op \in Ops$ is defined by the strongest-post operator $SP_{op}(\cdot)$: given an abstract variable assignment v , $SP_{op}(v)$ represents the set of concrete data states that are reachable from the concrete data states in the set $\llbracket v \rrbracket$ by executing op . Formally, given an abstract variable assignment v and an assignment operation $x := exp$, we have $SP_{x:=exp}(v) = v|_{X \setminus \{x\}} \wedge v_x$ with

$$v_x = \begin{cases} \{(x, c)\} & \text{if } c \in \mathbb{Z} \text{ is the result of the arithmetic} \\ & \text{evaluation of expression } exp|_v \\ \{\} & \text{otherwise (if } exp|_v \text{ cannot be evaluated)} \end{cases}$$

where $exp|_v$ denotes the interpretation of expression exp for the abstract variable assignment v . Given an abstract variable assignment v and an assume operation $[p]$, we have $SP_{[p]}(v) = \perp$ if $v = \perp$ or the predicate $p|_v$ is unsatisfiable, or we have $SP_{[p]}(v) = v \wedge v_p$ with $v_p = \{(x, c) \in (X \setminus \text{def}(v) \times \mathbb{Z}) \mid p|_v \Rightarrow (x = c)\}$, and $p|_v = p \wedge \bigwedge_{y \in \text{def}(v)} y = v(y)$.

A *path* σ is a sequence $\langle (op_1, l_1), \dots, (op_n, l_n) \rangle$ of pairs of an operation and a location. The path σ is called *program path* if for every i with $1 \leq i \leq n$ there exists a CFA edge $g = (l_{i-1}, op_i, l_i)$ and l_0 is the initial program location, i.e., σ represents a syntactic walk through the CFA. The result of appending the pair (op_n, l_n) to a path $\sigma = \langle (op_1, l_1), \dots, (op_m, l_m) \rangle$ is defined as $\sigma \wedge (op_n, l_n) = \langle (op_1, l_1), \dots, (op_m, l_m), (op_n, l_n) \rangle$. Every path $\sigma = \langle (op_1, l_1), \dots, (op_n, l_n) \rangle$ defines a *constraint sequence* $\gamma_\sigma = \langle op_1, \dots, op_n \rangle$. The *conjunction* $\gamma \wedge \gamma'$ of two constraint sequences $\gamma = \langle op_1, \dots, op_n \rangle$ and $\gamma' = \langle op'_1, \dots, op'_m \rangle$ is defined as their concatenation, i.e., $\gamma \wedge \gamma' = \langle op_1, \dots, op_n, op'_1, \dots, op'_m \rangle$, the *implication* of γ and γ' (denoted by $\gamma \Rightarrow \gamma'$) as the implication of their strongest-post assignments $SP_\gamma(\top) \Rightarrow SP_{\gamma'}(\top)$, and γ is *contradicting* if $SP_\gamma(\top) = \perp$. The *semantics of a path* $\sigma = \langle (op_1, l_1), \dots, (op_n, l_n) \rangle$ is defined as the successive application of the strongest-post operator to each operation of the corresponding constraint sequence γ_σ : $SP_{\gamma_\sigma}(v) = SP_{op_n}(\dots SP_{op_1}(v) \dots)$. The set of concrete program states that result from running a program path σ is represented

¹Our implementation is based on CPACHECKER, which operates on C programs; non-recursive function calls are supported.

Algorithm 1 CEGAR(\mathbb{D}, e_0, π_0), cf. [11]

Input: a CPA with dynamic precision adjustment \mathbb{D} and an initial abstract state $e_0 \in E$ with precision $\pi_0 \in (L \rightarrow 2^\Pi)$
Output: verification result TRUE (property holds) or FALSE
Variables: a set reached of elements of $E \times (L \rightarrow 2^\Pi)$, a set waitlist of elements of $E \times (L \rightarrow 2^\Pi)$, and an error path $\sigma = \langle (op_1, l_1), \dots, (op_n, l_n) \rangle$

```
1: reached :=  $\{(e_0, \pi_0)\}$ ; waitlist :=  $\{(e_0, \pi_0)\}$ ;  $\pi := \pi_0$ 
2: while true do
3:   (reached, waitlist) := CPA( $\mathbb{D}$ , reached, waitlist)
4:   if waitlist =  $\{\}$  then
5:     return TRUE
6:   else
7:      $\sigma := \text{ExtractErrorPath}(\text{reached})$ 
8:     if IsFeasible( $\sigma$ ) then // error path is feasible: report bug
9:       return FALSE
10:    else // error path is infeasible: refine and restart
11:       $\pi := \pi \cup \text{Refine}(\sigma)$ 
12:      reached :=  $\{(e_0, \pi)\}$ ; waitlist :=  $\{(e_0, \pi)\}$ 
```

