Deriving Invariants by Algorithmic Learning, Decision Procedures, and Predicate Abstraction*

Yungbum Jung[1], Soonho Kong[1], Bow-Yaw Wang[2], and Kwangkeun Yi[1]

[1] School of Computer Science and Engineering, Seoul National University
{dreameye,soon,kwang}@ropas.snu.ac.kr
[2] Institute of Information Science, Academia Sinica
bywang@iis.sinica.edu.tw

Abstract. By combining algorithmic learning, decision procedures, and predicate abstraction, we present an automated technique for finding loop invariants in propositional formulae. Given invariant approximations derived from pre- and post-conditions, our new technique exploits the flexibility in invariants by a simple randomized mechanism. The proposed technique is able to generate invariants for some Linux device drivers and SPEC2000 benchmarks in our experiments.

1 Introduction

Algorithmic learning has been applied to assumption generation in compositional reasoning [9]. In contrast to traditional techniques, the learning approach does not derive assumptions in an off-line manner. It instead finds assumptions by interacting with a model checker progressively. Since assumptions in compositional reasoning are generally not unique, algorithmic learning can exploit the flexibility in assumptions to attain preferable solutions. Applications in formal verification and interface synthesis have also been reported [9, 1, 2, 18, 7].

Finding loop invariants follows a similar pattern. Invariants are often not unique. Indeed, programmers derive invariants incrementally. They usually have their guesses of invariants in mind, and gradually refine their guesses by observing program behavior more. Since in practice there are many invariants for given pre- and post-conditions, programmers have more freedom in deriving invariants. Yet traditional invariant generation techniques do not exploit the flexibility. They have a similar impediment to traditional assumption generation.

This article reports our first findings in applying algorithmic learning to invariant generation. We show that the three technologies (algorithmic learning,

* This work was supported by (A) the Engineering Research Center of Excellence Program of Korea Ministry of Education, Science and Technology(MEST) / Korea Science and Engineering Foundation(KOSEF) Grant Number R11-2008-007-01002-0, (B) the Brain Korea 21 Project, School of Electrical Engineering and Computer Science, Seoul National University, (C) SK Telecom, and (D) National Science Council of Taiwan Grant Numbers 95-2221-E-001-024-MY3 and 97-2221-E-001-006-MY3.

decision procedures, and predicate abstraction) can be arranged in concert to derive loop invariants in propositional (or, quantifier-free) formulae. The new technique is able to generate invariants for some Linux device drivers and SPEC2000 benchmarks without any help from static or dynamic analyses.

For a while loop, an exact learning algorithm for Boolean formulae searches for invariants by asking queries. Queries can be resolved (not always, see below) by decision procedures automatically. Recall that the learning algorithm generates only Boolean formulae but decision procedures work in propositional formulae. We thus perform predicate abstraction and concretization to integrate the two components.

In reality, information about loop invariant is incomplete. Queries may not be resolvable due to insufficient information. One striking feature of our learning approach is to exploit the flexibility in invariants. When query resolution requires information unavailable to decision procedures, we simply give a random answer. We surely could use static analysis to compute soundly approximated information other than random answers. Yet there are so many invariants for the given pre- and post-conditions. A little bit of random information does not prevent algorithmic learning from inferring invariants. Indeed, the learning algorithm is able to derive invariants in our experiments by coin tossing.

Example

$$\{i = 0\}\ \texttt{while}\ i < 10\ \texttt{do}\ b := \texttt{nondet};\ \texttt{if}\ b\ \texttt{then}\ i := i + 1\ \texttt{end}\ \{i = 10\ \wedge\ b\}$$

The while loop assigns a random truth value to the variable b in the beginning of its body. It increases the variable i by 1 if b is true. Observe that the variable b must be true after the while loop. We would like to find an invariant which proves the postcondition $i = 10 \wedge b$. Heuristically, we choose $i = 0$ and $(i = 10 \wedge b) \vee i < 10$ as under- and over-approximations to invariants respectively. With the help of a decision procedure, these invariant approximations are used to resolve queries made by the learning algorithm. After resolving a number of queries, the learning algorithm asks whether $i \neq 0 \wedge i < 10 \wedge \neg b$ should be included in the invariant. Note that the query is not stronger than the under-approximation, nor weaker than the over-approximation. Hence decision procedures cannot resolve it due to lack of information. At this point, one could apply static analysis and see that it is possible to have this state at the beginning of the loop. Instead of employing static analysis, we simply give a random answer to the learning algorithm. For this example, this information is crucial: the learning algorithm will ask us to give a counterexample to its best guess $i = 0 \vee (i = 10 \wedge b)$ after it processes the incorrect answer. Since the guess is not an invariant and flipping coins does not generate a counterexample, we restart the learning process. If the query $i \neq 0 \wedge i < 10 \wedge \neg b$ is answered correctly, the learning algorithm infers the invariant $(i = 10 \wedge b) \vee i < 10$ with two more resolvable queries.

Contribution

- We prove that algorithmic learning, decision procedures, and predicate abstraction in combination can automatically infer invariants in propositional formulae for programs in our simple language.
- We demonstrate that the technique works in realistic settings: we are able to generate invariants for some Linux device drivers and SPEC2000 benchmarks in our experiments.
- The technique can be seen as a framework for invariant generation. Static analyzers can contribute by providing information to algorithmic learning. Ours is hence orthogonal to existing techniques.

We organize this paper as follows. After preliminaries (Section 2), we present an overview of the framework in Section 3. In Section 4, we review the exact learning algorithm introduced in [6]. Section 5 gives the details of our learning approach. We report experiments in Section 6. Section 7 briefly discusses our learning approach, future work, and related work. Section 8 concludes our work.

2 The Target Language and Notation

The syntax of statements in our simple imperative language is as follows.

$$\begin{aligned}\mathsf{Stmt} \triangleq\ & \texttt{nop} \mid \texttt{assume}\ \mathsf{Prop} \mid \mathsf{Stmt};\ \mathsf{Stmt} \mid \\ & x := \mathsf{Exp} \mid x := \texttt{nondet} \mid b := \mathsf{Bool} \mid b := \texttt{nondet} \mid \\ & \texttt{if}\ \mathsf{Prop}\ \texttt{then}\ \mathsf{Stmt}\ \texttt{else}\ \mathsf{Stmt} \mid \texttt{switch}\ \mathsf{Exp}\ \texttt{do}\ \texttt{case}\ \mathsf{Exp}:\ \mathsf{Stmt}\ \cdots \mid \\ & \{\ \mathsf{Prop}\ \}\ \texttt{while}\ \mathsf{Prop}\ \texttt{do}\ \mathsf{Stmt}\ \{\ \mathsf{Prop}\ \}\end{aligned}$$

Natural number variables and Boolean variables are allowed. They assign to arbitrary values in their respective domains by the keyword `nondet`. Note that `while` statements are annotated. Programmers are asked to specify a *precondition* before a `while` statement, and a *postcondition* after the statement.

