

A Forward-Backward Abstraction Refinement Algorithm

Francesco Ranzato, Olivia Rossi Doria, and Francesco Tapparo

Dipartimento di Matematica Pura ed Applicata
Università di Padova, Italy

Abstract. Abstraction refinement-based model checking has become a standard approach for efficiently verifying safety properties of hardware/software systems. Abstraction refinement algorithms can be guided by counterexamples generated from abstract transition systems or by fixpoints computed in abstract domains. Cousot, Ganty and Raskin recently put forward a new fixpoint-guided abstraction refinement algorithm that is based on standard abstract interpretation and improves the state-of-the-art, also for counterexample-driven methods. This work presents a new fixpoint-guided abstraction refinement algorithm that enhances the Cousot-Ganty-Raskin’s procedure. Our algorithm is based on three main ideas: (1) within each abstraction refinement step, we perform multiple forward-backward abstract state space traversals; (2) our abstraction is a disjunctive abstract domain that is used both as an overapproximation and an underapproximation; (3) we maintain and iteratively refine an overapproximation M of the set of states that belong to some minimal (i.e. shortest) counterexample to the given safety property so that each abstract state space traversal is limited to the states in M .

1 Introduction

Abstraction techniques are widely used in model checking to blur some properties of the concrete model and then to design a reduced abstract model where to run the verification algorithm [3]. Abstraction provides a successful solution to the state-explosion problem that arises in model checking systems with parallel components [4]. CounterExample-Guided Abstraction Refinement (CEGAR), pioneered by Clarke et al. [5], is become the standard methodology for applying abstraction to model checking. The basic idea of the CEGAR approach is as follows: if the abstract model checker return “YES” then the system satisfies the property; otherwise the abstract model checker returns an abstract counterexample to the property, that is checked to determine whether it corresponds to a real counterexample or not; if it does then return “NO” otherwise refine the abstract model in order to remove that spurious counterexample. Many different algorithms that implement the CEGAR approach have been suggested. Most CEGAR-based model checkers — like BLAST [16, 17], MAGIC [5, 2] and SLAM [1] — deal with counterexamples that are paths of abstract states, i.e. paths in an abstract transition system defined by an abstract state space and an abstract transition relation. Most often, model checkers aim at verifying so-called safety properties, i.e., states that can be reached from an initial state are always safe. Hence, safety verification consists in

automatically proving that systems cannot go wrong.

Recently, Cousot, Ganty and Raskin [10] (more details are given by the PhD thesis [12]) put forward a new fixpoint-guided abstraction refinement algorithm, here called CGR, for checking safety properties. The CGR algorithm is based on a number of interesting features. (1) CGR maintains and refines generic abstract domains that are defined within the standard abstract interpretation framework, as opposed to most other CEGAR-based algorithms that consider as abstract models a partition of the state space. (2) The refinement of the current abstract domain A is driven by the abstract fixpoint computed within A and not by a path-based counterexample. (3) CGR computes overapproximations of both least and greatest fixpoints, and these two analyses are made iteratively synergic since the current abstract fixpoint computation is limited by the abstract value provided by the previous abstract fixpoint computation.

We isolated a number of examples where the behavior of the CGR algorithm could be improved, in particular where one could abstractly conclude that the system is safe or unsafe without resorting to abstraction refinements. This work puts forward a new fixpoint-guided abstraction refinement algorithm for safety verification, called FBAR (Forward-Backward Abstraction Refinement), that is designed as an enhancement of the CGR procedure that integrates some new ideas.