by the pair $(l_n, \text{SP}_{\gamma_\sigma}(\top))$, where \top is the initial abstract variable assignment. A path σ is *feasible* if $\text{SP}_{\gamma_\sigma}(\top)$ is not contradicting, i.e., $\text{SP}_{\gamma_\sigma}(\top) \neq \perp$. A concrete state (l_n, cd_n) is *reachable*, denoted by $(l_n, cd_n) \in \text{Reach}$, if there exists a feasible program path $\sigma = \langle (op_1, l_1), \dots, (op_n, l_n) \rangle$ with $cd_n \in \llbracket \text{SP}_{\gamma_\sigma}(\top) \rrbracket$. A location l is reachable if there exists a concrete data state cd such that (l, cd) is reachable. A program is *safe* (the specification is satisfied) if l_e is not reachable. A path $\sigma = \langle (op_1, l_1), \dots, (op_n, l_e) \rangle$, which ends in l_e , is called *error path*.

The *precision* is a function $\pi : L \rightarrow 2^\Pi$, where Π depends on the abstract domain used by the analysis. It assigns to each program location some analysis-dependent information that defines the level of abstraction of the analysis. For example, when using predicate abstraction, the set Π is a set of predicates over program variables. When using a value domain, the set Π is the set X of program variables, and a precision defines which program variables should be tracked by the analysis at which program location.

Counterexample-Guided Abstraction Refinement (CEGAR). CEGAR is a technique for automatic iterative refinement of an abstract model [14]. CEGAR is based on three concepts: (1) a *precision*, which determines the current level of abstraction, (2) a *feasibility check*, deciding if an error path (the counterexample) is feasible, and (3) a *refinement* procedure, which takes as input an infeasible error path and extracts a precision to refine the abstract model such that the infeasible error path is eliminated from further exploration. Algorithm 1 shows an outline of a generic and simple CEGAR algorithm. It uses the CPA algorithm [8, 11] for program analysis with dynamic precision adjustment and an abstract domain \mathbb{D} that is formalized as a configurable program analysis (CPA) with dynamic precision adjustment. The CPA uses a set E of abstract states and a set $L \rightarrow 2^\Pi$ of precisions. The analysis algorithm computes the sets reached and waitlist, which represent the current reachable

Algorithm 2 Refine(σ)

Input: an infeasible error path $\sigma = \langle (op_1, l_1), \dots, (op_n, l_n) \rangle$
Output: a precision π
Variables: a constraint sequence Γ

```
1:  $\Gamma := \langle \rangle$ 
2:  $\pi(l) := \{\}$ , for all program locations  $l$ 
3: for  $i := 1$  to  $n - 1$  do
4:    $\gamma^+ := \langle op_{i+1}, \dots, op_n \rangle$ 
5:    $\Gamma := \text{Interpolate}(\Gamma \wedge op_i, \gamma^+)$  // inductive interpolation
6:    $\pi(l_i) := \text{ExtractPrecision}(\Gamma)$  // create precision based on  $\Gamma$ 
7: return  $\pi$ 
```

abstract states with precisions and the frontier, respectively. The analysis algorithm is run first with π_0 as coarse initial precision (usually $\pi_0(l) = \{\}$ for all $l \in L$). If all program states have been exhaustively checked and no error was reached, indicated by an empty waitlist, then the CEGAR algorithm terminates and reports TRUE (program is safe). If the analysis algorithm finds an error in the abstract state space, then it stops and returns the yet incomplete sets reached and waitlist. Now the corresponding abstract error path is extracted from the set reached, using the procedure ExtractErrorPath, and passed to the procedure IsFeasible for the *feasibility check*. If the abstract error path is feasible, meaning there exists a corresponding concrete error path, then this error path represents a violation of the specification and the algorithm terminates, reporting FALSE. If the error path is infeasible, i.e., is not corresponding to a concrete program path, then the precision was too coarse and needs to be refined. The *refinement* step is performed by the procedure Refine (cf. Alg. 2) which returns a precision π that makes the analysis strong enough to exclude the infeasible error path from future state-space explorations. This returned precision is used to extend the current precision of the CEGAR algorithm, which starts its next iteration, delegating to the analysis algorithm the re-computation of the sets reached and waitlist based on this refined precision. CEGAR is often used with lazy abstraction [19] to avoid re-discovering the whole state space after each refinement, but instead removing only those parts of reached and waitlist that need to be re-analyzed with the new precision.

Interpolation for Constraint Sequences. An *interpolant* for two constraint sequences γ^- and γ^+ , such that $\gamma^- \wedge \gamma^+$ is contradicting, is a constraint sequence Γ for which 1) the implication $\gamma^- \Rightarrow \Gamma$ holds, 2) the conjunction $\Gamma \wedge \gamma^+$ is contradicting, and 3) the interpolant Γ contains in its constraints only variables that occur in both γ^- and γ^+ [11].