An *expression* Exp is a natural number ($n \in \mathbb{N}$), a variable (x), or a summation or the difference of two expressions. Due to the limitation of decision procedures, only linear arithmetic is allowed. It ensures complete answers from decision procedures.

$$\mathsf{Exp} \triangleq n \mid x \mid \mathsf{Exp} + \mathsf{Exp} \mid \mathsf{Exp} - \mathsf{Exp}$$

A *propositional formula* Prop is either: the falsehood symbol (`F`), a Boolean variable (b), the negation of a propositional formula, the conjunction of two propositional formulae, or comparisons ($E_0 < E_1$ or $E_0 = E_1$).

$$\mathsf{Prop} \triangleq \texttt{F} \mid b \mid \neg\mathsf{Prop} \mid \mathsf{Prop} \wedge \mathsf{Prop} \mid \mathsf{Exp} < \mathsf{Exp} \mid \mathsf{Exp} = \mathsf{Exp}$$

Let ρ_0 and ρ_1 be propositional formulae, π_0 and π_1 be expressions. We write `T` for $\neg$`F`, $\rho_0 \vee \rho_1$ for $\neg(\neg\rho_0 \wedge \neg\rho_1)$, $\rho_0 \Rightarrow \rho_1$ for $\neg\rho_0 \vee \rho_1$, $\rho_0 \Leftrightarrow \rho_1$ for $(\rho_0 \Rightarrow \rho_1) \wedge (\rho_1 \Rightarrow \rho_0)$, $\rho_0 \oplus \rho_1$ for $\neg(\rho_0 \Leftrightarrow \rho_1)$, $\pi_0 \leq \pi_1$ for $\pi_0 < \pi_1 \vee \pi_0 = \pi_1$, and $\pi_0 \neq \pi_1$ for $\neg(\pi_0 = \pi_1)$. Propositional formulae of the forms b, $\pi_0 < \pi_1$, and

$\pi_0 = \pi_1$ are called *atomic propositions*. If A is a set of atomic propositions, Prop_A denotes the set of propositional formulae generated from A.

A *Boolean formula* Bool is a restricted propositional formula constructed from truth values and Boolean variables.

$$\mathsf{Bool} \triangleq \mathtt{F} \mid b \mid \neg\mathsf{Bool} \mid \mathsf{Bool} \wedge \mathsf{Bool}$$

A *valuation* ν is an assignment of natural numbers to variables and truth values to Boolean variables. A *Boolean valuation* μ is an assignment of truth values to Boolean variables. If A is a set of atomic propositions and $Var(A)$ is the set of variables occurred in A, $Val_{Var(A)}$ denotes the set of valuations for $Var(A)$. Let ρ be a propositional formula. The valuation ν is a *model* of ρ (written $\nu \models \rho$) if ρ evaluates to $\mathtt{T}$ under the valuation ν. Similarly, the Boolean valuation μ is a *Boolean model* of the Boolean formula β (written $\mu \models \beta$) if β evaluates to $\mathtt{T}$ under μ. If B is a set of Boolean variables, the set of Boolean valuations for B is denoted by Val_B. Given a propositional formula ρ, a *satisfiability modulo theories (SMT) solver* returns a model of ρ if it exists (written $SMT(\rho) \rightarrow \nu$); otherwise, it returns $UNSAT$ (written $SMT(\rho) \rightarrow UNSAT$) [11, 22].

A *precondition* $Pre(\phi, S)$ for $\phi \in \mathsf{Prop}$ with respect to a statement S is a universally quantified formula that guarantees ϕ after the execution of the statement S.

$$\begin{aligned}
Pre(\phi, \mathtt{nop}) &= \phi \\
Pre(\phi, \mathtt{assume}\ \theta) &= \theta \Rightarrow \phi \\
Pre(\phi, S_0;\ S_1) &= Pre(Pre(\phi, S_1), S_0) \\
Pre(\phi, x := \pi) &= \begin{cases} \forall x.\phi & \text{if } \pi = \mathtt{nondet} \\ \phi[x \mapsto \pi] & \text{otherwise} \end{cases} \\
Pre(\phi, b := \rho) &= \begin{cases} \forall b.\phi & \text{if } \rho = \mathtt{nondet} \\ \phi[b \mapsto \rho] & \text{otherwise} \end{cases} \\
Pre(\phi, \mathtt{if}\ \rho\ \mathtt{then}\ S_0\ \mathtt{else}\ S_1) &= (\rho \Rightarrow Pre(\phi, S_0)) \wedge (\neg\rho \Rightarrow Pre(\phi, S_1)) \\
Pre(\phi, \mathtt{switch}\ \pi\ \mathtt{case}\ \pi_i\colon S_i) &= \bigwedge_i (\pi = \pi_i \Rightarrow Pre(\phi, S_i)) \\
Pre(\phi, \{\delta\}\ \mathtt{while}\ \rho\ \mathtt{do}\ S\ \{\epsilon\}) &= \begin{cases} \delta \text{ if } \epsilon \text{ implies } \phi \\ \mathtt{F} \text{ otherwise} \end{cases}
\end{aligned}$$

Observe that all universal quantifiers occur positively in $Pre(\phi, S)$ for any S. They can be eliminated by Skolem constants [12, 23].

3 Framework Overview

We combine algorithmic learning, decision procedures [11], and predicate abstraction [13] in our framework. Figure 1 illustrates the relation among these technologies. In the figure, the left side represents the concrete domain; the right side represents the abstract domain. Assume there is an invariant for a `while` statement with respect to the given pre- and post-conditions in the concrete domain. We would like to apply algorithmic learning to find such an invariant.

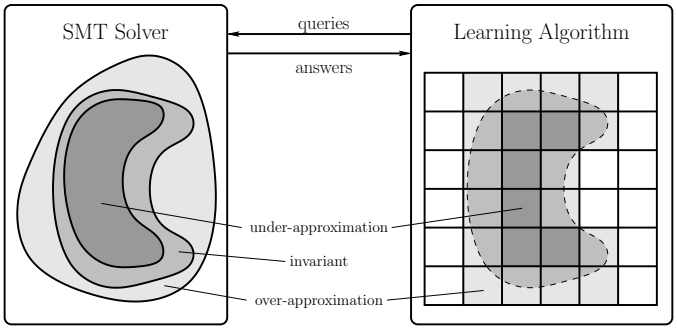


Fig. 1. Overview

To this purpose, we use the CDNF algorithm [6]. The CDNF algorithm is an exact learning algorithm for Boolean formulae. It is an active learning algorithm that makes queries about an unknown Boolean formula and outputs a Boolean formula that is equivalent to the unknown one [3, 6]. We perform predicate abstraction to represent propositional formulae as Boolean formulae in the abstract domain. Since the CDNF algorithm is able to learn arbitrary Boolean formulae, our technique can infer arbitrary invariants in propositional formulae by answering queries.