- (i) FBAR maintains and refines a disjunctive abstract domain μ that overapproximates any set S of states by $\mu(S) \supseteq S$. While a generic abstract domain can be viewed as a set of subsets of states that is closed under arbitrary intersections, a disjunctive abstract domain must also be closed under arbitrary unions. The advantage of dealing with a disjunctive abstraction μ is given by the fact that μ can be simultaneously used both as an over- and under-approximating abstraction. As an additional advantage, it turns out that disjunctive abstractions can be efficiently represented and refined, as shown in [20].
- (ii) FBAR computes and maintains an overapproximation M of the set of states that occur in some *minimal* safety counterexample. A safety counterexample is simply a path from an initial state to an unsafe state. However, counterexamples may be redundant, namely may contain shorter sub-counterexamples. A counterexample is thus called minimal when it cannot be reduced. It can be therefore helpful to focus on minimal counterexamples rather than on generic counterexamples. In FBAR, abstract fixpoints are always computed within the overapproximation M and other than being used for safety checking they are also used for refining M .
- (iii) Each abstraction refinement step in FBAR consists of two loops that check whether the system can be proved safe/unsafe by using the current abstraction. The safety loop is based on a combined forward-backward abstract exploration of the portion of the state space limited by M . This combined forward-backward abstract computation was first introduced by Cousot [6]. The unsafety loop relies on an iterated combination of two abstract fixpoints: the first one is an overapproximation of the states in M that are globally safe and is computed by using the current abstraction μ as an overapproximation; the second one is instead an underapproximation of the states in M that can reach an unsafe state and is computed by viewing μ as an underapproximating abstraction.

We prove that FBAR is a correct algorithm for safety verification and that, analogously to CGR, it terminates when the concrete domain satisfies the descending chain condition. We formally compare FBAR and CGR by showing that FBAR improves CGR in the following sense: if CGR terminates on a given disjunctive abstraction μ with no refinement then FBAR also terminates on μ , while the converse is not true.

Related Work. We discussed above the relationship with the CGR algorithm in [10]. Gulavani and Rajamani [15] also describe a fixpoint-driven abstraction refinement algorithm for safety verification. This algorithm relies on using widening operators during abstract fixpoint computations. When the abstract fixpoint is inconclusive, this algorithm does not refine the abstract domain but determines which iteration of the abstract fixpoint computation was responsible of the loss of precision so that widening is replaced with a concrete union and the abstract computation is re-started from that iteration. Manevich et al. [18] put forward an abstraction refinement algorithm for safety verification that runs over disjunctive abstract domains. However, this algorithm does not compute abstract fixpoints but instead computes paths of abstract values so that the abstraction refinement is based on counterexamples defined as sequences of abstract values.

2 Background

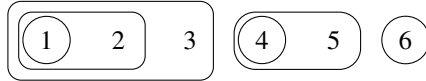
Notation and Orders. Let Σ be any set. If $S \subseteq \Sigma$ then $\neg S$ denotes the complement set $\Sigma \setminus S$ when Σ is clear from the context. A set S of one-digit integers is often written in a compact form without brackets and commas like $S = 1357$ that stands for $\{1, 3, 5, 7\}$. $\text{Part}(\Sigma)$ denotes the set of partitions of Σ . If $R \subseteq \Sigma \times \Sigma$ is any relation then $R^* \subseteq \Sigma \times \Sigma$ denotes the reflexive and transitive closure of R . Posets and complete lattices are denoted by P_{\leq} where \leq is the partial order. A function f between complete lattices is additive (co-additive) when f preserves least upper (greatest lower) bounds. If $f : P \rightarrow P$ then $\text{lfp}(f)$ and $\text{gfp}(f)$ denote, resp., the least and greatest fixpoints of f , when they exist.