In the following, we will introduce our novel approach, which extends the procedure Refine to not only perform interpolation on a single infeasible error path, and returning an arbitrary interpolant, but instead, interpolate a set of infeasible sliced prefixes stemming from this single infeasible error path, and offering a set of interpolants from which the most suitable precision may be chosen.

III. SLICED PREFIXES

Our novel technique can be used to extend any approach that is based on CEGAR. *Slice-based refinement selection* extracts from a given infeasible error path not only one single interpolation problem for obtaining a refined precision, but a set of (more abstract, sliced) infeasible error paths and thus a set of interpolation problems, from which the refined precision can be derived. The interpolation problems for the extracted paths are given, one by one, to the interpolation engine, in order to derive interpolants for each path individually. Hence, the abstraction refinement of the analysis is no longer dependent on what the interpolation engine produces, but instead it is free to choose from a set of interpolants the one it finds most suitable. The move from solving a single interpolation problem to solving multiple interpolation problems to enable refinement selection, and in the process transforming the refinement selection into an optimization problem, is a key insight of our approach.

Infeasible Sliced Prefixes. A CEGAR-based analysis usually encounters an infeasible error path due to the coarse precision that it starts with. This occurs when there exists a path to the error location that contains at least one assume operation that is feasible when the reachability algorithm computes *abstract* successors based on the current precision, but is actually contradicting under the *concrete* semantics of the program. Every infeasible error path contains at least one such contradicting assume operation, but often, there exist several independent contradicting assume operations in an infeasible error path, which leads to the notion of infeasible sliced prefixes: A path $\phi = \langle (op_1, l_1), \dots, (op_w, l_w) \rangle$ is a *sliced prefix* for a program path $\sigma = \langle (op_1, l_1), \dots, (op_n, l_n) \rangle$ if $w \leq n$ and for all $1 \leq i \leq w$, we have $\phi.l_i = \sigma.l_i$ and $(\phi.op_i = \sigma.op_i$ or $(\phi.op_i = [true]$ and $\sigma.op_i$ is assume op)), i.e., a sliced prefix results from a path by omitting pairs of operations and locations from the end, and possibly replacing some assume operations by no-op operations. If a sliced prefix for σ is infeasible, then σ is infeasible.

Extracting Infeasible Sliced Prefixes from an Infeasible Error Path. Algorithm 3 is capable of extracting from an infeasible error path all its infeasible sliced prefixes, i.e., all paths from the initial program operation to a contradicting assume operation. The algorithm iterates through the given infeasible error path σ . It keeps incrementing a sliced path prefix σ_f that contains all operations from σ that were seen so far, except the contradicting assume operations, which are replaced by no-op operations. Thus, σ_f always stays feasible. For every element (op, l) from the original path σ (iterating in order from the first to the last pair), we check whether it contradicts σ_f , which is the case if the result of the strongest-post operator for the path $\sigma_f \wedge (op, l)$ is contradicting (denoted by \perp). If so, the algorithm has found a new infeasible sliced prefix. In any case, it continues with the next element after extending σ_f (either by the current operation, or by a no-op operation if the current operation is contradicting). When the algorithm terminates, which is guaranteed because σ is finite, the set Σ contains all infeasible sliced prefixes of σ . There

Algorithm 3 ExtractSlicedPrefixes(σ)

Input: an infeasible path $\sigma = \langle (op_1, l_1), \dots, (op_n, l_n) \rangle$

Output: a non-empty set $\Sigma = \{\sigma_1, \dots, \sigma_n\}$ of infeasible sliced prefixes of σ

Variables: a path σ_f that is always feasible

```

1:  $\Sigma := \{\}; \sigma_f := \langle \rangle$ 
2: for each  $(op, l) \in \sigma$  // iterate in order from  $(op_1, l_1)$  to  $(op_n, l_n)$ 
   do
3:   if  $SP_{\sigma_f \wedge (op, l)}(\top) = \perp$  then
4:     // add  $\sigma_f \wedge (op, l)$  to the set of infeasible sliced prefixes
5:      $\Sigma := \Sigma \cup \{\sigma_f \wedge (op, l)\}$ 
6:      $\sigma_f := \sigma_f \wedge ([true], l)$  // append no-op
7:   else
8:      $\sigma_f := \sigma_f \wedge (op, l)$  // append original pair
9: return  $\Sigma$ 

```

is always at least one infeasible sliced prefix because σ is infeasible.

Algorithm 3 returns the set of all infeasible sliced prefixes. Each of these sliced prefixes has some interesting characteristics: (1) Each sliced prefix ϕ starts with the initial operation op_1 , and ends with an assume operation that contradicts the previous operations of the sliced prefix, i.e., $SP_\phi(\top) = \perp$. (2) The i -th sliced prefix, excluding its (final and only) contradicting assume operation and location, is a prefix of the $(i + 1)$ -st sliced prefix. (3) All sliced prefixes differ from a prefix of the original infeasible error path σ only in their no-op operations.