To realize this idea, we devise a mechanism (a teacher) to resolve queries in the abstract domain. There are two types of queries: membership queries ask whether a Boolean valuation is a model of an invariant; equivalence queries ask whether a Boolean formula is an invariant and demand a counterexample if it is not. It is not difficult to concretize queries in the abstract domain. Answering queries however requires information about invariants yet to be computed.

Although an invariant is unknown, its approximations can be derived from the pre- and post-conditions, or computed by static analysis. Hence, we estimate invariant approximations heuristically and adopt decision procedures for query resolution. For a membership query, we check if its concretization is in the under-approximation or outside the over-approximation by an SMT solver. If it is in the under-approximation, the answer is affirmative; if it is out of the over-approximation, the answer is negative. Otherwise, we simply give a random answer. Equivalence queries are resolved similarly, but we restart the learning process when equivalence queries are not resolvable. If the concretization is not weaker than the under-approximation or not stronger than the over-approximation, a counterexample can be generated by an SMT solver. Otherwise, the learning process is restarted instead of giving random answers.

4 The CDNF Algorithm

In [6], an exact learning algorithm for Boolean formulae over a finite set B of Boolean variables is introduced. The CDNF algorithm generates a conjunction of

formulae in disjunctive normal form equivalent to the unknown Boolean formula λ. It assumes a teacher to answer the following queries:

1. *Membership queries.* Let μ be a Boolean valuation for B. The membership query $MEM(\mu)$ asks if μ is a model of the unknown Boolean formula λ. If $\mu \models \lambda$, the teacher answers *YES* (denoted by $MEM(\mu) \to YES$). Otherwise, the teacher answers *NO* (denoted by $MEM(\mu) \to NO$).
2. *Equivalence queries.* Let $\beta \in \mathsf{Bool}_B$. The equivalence query $EQ(\beta)$ asks if β is equivalent to the unknown Boolean formula λ. If so, the teacher answers *YES* (denoted by $EQ(\beta) \to YES$). Otherwise, the teacher returns a Boolean valuation μ for B such that $\mu \models \beta \oplus \lambda$ as a counterexample (denoted by $EQ(\beta) \to \mu$).

```
(* B = {b_1, b_2, ..., b_m}: a finite set of Boolean variables *)
Input: A teacher answers membership and equivalence queries for an unknown
       Boolean formula λ
Output: A Boolean formula equivalent to λ
t := 0;
if EQ(T) → YES then return T;
let μ be such that EQ(T) → μ;
0 t := t + 1;  (H_t, S_t, a_t) := (F, ∅, μ);
1 if EQ(⋀_{i=1}^{t} H_i) → YES then return ⋀_{i=1}^{t} H_i;
  let μ be such that EQ(⋀_{i=1}^{t} H_i) → μ;
  I := {i : μ ⊭ H_i};
2 if I = ∅ then goto 0;
  foreach i ∈ I do
      μ_i := μ;
      walk from μ_i towards a_i while keeping μ_i ⊨ λ;
      S_i := S_i ∪ {μ_i ⊕ a_i};
  end
  H_i := M_DNF(S_i)[B ↦ B ⊕ a_i] for i = 1, ..., t;
3 goto 1;
```

Algorithm 1: The CDNF Algorithm [6]

Let μ and a be Boolean valuations for B. The Boolean valuation $\mu \oplus a$ is defined by $(\mu \oplus a)(b_i) = \mu(b_i) \oplus a(b_i)$ for $b_i \in B$. For any Boolean formula β, $\beta[B \mapsto B \oplus a]$ is the Boolean formula obtained from β by replacing $b_i \in B$ with $\neg b_i$ if $a(b_i) = \mathtt{T}$. For a set S of Boolean valuations for B, define

$$M_{DNF}(\mu) = \bigwedge_{\mu(b_i)=\mathtt{T}} b_i \quad \text{and} \quad M_{DNF}(S) = \bigvee_{\mu \in S} M_{DNF}(\mu).$$

For the degenerate cases, $M_{DNF}(\mu) = \mathtt{T}$ when $\mu \equiv \mathtt{F}$ and $M_{DNF}(\emptyset) = \mathtt{F}$. Algorithm 1 shows the CDNF algorithm [6]. In the algorithm, the step "walk from μ

towards a while keeping $\mu \models \lambda$" takes two Boolean valuations μ and a. It flips the assignments in μ different from those of a and maintains $\mu \models \lambda$. Algorithm 2 implements the walking step by membership queries.

```
(* B = {b_1, b_2, ..., b_m}: a finite set of Boolean variables *)
Input: valuations μ and a for B
Output: a model μ of λ by walking towards a
i := 1;
while i ≤ m do
    if μ(b_i) ≠ a(b_i) then
        μ(b_i) := ¬μ(b_i);
        if MEM(μ) → YES then  i := 0  else  μ(b_i) := ¬μ(b_i);
    end
    i := i + 1;
end
return μ
```

Algorithm 2: Walking towards a

Intuitively, the CDNF algorithm computes the conjunction of approximations to the unknown Boolean formula. In Algorithm 1, H_i records the approximation generated from the set S_i of Boolean valuations with respect to the Boolean valuation a_i. The algorithm checks if the conjunction of approximations H_i's is the unknown Boolean formula (line **1**). If it is, we are done. Otherwise, the algorithm tries to refine H_i by expanding S_i. If none of H_i's can be refined (line **2**), another approximation is added (line **0**). The algorithm reiterates after refining the approximations H_i's (line **3**). Let λ be a Boolean formula, $|\lambda|_{DNF}$ and $|\lambda|_{CNF}$ denote the minimum sizes of λ in disjunctive and conjunctive normal forms respectively. The CDNF algorithm learns any Boolean formula λ with a polynomial number of queries in $|\lambda|_{DNF}$, $|\lambda|_{CNF}$, and the number of Boolean variables [6]. Appendix A gives a sample run of the CDNF algorithm.

5 Learning Invariants

Consider the `while` statement

$$\{\delta\}\ \texttt{while}\ \rho\ \texttt{do}\ S\ \{\epsilon\}.$$

The propositional formula ρ is called the *guard* of the `while` statement; the statement S is called the *body* of the `while` statement. The annotation is intended to denote that if the precondition δ holds, then the postcondition ϵ must hold after the execution of the `while` statement. The *invariant generation problem* is to compute an invariant to justify the pre- and post-conditions.