Abstract Domains. In standard Cousot and Cousot's abstract interpretation, abstract domains (or abstractions) can be equivalently specified either by Galois connections/insertions through α/γ abstraction/concretization maps or by upper closure operators (uco's) [7]. These two approaches are equivalent, modulo isomorphic representations of domain's objects. The closure operator approach has the advantage of being independent from the representation of domain's objects and is therefore appropriate for reasoning on abstract domains independently from their representation. Given a state space Σ , the complete lattice $\wp(\Sigma)_{\subseteq}$, i.e. the powerset of Σ ordered by the subset relation, plays the role of concrete domain. Let us recall that an operator $\mu : \wp(\Sigma) \rightarrow \wp(\Sigma)$ is a uco on $\wp(\Sigma)$, that is an overapproximating abstract domain of $\wp(\Sigma)$, when μ is monotone, idempotent and extensive (i.e., overapproximating: $X \subseteq \mu(X)$). Each closure μ is uniquely determined by its image $\text{img}(\mu) = \{\mu(X) \in \wp(\Sigma) \mid X \in \wp(\Sigma)\}$ as follows: for any $X \subseteq \Sigma$, $\mu(X) = \bigcap \{Y \in \text{img}(\mu) \mid X \subseteq Y\}$. On the other hand, a set of subsets $A \subseteq \wp(\Sigma)$ is the image of some closure on $\wp(\Sigma)$ iff A is closed under arbitrary intersections, i.e. $A = \text{Cl}_{\cap}(A) \stackrel{\text{def}}{=} \{\bigcap S \mid S \subseteq A\}$ (in particular, note that $\text{Cl}_{\cap}(A)$ always

contains $\Sigma = \cap \emptyset$). This makes clear that an abstract domain μ guarantees that for any concrete set of states X , $\mu(X)$ is the best (i.e., more precise) overapproximation of X in μ . We denote by $\text{Abs}(\Sigma)$ the set of abstract domains of $\wp(\Sigma)_{\subseteq}$. Capital letters like $A, A' \in \text{Abs}(\Sigma)$ are sometimes used for denoting abstract domains. By a slight abuse of notation, a given abstract domain $A \in \text{Abs}(\Sigma)$ can be viewed and used both as a set of subsets of Σ and as an operator on $\wp(\Sigma)$ when the context allows us to disambiguate this use. If $A_1, A_2 \in \text{Abs}(\Sigma)$ then A_1 is more precise than (or is a refinement of) A_2 when $A_1 \supseteq A_2$. $\text{Abs}(\Sigma)_{\supseteq}$ is called the (complete) lattice of abstract domains of $\wp(\Sigma)$.

Let $f : \wp(\Sigma) \rightarrow \wp(\Sigma)$ be a concrete semantic function, like a predicate transformer, and, given an abstraction $\mu \in \text{Abs}(\Sigma)$, let $f^\# : \mu \rightarrow \mu$ be a corresponding abstract function on μ . Then, $f^\#$ is a correct approximation of f in μ when for any $X \in \wp(\Sigma)$, $\mu(f(X)) \subseteq f^\#(\mu(X))$. The abstract function $f^\mu : \mu \rightarrow \mu$ defined as $f^\mu \stackrel{\text{def}}{=} \mu \circ f$ is called the best correct approximation of f in μ because for any correct approximation $f^\#$, for any $X \in \mu$, $f^\mu(X) \subseteq f^\#(X)$ always holds.

Disjunctive Abstract Domains. An abstract domain $\mu \in \text{Abs}(\wp(\Sigma))$ is disjunctive (or additive or a powerset abstract domain) when μ is additive, i.e. μ preserves arbitrary unions. This happens exactly when the image $\text{img}(\mu)$ is closed under arbitrary unions, i.e., $\mu = \text{Cl}_{\cup}(\mu) \stackrel{\text{def}}{=} \{\cup S \mid S \subseteq \mu\}$ (in particular, note that $\text{Cl}_{\cup}(\mu)$ always contains $\emptyset = \cup \emptyset$). Hence, a disjunctive abstract domain is a set of subsets of states that is closed under both arbitrary intersections and unions. The intuition is that a disjunctive abstract domain does not loose precision in approximating concrete set unions. We denote by $\text{dAbs}(\wp(\Sigma)) \subseteq \text{Abs}(\wp(\Sigma))$ the set of disjunctive abstract domains. A disjunctive abstraction μ can be specified just by defining how any singleton $\{x\} \subseteq \Sigma$ is approximated by $\mu(\{x\})$, because the approximation of a generic subset $X \subseteq \Sigma$ can be obtained through set unions as $\mu(X) = \cup_{x \in X} \mu(\{x\})$. We exploit this property for representing disjunctive abstractions through diagrams. As an example, the following diagram:



denotes the disjunctive abstract domain μ that is determined by the following behaviour on singletons: $\mu(1) = 1$, $\mu(2) = 12$, $\mu(3) = 123$, $\mu(4) = 4$, $\mu(5) = 45$, $\mu(6) = 6$, so that μ is the closure under unions of the set $\{1, 12, 123, 4, 45, 6\}$.

Underapproximating Abstract Domains. It turns out that a disjunctive abstract domain $\mu \in \text{Abs}(\wp(\Sigma)_{\subseteq})$ can be also viewed as an underapproximating abstract domain, namely an abstraction of the concrete domain $\wp(\Sigma)_{\supseteq}$ where the approximation order is reversed. Formally, this is specified by the closure $\tilde{\mu} \in \text{Abs}(\wp(\Sigma)_{\supseteq})$ that is defined as the adjoint of μ as follows: for any $X \subseteq \Sigma$, $\tilde{\mu}(X) \stackrel{\text{def}}{=} \cup \{Y \subseteq \Sigma \mid Y = \mu(Y) \subseteq X\}$. An underapproximating abstraction is thus determined by the behaviour on the sets $\neg x \stackrel{\text{def}}{=} \Sigma \setminus \{x\}$ because, for any $X \subseteq \Sigma$, $\tilde{\mu}(X) = \cap_{x \notin X} \tilde{\mu}(\neg x)$. For example, for the above disjunctive abstraction μ , we have that $\tilde{\mu}(\neg 1) = 456$, $\tilde{\mu}(\neg 2) = 1456$, $\tilde{\mu}(\neg 3) = 12456$, $\tilde{\mu}(\neg 4) = 1236$, $\tilde{\mu}(\neg 5) = 12346$ and $\tilde{\mu}(\neg 6) = 12345$.

Transition Systems. A transition system $\mathcal{T} = (\Sigma, R)$ consists of a set Σ of states and a transition relation $R \subseteq \Sigma \times \Sigma$, that is also denoted in infix notation by \rightarrow . As usual in model checking, we assume that the relation R is total, i.e., for any $s \in \Sigma$ there exists some $t \in \Sigma$ such that $s \rightarrow t$. The set of finite and infinite paths in \mathcal{T} is denoted by $\text{Path}(\mathcal{T})$. For any $\Pi \subseteq \text{Path}(\mathcal{T})$, $\text{states}(\Pi) \subseteq \Sigma$ denotes the set of states that occur in some path $\pi \in \Pi$. If $\pi \in \text{Path}(\mathcal{T})$ is finite then $\text{first}(\pi), \text{last}(\pi) \in \Sigma$ denote, resp., the first and last states of π .

Standard predicate transformers $\text{pre}, \widetilde{\text{pre}}, \text{post}, \widetilde{\text{post}} : \wp(\Sigma) \rightarrow \wp(\Sigma)$ are defined as usual:

- $\text{pre}(X) \stackrel{\text{def}}{=} \{a \in \Sigma \mid \exists b \in X. a \rightarrow b\}$;
- $\text{post}(X) \stackrel{\text{def}}{=} \{b \in \Sigma \mid \exists a \in X. a \rightarrow b\}$;
- $\widetilde{\text{pre}}(X) \stackrel{\text{def}}{=} \neg \text{pre}(\neg X) = \{a \in \Sigma \mid \forall b. a \rightarrow b \Rightarrow b \in X\}$;
- $\widetilde{\text{post}}(X) \stackrel{\text{def}}{=} \neg \text{post}(\neg X) = \{b \in \Sigma \mid \forall a. a \rightarrow b \Rightarrow a \in X\}$.