The visualizations in Fig. 4 capture the details of this process. Figure 4a shows the original error path. Nodes represent program locations and edges represent operations between these locations (assignments to variables or assume operations over variables, the latter denoted with brackets). To allow easier distinction, program locations that are followed by assume operations are drawn as diamonds, while other program locations are drawn as squares. Contradicting assume operations are drawn with a filled background. The sequence of operations ends in the error state, denoted by l_e . Figure 4b depicts the cascade-like sliced prefixes that the algorithm encounters during its progress. Figure 4c shows the three infeasible sliced prefixes that Alg. 3 returns for this example.

The refinement procedure can now use any of these sliced prefixes to create interpolation problems, and is not bound to a single sequence of interpolants for a single infeasible error path; a refinement selection from different precisions is now possible. The following proposition states that this is a valid refinement process.

Proposition. Let σ be an infeasible error path and ϕ be the i -th infeasible sliced prefix for σ that is extracted by Alg. 3, then all interpolant sequences for ϕ are also interpolant sequences for σ .

Proof. Let $\sigma = \langle (op_1, l_1), \dots, (op_n, l_n) \rangle$ and $\phi = \langle (op_1, l_1), \dots, (op_w, l_w) \rangle$. Let Γ_{ϕ^j} be the j -th interpolant of an interpolant sequence for ϕ , i.e., for the two constraint sequences $\gamma_{\phi^j}^- = \langle op_1, \dots, op_j \rangle$ and $\gamma_{\phi^j}^+ = \langle op_{j+1}, \dots, op_w \rangle$, with $1 \leq j < w$. Because ϕ is infeasible, the two constraint sequences $\gamma_{\phi^j}^-$ and $\gamma_{\phi^j}^+$ are contradicting, and therefore, Γ_{ϕ^j} exists [11]. The interpolant