Definition 1. *Let* $\{\delta\}$ `while` ρ `do` S $\{\epsilon\}$ *be a* `while` *statement. An* invariant ι *is a propositional formula such that*

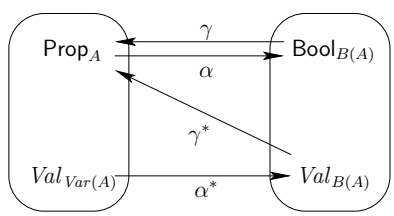
$$(a)\ \delta \Rightarrow \iota \qquad (b)\ \rho \wedge \iota \Rightarrow Pre(\iota, S) \qquad (c)\ \neg\rho \wedge \iota \Rightarrow \epsilon.$$

An invariant allows us to prove that the `while` statement fulfills the annotated requirements. Observe that Definition 1 (c) is equivalent to $\iota \Rightarrow \epsilon \vee \rho$. Along with Definition 1 (a), we see that any invariant must be weaker than δ but stronger than $\epsilon \vee \rho$. Hence δ and $\epsilon \vee \rho$ are called the *strongest* and *weakest* approximations to invariants for $\{\delta\}$ `while` ρ `do` S $\{\epsilon\}$ respectively.

Our goal is to apply the CDNF algorithm (Algorithm 1) to "learn" an invariant for an annotated `while` statement. To achieve this goal, we first lift the invariant generation problem to the abstract domain by predicate abstraction. Moreover, we need to devise a mechanism to answer queries from the learning algorithm in the abstract domain. In the following, we show how to answer queries by an SMT solver and invariant approximations.

5.1 Predicate Abstraction to Connect Algorithmic Learning and SMT Solvers

Domains for an SMT solver and algorithmic learning are adjoined via the predicate abstraction [13]. The α, α^*, γ, and γ^* are the abstraction (α, α^*) and concretization (γ, γ^*) maps between the two domains. SMT solvers work in propositional formulae. Algorithmic learning works in Boolean formulae.



Let A be a fixed set of atomic propositions. For each atomic proposition $p \in A$, we use a Boolean variable b_p to represent p. Let $B(A) = \{b_p : p \in A\}$ be the set of Boolean variables corresponding to the atomic propositions in A. Consider the *concrete domain* Prop_A and the *abstract domain* $\mathsf{Bool}_{B(A)}$. A Boolean formula $\beta \in \mathsf{Bool}_{B(A)}$ is called a *canonical monomial* if it is a conjunction of literals such that each Boolean variable in $B(A)$ appears exactly once. Define the mappings $\gamma : \mathsf{Bool}_{B(A)} \to \mathsf{Prop}_A$ and $\alpha : \mathsf{Prop}_A \to \mathsf{Bool}_{B(A)}$:

$\gamma(\beta) = \beta[\overline{b}_p \mapsto \overline{p}]$; and
$\alpha(\theta) = \bigvee\{\beta \in \mathsf{Bool}_{B(A)} : \beta$ is a canonical monomial and $\theta \wedge \gamma(\beta)$ is satisfiable$\}$.

where $\overline{b}_p$ and $\overline{p}$ are the Boolean variables in $B(A)$ and their corresponding atomic propositions respectively.

The following lemmas are useful in proving our technical results:

Lemma 1. *Let A be a set of atomic propositions, $\theta, \rho \in \mathsf{Prop}_A$. Then*

$$\theta \Rightarrow \rho \text{ implies } \alpha(\theta) \Rightarrow \alpha(\rho).$$

Lemma 2. *Let A be a set of atomic propositions, $\theta \in \mathsf{Prop}_A$, and β a canonical monomial in $\mathsf{Bool}_{B(A)}$. Then $\theta \wedge \gamma(\beta)$ is satisfiable if and only if $\gamma(\beta) \Rightarrow \theta$.*[3]

Recall that a teacher for the CDNF algorithm answers queries in the abstract domain, and an SMT solver computes models in the concrete domain. In order to let an SMT solver play the role of a teacher, more transformations are needed. A valuation induces a natural Boolean valuation. Precisely, define the Boolean valuation $\alpha^*(\nu)$ for the valuation ν as follows.

$$(\alpha^*(\nu))(b_p) = \begin{cases} \mathtt{T} & \text{if } \nu \models p \\ \mathtt{F} & \text{otherwise} \end{cases}$$

Lemma 3. *Let A be a set of atomic propositions, $\theta \in \mathsf{Prop}_A$, $\beta \in \mathsf{Bool}_{B(A)}$, and ν a valuation for $Var(A)$. Then*

1. *$\nu \models \theta$ if and only if $\alpha^*(\nu) \models \alpha(\theta)$; and*
2. *$\nu \models \gamma(\beta)$ if and only if $\alpha^*(\nu) \models \beta$.*

A Boolean valuation on the other hand induces a propositional formula. Define the propositional formula $\gamma^*(\mu)$ for the Boolean valuation μ as follows.

$$\gamma^*(\mu) = \bigwedge_{p \in A} \{p : \mu(b_p) = \mathtt{T}\} \wedge \bigwedge_{p \in A} \{\neg p : \mu(b_p) = \mathtt{F}\}$$

Lemma 4. *Let A be a set of atomic propositions, $\theta \in \mathsf{Prop}_A$, and μ a Boolean valuation for $B(A)$. Then $\gamma^*(\mu) \Rightarrow \theta$ if and only if $\mu \models \alpha(\theta)$.*

5.2 Answering Queries from Algorithmic Learning

Suppose $\iota \in \mathsf{Prop}_A$ is an invariant for the statement $\{\delta\}$ `while` ρ `do` S $\{\epsilon\}$. Let $\underline{\iota}, \overline{\iota} \in \mathsf{Prop}_A$. We say $\underline{\iota}$ is an *under-approximation* to an invariant ι if $\delta \Rightarrow \underline{\iota}$ and $\underline{\iota} \Rightarrow \iota$. Similarly, $\overline{\iota}$ is an *over-approximation* to an invariant ι if $\iota \Rightarrow \overline{\iota}$ and $\overline{\iota} \Rightarrow \epsilon \vee \rho$. The strongest ($\delta$) and weakest ($\epsilon \vee \rho$) approximations are trivial under- and over-approximations to any invariant respectively.

Recall that the CDNF algorithm makes the following queries: (1) membership queries $MEM(\mu)$ where $\mu \in Val_{B(A)}$, and (2) equivalence queries $EQ(\beta)$ where $\beta \in \mathsf{Bool}_{B(A)}$. In the following, we show how to resolve these queries by means of an SMT solver and the invariant approximations ($\underline{\iota}$ and $\overline{\iota}$).