Let us remark that pre and post are additive while $\widetilde{\text{pre}}$ and $\widetilde{\text{post}}$ are co-additive. We use the notation $\text{pre}^*, \widetilde{\text{pre}}^*, \text{post}^*, \widetilde{\text{post}}^*$ when the reflexive-transitive closure R^* is considered instead of R . Let us recall the following standard fixpoint characterizations:

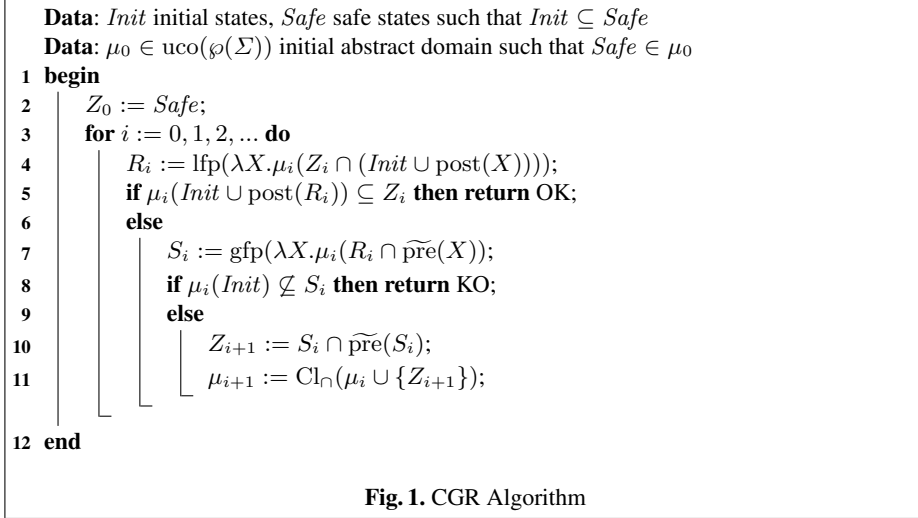
$$\begin{aligned} \text{pre}^*(X) &= \text{lfp}(\lambda Z. X \cup \text{pre}(Z)); & \text{post}^*(X) &= \text{lfp}(\lambda Z. X \cup \text{post}(Z)); \\ \widetilde{\text{pre}}^*(X) &= \text{gfp}(\lambda Z. X \cap \widetilde{\text{pre}}(Z)); & \widetilde{\text{post}}^*(X) &= \text{gfp}(\lambda Z. X \cap \widetilde{\text{post}}(Z)). \end{aligned}$$

Safety Verification Problems. Let $\text{Init} \subseteq \Sigma$ be a set of initial states and $\text{Safe} \subseteq \Sigma$ a set of safe states. We denote by $\text{NInit} \stackrel{\text{def}}{=} \neg \text{Init}$ the set of noninitial states and by $\text{Bad} = \neg \text{Safe}$ the set of bad (i.e. unsafe) states. The set of reachable states is $\text{post}^*(\text{Init})$. The set of states that are globally safe is $\widetilde{\text{pre}}^*(\text{Safe})$. The set of states that can reach a bad state is $\text{pre}^*(\text{Bad})$. The set of states that can be reached only from noninitial states is $\widetilde{\text{post}}^*(\text{NInit})$. Note that $\text{pre}^*(\text{Bad}) = \neg \widetilde{\text{pre}}^*(\text{Safe})$ and $\widetilde{\text{post}}^*(\text{NInit}) = \neg \text{post}^*(\text{Init})$.