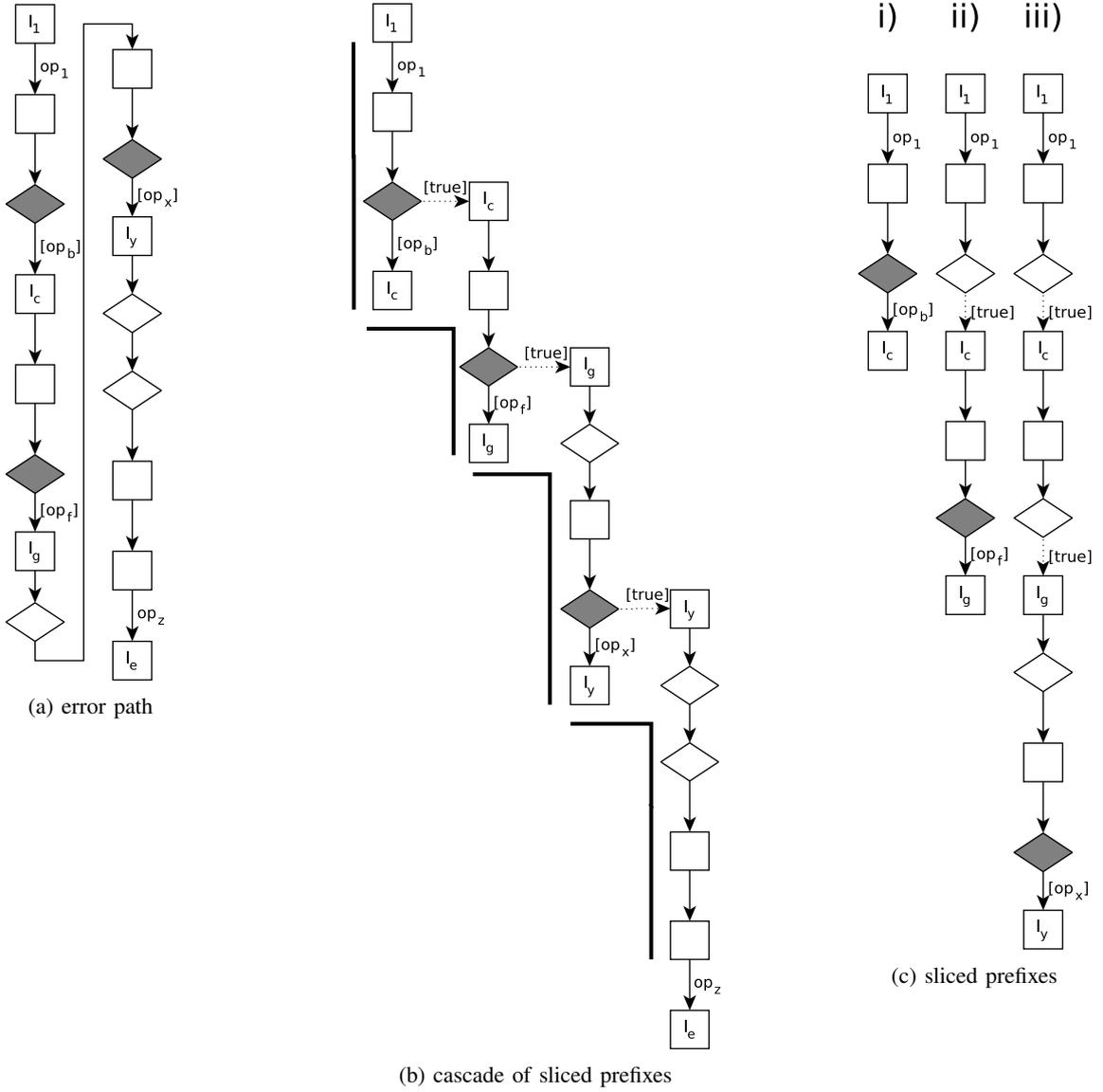


Fig. 4: From one infeasible error path to a set of infeasible sliced prefixes

Γ_{ϕ^j} is also an interpolant for $\gamma_{\sigma^j}^- = \langle op_1, \dots, op_j \rangle$ and $\gamma_{\sigma^j}^+ = \langle op_{j+1}, \dots, op_e \rangle$, if (1) the implication $\gamma_{\sigma^j}^- \Rightarrow \Gamma_{\phi^j}$ holds, (2) the conjunction $\Gamma_{\phi^j} \wedge \gamma_{\sigma^j}^+$ is contradicting, and (3) the interpolant Γ_{ϕ^j} contains only variables that occur in both $\gamma_{\sigma^j}^-$ and $\gamma_{\sigma^j}^+$. Consider that $\gamma_{\phi^j}^-$ was created from $\gamma_{\sigma^j}^-$ by replacing some assume operations by no-op operations, and that $\gamma_{\phi^j}^+$ was created from $\gamma_{\sigma^j}^+$ by replacing some assume operations by no-op operations and by removing the operations $\langle op_{w+1}, \dots, op_n \rangle$ at the end. Thus, both $\gamma_{\phi^j}^-$ and $\gamma_{\phi^j}^+$ do not contain any additional constraints (except for no-op operations) than $\gamma_{\sigma^j}^-$ and $\gamma_{\sigma^j}^+$, respectively. Because Γ_{ϕ^j} is an interpolant for $\gamma_{\phi^j}^-$ and $\gamma_{\phi^j}^+$, we know that $\gamma_{\phi^j}^- \Rightarrow \Gamma_{\phi^j}$ holds, and because $\gamma_{\sigma^j}^-$ can only be stronger than $\gamma_{\phi^j}^-$, Claim (1) follows. The conjunction $\Gamma_{\phi^j} \wedge \gamma_{\phi^j}^+$ is contradicting, and $\gamma_{\sigma^j}^+$ can only be stronger than $\gamma_{\phi^j}^+$. Thus, Claim (2) holds. Because Γ_{ϕ^j} references only variables that occur in both $\gamma_{\phi^j}^-$ and $\gamma_{\phi^j}^+$, which do not contain more variables than $\gamma_{\sigma^j}^-$ and $\gamma_{\sigma^j}^+$, resp., Claim (3) holds.

IV. SLICE-BASED REFINEMENT SELECTION

As described earlier, extracting good precisions from the infeasible error paths is key to the CEGAR technique, and the choice of interpolants influences the quality of the precision, and thus, the effectiveness of the analysis algorithm. By using the results introduced in the previous section, the refinement procedure can now be improved by selecting a precision that is derived via interpolation from a selected sliced prefix.

Algorithm 4 shows our algorithm for slice-based refinement selection, which can be used as a replacement for Alg. 2 in the CEGAR algorithm and chooses a suitable interpolant sequence during the refinement step. First, this algorithm uses ExtractSlicedPrefixes to extract all infeasible sliced prefixes. Second, it computes interpolant sequences for all of them and stores them in the mapping ι . Third, one sliced prefix is chosen by a heuristic (in function ChooseSlicedPrefix) and fourth, the returned precision is created from the interpolants for the chosen sliced prefix. The heuristic can decide based