Membership Queries In the membership query $MEM(\mu)$, the teacher is required to answer whether $\mu \models \alpha(\iota)$. We concretize the Boolean valuation μ and check it against the approximations. If the concretization $\gamma^*(\mu)$ is inconsistent (that is, $\gamma^*(\mu)$ is unsatisfiable), we simply answer *NO* for the membership query. Otherwise, there are three cases:

[3] Complete proofs are in [20]

1. $\gamma^*(\mu) \Rightarrow \underline{\iota}$. Thus $\mu \models \alpha(\underline{\iota})$ (Lemma 4). And $\mu \models \alpha(\iota)$ by Lemma 1.
2. $\gamma^*(\mu) \not\Rightarrow \overline{\iota}$. Thus $\mu \not\models \alpha(\overline{\iota})$ (Lemma 4). That is, $\mu \models \neg\alpha(\overline{\iota})$. Since $\iota \rightarrow \overline{\iota}$, we have $\mu \not\models \alpha(\iota)$ by Lemma 1.
3. Otherwise, we cannot determine whether $\mu \models \alpha(\iota)$ by the approximations.

```
(* ι̲: an under-approximation; ῑ: an over-approximation *)
```
Input: a valuation μ for $B(A)$
$\theta := \gamma^*(\mu)$;
if $SMT(\theta) \rightarrow UNSAT$ **then return** NO;
if $SMT(\theta \wedge \neg\underline{\iota}) \rightarrow UNSAT$ **then return** YES;
if $SMT(\theta \wedge \neg\overline{\iota}) \rightarrow \nu$ **then return** NO;
abort with θ;

Algorithm 3: Resolving Membership Queries

Algorithm 3 shows our membership query resolution algorithm. Note that when a membership query cannot be resolved by an SMT solver given invariant approximations, one can use better approximations from static analyzers. Our framework is therefore orthogonal to existing static analysis techniques.

Equivalence Queries To answer the equivalence query $EQ(\beta)$, we concretize the Boolean formula β and check if $\gamma(\beta)$ is indeed an invariant of the `while` statement for the given pre- and post-conditions. If it is, we are done. Otherwise, we use an SMT solver to find a witness to $\alpha(\iota) \oplus \beta$. There are three cases:

1. There is a ν such that $\nu \models \neg(\underline{\iota} \Rightarrow \gamma(\beta))$. Then $\nu \models \underline{\iota} \wedge \neg\gamma(\beta)$. By Lemma 3 and 1, we have $\alpha^*(\nu) \models \alpha(\iota)$ and $\alpha^*(\nu) \models \neg\beta$. Thus, $\alpha^*(\nu) \models \alpha(\iota) \wedge \neg\beta$.
2. There is a ν such that $\nu \models \neg(\gamma(\beta) \Rightarrow \overline{\iota})$. Then $\nu \models \gamma(\beta) \wedge \neg\overline{\iota}$. By Lemma 3, $\alpha^*(\nu) \models \beta$. $\alpha^*(\nu) \models \neg\alpha(\iota)$ by Lemma 3 and 1. Hence $\alpha^*(\nu) \models \beta \wedge \neg\alpha(\iota)$.
3. Otherwise, we cannot find a witness to $\alpha(\iota) \oplus \beta$ by the approximations.

```
(* {δ} while ρ do S {ϵ} *)
(* ι̲: an under-approximation; ῑ: an over-approximation *)
```
Input: $\beta \in \mathsf{Bool}_{B(A)}$
$\theta := \gamma(\beta)$;
if $SMT(\underline{\iota} \wedge \neg\theta) \rightarrow UNSAT$ **and** $SMT(\theta \wedge \neg\overline{\iota}) \rightarrow UNSAT$ **and**
$SMT(\rho \wedge \theta \wedge \neg Pre(\theta, S)) \rightarrow UNSAT$ **then**
  return YES;
if $SMT(\underline{\iota} \wedge \neg\theta) \rightarrow \nu$ **then return** $\alpha^*(\nu)$;
if $SMT(\theta \wedge \neg\overline{\iota}) \rightarrow \nu$ **then return** $\alpha^*(\nu)$;
abort with θ;

Algorithm 4: Resolving Equivalence Queries

Algorithm 4 shows our equivalence query resolution algorithm. Note that Algorithm 4 returns *YES* only if an invariant is found.

Similar to membership query resolution, one can refine approximations by static analysis when an equivalence query is not resolvable by an SMT solver given invariant approximations. For simplicity, Algorithm 4 aborts the learning algorithm with the unresolved equivalence query.

5.3 Main Loop of Our Approach

Algorithm 5 gives the top-level loop of our framework. Initially, we use the disjunction of strongest approximation and the postcondition as the under-approximation; the weakest approximation is the over-approximation. The under-approximation aims to find an invariant that establishes the postcondition. This heuristic is proved very useful in practice.

```
(* {δ} while ρ do S {ε} *)
function randomized_membership μ =
   try Algorithm 3 with input μ when abort → return YES or NO randomly;

ι := δ ∨ ε; ῑ := ε ∨ ρ;
repeat
    try ι := Algorithm 1 with randomized_membership and Algorithm 4
    when abort → continue
until an invariant ι is found ;
```

Algorithm 5: Main Loop

After determining the approximations, Algorithm 1 is used to find an invariant. We use Algorithms 3 and 4 to resolve queries with an SMT solver given the invariant approximations. If Algorithm 3 aborts with an unresolved membership query, a random answer is returned by *randomized_membership*. If Algorithm 4 aborts with an unresolved equivalence query, the learning algorithm is restarted.

Since algorithmic learning does not commit to any specific target, it always finds an invariant consistent with answers to previous queries. In other words, the learning algorithm will always generate an invariant if there is one consistent with our random answers. Although our random answers may exclude certain invariants, an invariant can still be inferred. Verifying whether a formula is an invariant is done by checking the sufficient conditions of Definition 1 in our equivalence query resolution algorithm (Algorithm 4).