A system \mathcal{T} is safe when any reachable state is safe, i.e. $\text{post}^*(\text{Init}) \subseteq \text{Safe}$, or, equivalently, when one of the following equivalent conditions holds: $\text{Init} \subseteq \widetilde{\text{pre}}^*(\text{Safe}) \Leftrightarrow \text{pre}^*(\text{Bad}) \subseteq \text{NInit} \Leftrightarrow \text{Bad} \subseteq \widetilde{\text{post}}^*(\text{NInit})$. A safety verification problem is then specified by a transition system $\mathcal{T} = \langle \Sigma, R, \text{Init}, \text{Safe} \rangle$ that also defines initial and safe states and consists in checking whether \mathcal{T} is safe (OK) or not (KO).

3 Cousot-Ganty-Raskin's Algorithm

The Cousot-Ganty-Raskin's algorithm, here denoted by CGR, is recalled in Fig. 1. In each abstraction refinement step $i \geq 0$, CGR abstractly computes two overapproximations R_i and S_i of least/greatest fixpoints and a concrete value Z_{i+1} that is added to the current abstract domain μ_i for the purpose of refining it. The correctness of CGR follows from the following three main invariants: for all $i \geq 0$: (1) $Z_{i+1} \subseteq S_i \subseteq R_i \subseteq Z_i$; (2) if the system is safe then $\text{post}^*(\text{Init}) \subseteq R_i$, i.e. R_i overapproximates the reachable states; (3) $R_i \subseteq \widetilde{\text{pre}}^i(\text{Safe})$, i.e. R_i underapproximates the states that remain inside Safe along paths of length $\leq i$.

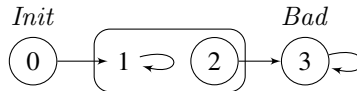


CGR admits a dual version, denoted by CGR^{\leftarrow} , where the transition relation is reversed, namely where R is replaced by R^{-1} , $Init$ by Bad and $Safe$ by $NInit$ (so that post and $\widetilde{\text{pre}}$ become, respectively, pre and $\widetilde{\text{post}}$). Thus, while CGR performs a forward abstract exploration of the state space through post , CGR^{\leftarrow} proceeds instead backward through pre .

3.1 Where CGR Could Be Improved

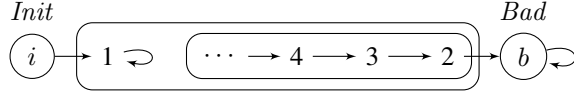
We isolated a number of examples where the CGR algorithm could be improved.

Example 3.1. Let us consider the safety problem represented by the following diagram that also specifies a disjunctive abstract domain $\mu_0 \in \text{Abs}(\wp(\Sigma))$.



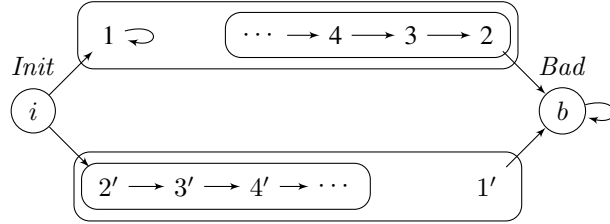
CGR computes the following sequence: $Z_0 = 012$, $R_0 = 012$, $S_0 = 012$, $Z_1 = 01$, $R_1 = 01$ and then outputs OK. Let us observe that $S_0 = 012$ because $\mu_0(R_0 \cap \widetilde{\text{pre}}(R_0)) = \mu_0(01) = 012$. Thus, CGR needs to refine the abstraction μ_0 by adding the singleton $\{1\}$ to μ_0 . However, one could abstractly conclude that the system is safe already through the abstraction μ_0 . In fact, one could abstractly explore backward the state space by computing the following fixpoint: $T_0 = \text{lfp}(\lambda X. \mu_0(NInit \cap (Bad \cup \text{pre}(X))))$. Thus, $T_0 = 23$ and since $\mu_0(Bad \cup \text{pre}(T_0)) \subseteq T_0$ one can conclude that the system is safe. Thus, the dual algorithm CGR^{\leftarrow} is able to conclude that the system is safe with no abstraction refinement.