Algorithm 4 $\text{Refine}^+(\sigma)$

Input: an infeasible error path $\sigma = \langle (op_1, l_1), \dots, (op_n, l_n) \rangle$
Output: a precision π
Variables: a constraint sequence Γ ,
a set Σ of infeasible sliced prefixes of σ ,
a mapping ι from infeasible sliced prefixes and program locations to interpolants

```
1:  $\Sigma := \text{ExtractSlicedPrefixes}(\sigma)$ 
2: // compute interpolants for each location in each prefix
3: for each  $\phi_j = \langle (op_1, l_1), \dots, (op_w, l_w) \rangle \in \Sigma$  do
4:    $\Gamma := \langle \rangle$ 
5:   for  $i := 1$  to  $w - 1$  do
6:      $\gamma^+ := \langle op_{i+1}, \dots, op_w \rangle$ 
7:      $\Gamma := \text{Interpolate}(\Gamma \wedge op_i, \gamma^+)$  // inductive interpolation
8:      $\iota(\phi_j, l_i) := \Gamma$ 
9: // choose suitable sliced prefix
10: // (based on the sliced prefixes and its interpolants)
11:  $\phi_{selected} := \text{ChooseSlicedPrefix}(\Sigma)$ 
12: // create precision based on chosen interpolants
13:  $\pi(l) := \{ \}$ , for all program locations  $l$ 
14: for each  $(op, l) \in \phi_{selected}$  do
15:    $\Gamma := \iota(\phi_{selected}, l)$ 
16:    $\pi(l) := \text{ExtractPrecision}(\Gamma)$  // create precision based on  $\Gamma$ 
17: return  $\pi$ 
```

on the information contained in the sliced prefixes as well as in the interpolants, e.g., which variables are referenced by the interpolants.

Refinement-Selection Heuristics. We regard the selection of interpolants for refinement as an independent direction for further research, but present several ideas on how to select interpolants here. There are two obvious options for interpolant selection that do not depend on the actual interpolants. Using the interpolant sequence derived from the very first, i.e., the shortest, infeasible prefix may rule out many similar infeasible error paths. The downside of this choice is that the analysis has to track information very early, possibly blowing up the state-space and making the analysis less efficient. The other straight-forward option (also known as counterexample minimization [2]) is to use the longest infeasible sliced prefix (containing the last contradicting assume operation) for computing an interpolant sequence. This may lead to a precision that is local to the error location and does not require refining large parts of the state space at the beginning of the error path. However, it may also lead to a larger number of refinements if many error paths with a common prefix exist. A more advanced strategy is to analyze the domain types [3] of the variables that are referenced in the interpolant sequence. Each interpolant sequence can be assigned a score that depends on the domain types of the variables in the interpolant sequence such that the score of the interpolant sequence is better if it references only ‘easy’ types of variables, e.g., boolean variables, and no integer variables or even loop counters. This allows to focus on variables that are inexpensive to track, avoid loop unrolling where possible, and keep the size of the abstract state space as small as possible. Furthermore, it is possible to estimate, by means of the use-def relation of the variables in

the interpolants, how much of the already explored state-space has to be recomputed depending on which interpolant sequence is chosen. Based on that insight, we can identify the interpolant sequence that would ensure that only as little as possible from the state space needs to be re-explored. In addition to that, many different refinement heuristics are conceivable. For example, it would be possible to avoid sliced error paths that contain non-linear arithmetic if using predicate abstraction with an SMT solver for linear arithmetic.

In general, any such heuristic can be used without changing the overall algorithm, but only the function `ChooseSlicedPrefix` in Alg. 4 needs to be replaced accordingly. Using a selection heuristic specifically developed for programs encoding an event-condition-action system improved the effectiveness of our tool `CPACHECKER` in the RERS challenge 2014 and allowed it to obtain two gold and one bronze medals, as well as two special achievements². This shows that optimizing the CEGAR loop by using domain knowledge in the refinement step can be rewarding, and that our approach provides a possibility to do so easily. In the following, we present detailed results for the effectiveness of our approach for a value analysis with the heuristic based on domain types.

V. EXPERIMENTS

We implemented our approach in the open-source verification framework `CPACHECKER`, which is available online³ under the Apache 2.0 license. `CPACHECKER` already has several analyses implemented that can be used for program analysis with CEGAR and lazy abstraction. We only extended the refinement process to work according to Alg. 4 (Refine^+), and did neither change the abstract domains nor the interpolation engines. Our implementation is available in the source-code repository of `CPACHECKER`. The tool, the benchmark programs, the configuration files, and the complete results are available on the supplementary web page⁴.

Setup. We used the same experimental setup as in the International Competition on Software Verification (SV-COMP’14) [6]: machines with Intel Core i7-2600 quad-core CPUs with 3.4 GHz, a memory limit of 15 GB, and a time limit of 15 min. We limited each verification run to one CPU core, because we are interested in the consumed CPU time and the consumed wall time was not important.