6 Experiments

We have implemented a prototype in OCaml. In our implementation, we use Yices as the SMT solver to resolve queries (Algorithm 3 and 4). From SPEC2000 benchmarks and Linux device drivers we chose five `while` statements. We translated them into our language and added postcondition manually. Table 1 shows the performance numbers of our experiments. Among five `while` statements,

the cases `parser` and `vpr` are extracted from PARSER and VPR in SPEC2000 benchmarks respectively. The other three cases are extracted from Linux 2.6.28 device drivers: both `ide-ide-tape` and `ide-wait-ireason` are from IDE driver; `usb-message` is from USB driver. For each case, we report the number of language constructs in the loop ($SIZE$), the number of atomic propositions (AP), the number of membership queries (MEM), the number of equivalence queries (EQ), the number of randomly resolved membership queries (coin tossing), the number of the CDNF algorithm invocations (iterations), and the execution time. The data are the average of 500 runs and collected on a 2.6GHz Intel E5300 Duo Core with 3GB memory running Linux 2.6.28.

case	$SIZE$	AP	MEM	EQ	coin tossing	iterations	time (sec)
`ide-ide-tape`	16	6	18.2	5.2	4.1	1.2	0.055
`ide-wait-ireason`	9	6	216.1	111.8	47.2	9.9	0.602
`parser`	37	20	6694.5	819.4	990.3	12.5	32.120
`usb-message`	18	10	20.1	6.8	1.0	1.0	0.128
`vpr`	8	7	14.5	8.9	11.8	2.9	0.055

Table 1. Performance Numbers

Our technique is able to find invariants for four cases within 1 second. Most interestingly, the learning algorithm is able to find an invariant for `usb-message` regardless of the outcomes of coin tossing. For the most complicated case `parser`, our technique is able to generate an invariant with 991 random membership resolutions in about 33 seconds.

6.1 `ide-ide-tape` from Linux IDE Driver

$\{\ ret = 0 \wedge bh_b_count \leq bh_b_size\ \}$
1 `while` $n > 0$ `do`
2 `if` $(bh_b_size - bh_b_count) < n$ `then` $count := bh_b_size - bh_b_count$
3 `else` $count := n$;
4 $b :=$ `nondet`;
5 `if` b `then` $ret := 1$;
6 $n := n - count$; $bh_b_count := bh_b_count + count$;
7 `if` $bh_b_count = bh_b_size$ `then`
8 $bh_b_size :=$ `nondet`; $bh_b_count :=$ `nondet`; $bh_b_count := 0$;
9 `end`
$\{\ n = 0 \wedge bh_b_count \leq bh_b_size\ \}$

Fig. 2. A Sample Loop in Linux IDE Driver

Figure 2 is a `while` statement extracted from Linux IDE driver.[4] It copies data of size n from tape records. The variable *count* contains the size of the data to be copied from the current record (*bh_b_size* and *bh_b_count*). If the current tape record runs out of data, more data are copied from the next record. The flexibility in invariants can be witnessed in the following run. After successfully resolving 3 equivalence and 7 membership queries, the CDNF algorithm makes the following membership query unresolvable by the invariant approximations:

$$\overbrace{n > 0 \wedge (bh_b_size - bh_b_count) < n \wedge ret \neq 0}^{\rho} \wedge bh_b_count = bh_b_size$$

Answering *NO* to this query leads to the following unresolvable membership query after successfully resolving two more membership query:

$$\rho \wedge bh_b_count \neq bh_b_size \wedge bh_b_count \leq bh_b_size$$

We proceed with a random answer *YES*. After successfully resolving one more membership queries, we reach the following unresolvable membership query:

$$\rho \wedge bh_b_count \neq bh_b_size \wedge bh_b_count > bh_b_size$$

For this query, both answers lead to invariants. Answering *YES* yields the following invariant:

$$n \neq 0 \vee (bh_b_size - bh_b_count) \geq n$$

Answering *NO* yields the following invariant:

$$(bh_b_count \leq bh_b_size \wedge n \neq 0) \vee (bh_b_size - bh_b_count) \geq n$$

Note that they are two different invariants. The equivalence query resolution algorithm (Algorithm 4) ensures that both fulfill the conditions in Definition 1.

6.2 `parser` from VPR in SPEC2000 Benchmarks

Figure 3 shows a sample `while` statement from the `parser` program in SPEC2000 benchmark.[5] In the `while` body, there are three locations where *give_up* or *success* is set to `T`. Thus one of these conditions in the `if` statements must hold (the first conjunct of postcondition). Variable *valid* may get an arbitrary value if *linkages* is not zero. But it cannot be greater than *linkages* by the `assume` statement (the second conjunct of postcondition). The variable *linkages* gets an arbitrary value near the end of the `while` body. But it cannot be greater than 5000 (the fourth conjunct), and always equal to the variable *canonical* (the third

[4] The source code can be found in function `idetape_copy_stage_from_user()` of `drivers/ide/ide-tape.c` in Linux 2.6.28

[5] The source code can be found in function `loop()` of `CINT2000/197.parser/main.c` in SPEC2000.

```
{ phase = F ∧ success = F ∧ give_up = F ∧ cutoff = 0 ∧ count = 0 }
 1 while ¬(success ∨ give_up) do
 2    entered_phase := F;
 3    if ¬phase then
 4       if cutoff = 0 then cutoff := 1;
 5       else if cutoff = 1 ∧ maxcost > 1 then cutoff := maxcost;
 6             else phase := T; entered_phase := T; cutoff := 1000;
 7       if cutoff = maxcost ∧ ¬search then give_up := T;
 8    else
 9       count := count + 1;
10       if count > words then give_up := T;
11    if entered_phase then count := 1;
12    linkages := nondet;
13    if linkages > 5000 then linkages := 5000;
14    canonical := 0; valid := 0;
15    if linkages ≠ 0 then
16       valid := nondet; assume 0 ≤ valid ∧ valid ≤ linkages;
17       canonical := linkages;
18    if valid > 0 then success := T;
19 end
{ (valid > 0 ∨ count > words ∨ (cutoff = maxcost ∧ ¬search))∧
  valid ≤ linkages ∧ canonical = linkages ∧ linkages ≤ 5000 }
```

Fig. 3. A Sample Loop in SPEC2000 Benchmark PARSER

conjunct of postcondition). Despite the complexity of the postcondition and the `while` body, our approach is able to compute an invariant in 13 iterations on average. The execution time and number of iterations vary significantly. They range from 2.22s to 196.52s and 1 to 84 with standard deviations 31.01 and 13.33 respectively. By Chebyshev's inequality [27], our technique infers an invariant within two minutes with probability 0.876.

One of the found invariants is the following:

$$
\begin{aligned}
&\mathit{success} \Rightarrow (\mathit{valid} \leq \mathit{linkages} \wedge \mathit{linkages} \leq 5000 \wedge \mathit{canonical} = \mathit{linkages}) \bigwedge \\
&\mathit{success} \Rightarrow (\neg \mathit{search} \vee \mathit{count} > \mathit{words} \vee \mathit{valid} \neq 0) \bigwedge \\
&\mathit{success} \Rightarrow (\mathit{count} > \mathit{words} \vee \mathit{cutoff} = \mathit{maxcost} \vee (\mathit{canonical} \neq 0 \wedge \mathit{valid} \neq 0 \wedge \mathit{linkages} \neq 0)) \bigwedge \\
&\mathit{give_up} \Rightarrow ((\mathit{valid} = 0 \wedge \mathit{linkages} = 0 \wedge \mathit{canonical} = \mathit{linkages}) \vee \\
&\qquad\qquad (\mathit{canonical} \neq 0 \wedge \mathit{valid} \leq \mathit{linkages} \wedge \mathit{linkages} \leq 5000 \wedge \mathit{canonical} = \mathit{linkages})) \bigwedge \\
&\mathit{give_up} \Rightarrow (\mathit{cutoff} = \mathit{maxcost} \vee \mathit{count} > \mathit{words} \vee \\
&\qquad\qquad (\mathit{canonical} \neq 0 \wedge \mathit{valid} \neq 0 \wedge \mathit{linkages} \neq 0)) \bigwedge \\
&\mathit{give_up} \Rightarrow (\neg \mathit{search} \vee \mathit{count} > \mathit{words} \vee \mathit{valid} \neq 0)
\end{aligned}
$$

This invariant describes the conditions when *success* or *give_up* are true. For instance, it specifies that $\mathit{valid} \leq \mathit{linkages} \wedge \mathit{linkages} \leq 5000 \wedge \mathit{canonical} = \mathit{linkages}$ should hold if *success* is true. In Figure 3, we see that *success* is assigned to `T` at line 18 when *valid* is positive. Yet *valid* is set to 0 at line 14. Hence line 16 and 17 must be executed. Thus, the first ($\mathit{valid} \leq \mathit{linkages}$) and the third

($canonical = linkages$) conjuncts hold. Moreover, line 13 ensures that the second conjunct ($linkages \leq 5000$) holds as well.

7 Discussion and Future Work

The complexity of our technique depends on the distribution of invariants. It works most effectively if invariants are abundant. The number of iterations depends on the outcomes of coin tossing. The main loop may reiterate several times or not even terminate. Our experiments suggest that there are sufficiently many invariants in practice. For each of the 2500 ($= 5 \times 500$) runs, our technique always generates an invariant. On average, it takes 12.5 iterations for the most complicated case `parser`.

Since plentiful of invariants are available, it may appear that one of them can be generated by merely coin tossing. But this is not the case. In `parser`, our technique does not terminate if the under- and over-approximations are the strongest and weakest approximations respectively. Indeed, 6695 membership and 820 equivalence queries are resolved by invariant approximations in this case. Invariant approximations are essential to our framework.

For simplicity, predicates are collected from program texts, pre- and post-conditions in our experiments. Existing predicate discovery techniques can certainly be deployed. Better invariant approximations ($\underline{\iota}$ and $\overline{\iota}$) computed by static analysis can be used in our framework. More precise approximations of $\underline{\iota}$ and $\overline{\iota}$ will improve the performance by reducing the number of iterations via increasing the number of resolvable queries. Also, a variety of techniques from static analysis or loop invariant generation [12, 17, 28, 16, 19, 21, 23, 25] in particular can be integrated to resolve queries in addition to one SMT solver with coin tossing. Such a set of multiple teachers will increase the number of resolvable queries because it suffices to have just one teacher to answer the query to proceed.

In comparison with previous invariant generation techniques [12, 17, 28, 16, 19, 21, 23, 25], we have the following distinguishing features. (1) We do not use fixed point computation nor any static or dynamic analyses. Instead, we use algorithmic learning [6] to search for loop invariants. (2) Templates for invariants are not needed. Our approach does not restrict to specific forms of invariants imposed by templates. (3) We employ SMT solvers instead of theorem provers in our technique. This allows us to take advantages of recent development in efficient SMT algorithms. (4) Our method can be extended and combined with the existing loop invariant techniques.

Related Work Existing impressive techniques for invariant generation can be adopted as the query resolution components (teachers) in our algorithmic learning-based framework. Srivastava and Gulwani [28] devise three algorithms, two of them use fixed point computation and the other uses a constraint based approach [17, 16] to derive quantified invariants. Gupta and Rybalchenko [19] present an efficient invariant generator. They apply dynamic analysis to make invariant generation more efficient. Flanagan and Qadeer use predicate abstraction

to infer universally quantified loop invariants [12]. Predicates over Skolem constants are used to handle unbounded arrays. McMillan [25] extends a paramodulation-based saturation prover to an interpolating prover that is complete for universally quantified interpolants. He also solves the problem of divergence in interpolated-based invariant generation.

8 Conclusions

By combining algorithmic learning, decision procedures, and predicate abstraction, we introduced a technique for invariant generation. The new technique finds invariants guided by query resolution algorithms. Algorithmic learning gives a platform to integrate various techniques for invariant generation; it suffices to design new query resolution algorithms based on existing techniques. The learning algorithm will utilize the information provided by these techniques.

To illustrate the flexibility of algorithmic learning, we deploy a randomized query resolution algorithm. When a membership query cannot be resolved, a random answer is returned to the learning algorithm. Since the learning algorithm does not commit to any specific invariant beforehand, it always finds a solution consistent with query results. Our experiments indeed show that algorithmic learning is able to infer non-trivial invariants with this naïve membership resolution. It is important to exploit the power of coin tossing in our technique.

Acknowledgment We would like to thank anonymous referees for their comments and suggestions. We are grateful to Wontae Choi, Deokhwan Kim, Will Klieber, Sasa Misailovic, Bruno Oliveira, Corneliu Popeea, Hongseok Yang, and Karen Zee for their detailed comments and helpful suggestions. We also thank Heejae Shin for implementing OCaml binding for Yices.