Benchmark Programs. For benchmarking we used the C programs of the category “DeviceDrivers64” of SV-COMP’14. This category contains 1 428 large programs based on real-world Linux-kernel device drivers with an average of 6 045 lines of code per program. We consider this category to be especially interesting because our approach focuses on improving refinements in large programs (with long and complex error paths, and many contradicting assume operations per error path). Verification of device drivers is a challenging research topic [12] and an important application domain [5, 21]. For

²Results are available at <http://www.rers-challenge.org/2014Isola/>

³<http://cpachecker.sosy-lab.org>

⁴<http://www.sosy-lab.org/~dbeyer/cpa-ref-sel/>

TABLE I: Results for slice-based refinement selection

Tasks	DeviceDrivers64 (1 428 tasks)		All (2 626 tasks)	
	Classic	Sliced	Classic	Sliced
# Solved	1328	1375	1932	1996
CPU time (h)	28.4	16.9	171	156

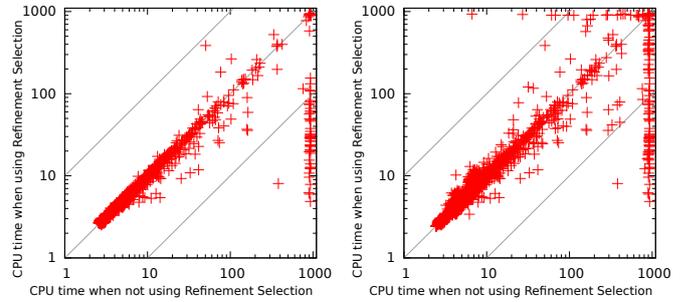
completeness, we also report the results for the 2 626 programs of all categories of SV-COMP’14 except “Concurrency”, “Heap-Manipulation”, “MemorySafety”, and “Recursive”, which rely on features that were not supported by the used configurations of our tool.

Configurations. Out of the several abstract domains that are supported by CPACHECKER, we choose the value analysis with refinement and lazy abstraction [11] for our experiments. This abstract domain tracks explicit values for each program variable, and in case the safety of the program depends on facts that cannot be handled by the value analysis, it delegates to an auxiliary predicate analysis, which is configured for single-block-encoding [10]. We used CPACHECKER in revision 15 509 of tag `cpachecker-1.3.10-refinementSelection`.

When using slice-based refinement selection, the heuristic for choosing sliced prefixes (function `ChooseSlicedPrefix` in Alg. 4) was configured to select the interpolant sequence with the best score based on the domain types of the variables [3] referenced in the interpolants, i.e., variables with a boolean character are favored over integer variables and loop counters.

Results. We now compare the results of running the analysis with both a classic refinement algorithm (as in Alg. 2) and our new refinement algorithm that is based on sliced prefixes (using Alg. 4). Table I shows a summary of the results. The new approach proves to be effective, by solving a total of 1 375 of 1 428 programs correctly in the category “DeviceDrivers64”. Compared to the existing approach, it solves 47 more programs correctly and verifies all programs that could be verified before, too (no regressions). At the same time, the total CPU time was reduced to 60 %. The reason for this vast improvement is that the heuristic for choosing sliced prefixes (guided by the domain type of the referenced variables) is especially effective for the highly complex and heterogeneous program code in Linux-kernel device drivers. On the set of all programs, slice-based refinement selection is effective, too. It can solve 64 more verification problems correctly and needs almost 10 % less time.

Figure 5 shows scatter plots for comparing the CPU time of slice-based refinement selection versus the existing approach on both sets of verification tasks. Only data points for successful verification runs and timeouts are shown (out-of-memory runs are omitted). The figures show that our approach in many cases makes the difference between solving the verification task within the time limit, and not solving the verification task at all (such instances are those at the right border of the plot). This illustrates that without slice-based refinement selection



(a) on category “DeviceDrivers64” (b) on all 2 626 verification tasks

Fig. 5: Scatter plots comparing the CPU when not using slice-based refinement selection (x-axis) with the CPU time when using slice-based refinement selection (y-axis)

and our heuristic for avoiding loop counters in the precision, the interpolants will sometimes be such that the analysis has to unroll long loops, which causes state-space explosion; this can often be avoided with the new approach. The plot also shows that for most of the remaining programs there is no difference in time. This is due to the fact that both sets also contain a large number of small programs, for which our approach does not make a difference, because the counterexamples are short and simple. Figure 5a shows that for the category “DeviceDrivers64”, there is not a single effectiveness regression, i.e., all verification tasks that the classic approach can solve can also be solved by slice-based refinement selection — plus 47 more. Figure 5b shows that on the set of all programs, there are a few regressions where there is a timeout when using the new approach. These are randomly created programs that belong to the “ECA” subset of SV-COMP’14. All variables in these programs have the same domain type, and thus, our heuristic for choosing interpolants based on the domain types of variables is not effective here. For this subset, a heuristic specifically developed for the ECA programs of RERS’14 was successful.

VI. CONCLUSION

In this work we presented our novel approach of *slice-based refinement selection*, which extracts several infeasible sliced prefixes from one single infeasible error path. From any of these infeasible sliced prefixes, an independent interpolation problem can be derived that can be solved by a standard interpolation engine, and the analysis can choose from the resulting interpolant sequences the one thought to be best for the verification. Our novel approach is independent from the abstract domain (in particular, does not depend on an SMT solver) and can be combined with any analysis that is based on CEGAR and interpolation-based abstraction refinement, while previous work on guided interpolation [22] is applicable only to SMT-based approaches. We experimentally demonstrated that the novel approach using a heuristic based on domain types can significantly improve the effectiveness and efficiency of the program analysis. We also discussed some possible further heuristics to select suitable interpolant sequences.