References

1. Alur, R., Cerný, P., Madhusudan, P., Nam, W.: Synthesis of interface specifications for java classes. In: POPL, ACM (2005) 98–109
2. Alur, R., Madhusudan, P., Nam, W.: Symbolic compositional verification by learning assumptions. In: CAV. Volume 3576 of LNCS., Springer (2005) 548–562
3. Angluin, D.: Learning regular sets from queries and counterexamples. Information and Computation **75**(2) (1987) 87–106
4. Balaban, I., Pnueli, A., Zuck, L.: Shape analysis by predicate abstraction. In: VMCAI. Volume 3385 of LNCS., Springer (2005)
5. Ball, T., Cook, B., Das, S., Rajamani, S.K.: Refining approximations in software predicate abstraction. In: TACAS. Volume 2988 of LNCS., Springer (2004) 388–403
6. Bshouty, N.H.: Exact learning boolean functions via the monotone theory. Information and Computation **123** (1995) 146–153
7. Chen, Y.F., Farzan, A., Clarke, E.M., Tsay, Y.K., Wang, B.Y.: Learning minimal separating DFA's for compositional verification. In: TACAS. Volume 5505 of LNCS., Springer (2009) 31–45

8. Clarke, E.M., Grumberg, O., Jha, S., Lu, Y., Veith, H.: Counterexample-guided abstraction refinement. In: CAV. Volume 1855 of LNCS., Springer (2000) 154–169
9. Cobleigh, J.M., Giannakopoulou, D., Păsăreanu, C.S.: Learning assumptions for compositional verification. In: TACAS. Volume 2619 of LNCS., Springer (2003) 331–346
10. Cousot, P., Halbwachs, N.: Automatic discovery of linear restraints among variables of a program. In: POPL, ACM (1978) 84–96
11. Dutertre, B., Moura, L.D.: The Yices SMT solver. Technical report, SRI International (2006)
12. Flanagan, C., Qadeer, S.: Predicate abstraction for software verification. In: POPL, ACM (2002) 191–202
13. Graf, S., Saïdi, H.: Construction of abstract state graphs with pvs. In: CAV. Volume 1254 of LNCS., Springer (1997) 72–83
14. Gulwani, S., Jain, S., Koskinen, E.: Control-flow refinement and progress invariants for bound analysis. In: PLDI, ACM (2009) 375–385
15. Gulwani, S., McCloskey, B., Tiwari, A.: Lifting abstract interpreters to quantified logical domains. In: POPL, ACM (2008) 235–246
16. Gulwani, S., Srivastava, S., Venkatesan, R.: Program analysis as constraint solving. In: PLDI, ACM (2008) 281–292
17. Gulwani, S., Srivastava, S., Venkatesan, R.: Constraint-based invariant inference over predicate abstraction. In: VMCAI. Volume 5403 of LNCS., Springer (2009) 120–135
18. Gupta, A., McMillan, K.L., Fu, Z.: Automated assumption generation for compositional verification. In: CAV. Volume 4590 of LNCS., Springer (2007) 420–432
19. Gupta, A., Rybalchenko, A.: Invgen: An efficient invariant generator. In: CAV. Volume 5643 of LNCS., Springer (2009) 634–640
20. Jung, Y., Kong, S., Bow-Yaw, W., Yi, K.: Deriving invariants by algorithmic learning, decision procedures, and predicate abstraction. Technical Memorandum ROSAEC-2009-004, Research On Software Analysis for Error-Free Computing (2009)
21. Kovács, L., Voronkov, A.: Finding loop invariants for programs over arrays using a theorem prover. In: FASE. LNCS, Springer (2009) 470–485
22. Kroening, D., Strichman, O.: Decision Procedures an algorithmic point of view. EATCS. Springer (2008)
23. Lahiri, S.K., Bryant, R.E., Bryant, A.E.: Constructing quantified invariants via predicate abstraction. In: VMCAI. Volume 2937 of LNCS., Springer (2004) 267–281
24. Lahiri, S.K., Bryant, R.E., Bryant, A.E., Cook, B.: A symbolic approach to predicate abstraction. In: CAV. Volume 2715 of LNCS., Springer (2003) 141–153
25. McMillan, K.L.: Quantified invariant generation using an interpolating saturation prover. In: TACAS, Springer (2008) 413–427
26. Podelski, A., Wies, T.: Boolean heaps. In: SAS. Volume 3672 of LNCS., Springer (2005) 268–283
27. Rosen, K.H.: Discrete Mathematics and Its Applications. McGraw-Hill Higher Education (2006)
28. Srivastava, S., Gulwani, S.: Program verification using templates over predicate abstraction. In: PLDI, ACM (2009) 223–234
29. Zee, K., Kuncak, V., Rinard, M.: Full functional verification of linked data structures. In: PLDI, ACM (2008) 349–361
30. Zee, K., Kuncak, V., Rinard, M.C.: An integrated proof language for imperative programs. In: PLDI, ACM (2009) 338–351

A An Example of the CDNF Algorithm

Let us apply Algorithm 1 to learn the Boolean formula $b_0 \oplus b_1$. The algorithm first makes the query $EQ(\mathtt{T})$ (Figure 4). The teacher responds by giving the valuation $\mu_1(b_0) = \mu_1(b_1) = 0$ (denoted by $\mu_1(b_0b_1) = 00$). Hence Algorithm 1 assigns $\emptyset$ to S_1, $\mathtt{F}$ to H_1, and μ_1 to a_1. Next, the query $EQ(H_1)$ is made and the teacher responds with the valuation $\mu_2(b_0b_1) = 01$. Since $\mu_2 \not\models \mathtt{F}$, we have $I = \{1\}$. Algorithm 1 now walks from μ_2 towards a_1. Since flipping $\mu_2(b_1)$ would not give us a model of $b_0 \oplus b_1$, we have $S_1 = \{\mu_2\}$ and $H_1 = b_1$. In this example, Algorithm 1 generates $(b_1 \vee b_0) \wedge (\neg b_0 \vee \neg b_1)$ as a representation for the unknown Boolean formula $b_0 \oplus b_1$. Observe that the generated Boolean formula is a conjunction of two Boolean formulae in disjunctive normal form.

equivalence query	answer	I	S_i	H_i	a_i
$\mathtt{T}$	$\mu_1(b_0b_1) = 00$		$S_1 = \emptyset$	$H_1 = \mathtt{F}$	$a_1 = \mu_1$
$\mathtt{F}$	$\mu_2(b_0b_1) = 01$	$\{1\}$	$S_1 = \{\mu_2\}$	$H_1 = b_1$	
b_1	$\mu_3(b_0b_1) = 11$	$\emptyset$	$S_2 = \emptyset$	$H_2 = \mathtt{F}$	$a_2 = \mu_3$
$b_1 \wedge \mathtt{F}$	$\mu_4(b_0b_1) = 01$	$\{2\}$	$S_2 = \{\mu_5\}^\dagger$	$H_2 = \neg b_0$	
$b_1 \wedge \neg b_0$	$\mu_6(b_0b_1) = 10$	$\{1, 2\}$	$S_1 = \{\mu_2, \mu_6\}$ $S_2 = \{\mu_5, \mu_7\}^\dagger$	$H_1 = b_1 \vee b_0$ $H_2 = \neg b_0 \vee \neg b_1$	
$(b_1 \vee b_0) \wedge (\neg b_0 \vee \neg b_1)$	*YES*				

† $\mu_5(b_0b_1) = 10$ and $\mu_7(b_0b_1) = 01$

Fig. 4. Learning $b_0 \oplus b_1$