REFERENCES

- [1] A. Albarghouthi, Y. Li, A. Gurfinkel, and M. Chechik. UFO: A framework for abstraction- and interpolation-based software verification. In *Proc. CAV*, LNCS 7358, pages 672–678. Springer, 2012.
- [2] F. Alberti, R. Bruttomesso, S. Ghilardi, S. Ranise, and N. Sharygina. An extension of lazy abstraction with interpolation for programs with arrays. *Formal Methods in System Design*, 45(1):63–109, 2014.
- [3] S. Apel, D. Beyer, K. Friedberger, F. Raimondi, and A. von Rhein. Domain types: Abstract-domain selection based on variable usage. In *Proc. HVC*, LNCS 8244, pages 262–278. Springer, 2013.
- [4] T. Ball, B. Cook, V. Levin, and S.K. Rajamani. SLAM and Static Driver Verifier: Technology transfer of formal methods inside Microsoft. In *Proc. IFM*, LNCS 2999, pages 1–20. Springer, 2004.
- [5] T. Ball and S. K. Rajamani. The SLAM project: Debugging system software via static analysis. In *Proc. POPL*, pages 1–3. ACM, 2002.
- [6] D. Beyer. Status report on software verification (competition summary SV-COMP 2014). In *Proc. TACAS*, LNCS 8413, pages 373–388. Springer, 2014.
- [7] D. Beyer, T. A. Henzinger, R. Jhala, and R. Majumdar. The software model checker BLAST. *Int. J. Softw. Tools Technol. Transfer*, 9(5-6):505–525, 2007.
- [8] D. Beyer, T. A. Henzinger, and G. Théoduloz. Program analysis with dynamic precision adjustment. In *Proc. ASE*, pages 29–38. IEEE, 2008.
- [9] D. Beyer and M. E. Keremoglu. CPACHECKER: A tool for configurable software verification. In *Proc. CAV*, LNCS 6806, pages 184–190. Springer, 2011.
- [10] D. Beyer, M. E. Keremoglu, and P. Wendler. Predicate abstraction with adjustable-block encoding. In *Proc. FMCAD*, pages 189–197. FMCAD, 2010.
- [11] D. Beyer and S. Löwe. Explicit-state software model checking based on CEGAR and interpolation. In *Proc. FASE*, LNCS 7793, pages 146–162. Springer, 2013.
- [12] D. Beyer and A. K. Petrenko. Linux driver verification. In *Proc. ISO/IEC JTC1 SC22 WG2 N1540*, LNCS 7610, pages 1–6. Springer, 2012.
- [13] B. Blanchet, P. Cousot, R. Cousot, J. Feret, L. Mauborgne, A. Miné, D. Monniaux, and X. Rival. A static analyzer for large safety-critical software. In *Proc. PLDI*, pages 196–207. ACM, 2003.
- [14] E. M. Clarke, O. Grumberg, S. Jha, Y. Lu, and H. Veith. Counterexample-guided abstraction refinement for symbolic model checking. *J. ACM*, 50(5):752–794, 2003.
- [15] W. Craig. Linear reasoning. A new form of the Herbrand-Gentzen theorem. *J. Symb. Log.*, 22(3):250–268, 1957.
- [16] V. D’Silva, D. Kröning, M. Purandare, and G. Weissenbacher. Interpolant strength. In *Proc. VMCAI*, LNCS 5944, pages 129–145. Springer, 2010.
- [17] S. Graf and H. Saïdi. Construction of abstract state graphs with Pvs. In *Proc. CAV*, LNCS 1254, pages 72–83. Springer, 1997.
- [18] T. A. Henzinger, R. Jhala, R. Majumdar, and K. L. McMillan. Abstractions from proofs. In *Proc. POPL*, pages 232–244. ACM, 2004.
- [19] T. A. Henzinger, R. Jhala, R. Majumdar, and G. Sutre. Lazy abstraction. In *Proc. POPL*, pages 58–70. ACM, 2002.
- [20] R. Jhala and R. Majumdar. Path slicing. In *Proc. PLDI*, pages 38–47. ACM, 2005.
- [21] A. V. Khoroshilov, V. S. Mutilin, A. K. Petrenko, and V. Zakharov. Establishing Linux driver verification process. In *Proc. Ershov Memorial Conference*, LNCS 5947, pages 165–176. Springer, 2009.
- [22] P. Rümmer and P. Subotic. Exploring interpolants. In *Proc. FMCAD*, pages 69–76. IEEE, 2013.