Along the same lines, it turns out that CGR even does not terminate when applied to the following infinite state system, although the abstraction is finite, while a backward abstract exploration would allow to conclude that the system is safe.



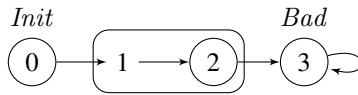
In fact, CGR does not terminate because it would compute the following infinite sequence: $Z_0 = i \cup 12345\dots$, $R_0 = i \cup 12345\dots$, $S_0 = i \cup 12345\dots$, $Z_1 = i \cup 1345\dots$, $R_1 = i \cup 1345\dots$, $S_1 = i \cup 1345\dots$, $Z_2 = i \cup 145\dots$. Instead, one could proceed backward by computing the following abstract fixpoint: $T_0 = \text{lfp}(\lambda X. \mu_0(N\text{Init} \cap (\text{Bad} \cup \text{pre}(X)))) = b \cup 2345\dots$. Hence, since $\mu_0(\text{Bad} \cup \text{pre}(T_0)) \subseteq T_0$ we can conclude that the system is safe. Thus, here again, CGR^\leftarrow is able to infer that the system is safe with no abstraction refinement.

Let us consider now the following infinite state system.



In this case, it turns out that neither CGR nor CGR^\leftarrow terminate. In fact, similarly to the above examples, it is simple to check that both CGR and CGR^\leftarrow would compute infinite sequences of abstract values. However, it is still possible to derive that the system is safe with no abstraction refinement. In fact, we can first compute the following forward abstract fixpoint $U_0 = \text{lfp}(\lambda X. \mu_0(\text{Safe} \cap (\text{Init} \cup \text{post}(X)))) = i \cup 1234\dots \cup 2'3'4'\dots$. Then, we can explore backward starting from *Bad* but remaining inside $U_0 \cup \text{Bad}$, namely we compute the following backward abstract fixpoint $V_0 = \text{lfp}(\lambda X. \mu_0((U_0 \cup \text{Bad}) \cap (\text{Bad} \cup \text{pre}(X)))) = b \cup 234\dots$. We can now conclude that the system is safe because $\mu_0(\text{Bad} \cup \text{pre}(V_0)) \subseteq V_0$. \square

Example 3.2. Let us consider the following safety problem and disjunctive abstract domain $\mu_0 \in \text{Abs}(\wp(\Sigma))$.



In this case, CGR computes the following sequence: $Z_0 = 012$, $R_0 = 012$, $S_0 = 012$, $Z_1 = 01$, $R_1 = 01$, $S_1 = \emptyset$ and then outputs KO. Thus, CGR needs to refine the abstraction μ_0 by adding the singleton $\{1\}$ to μ_0 . However, one could abstractly conclude that the system is not safe already through the abstraction μ_0 by viewing μ_0 as an underapproximating abstraction, i.e. by considering the underapproximating abstraction $\tilde{\mu}_0$. In fact, one could abstractly explore the state space backward by computing the following abstract fixpoint: $T_0 = \text{lfp}(\lambda X. \tilde{\mu}_0(\text{Bad} \cup \text{pre}(X))) = 0123$. Since T_0 is an underapproximation of the set of states that can reach a bad state and T_0 contains some initial state we can conclude that the system is unsafe. \square

These examples suggested us to design an abstraction refinement algorithm that improves the CGR algorithm by integrating a combined forward-backward abstract exploration of the state space and by using disjunctive abstract domains that can be exploited both as overapproximating and underapproximating abstractions.

4 Restricted Predicate Transformers

Let $M \subseteq \Sigma$ be a fixed set of states of interest. In our context, M will play the role of a portion of the state space that limits the abstract search space of our abstraction refinement algorithm. Let us define the following *restricted* (to the states in M) predicate transformers $\text{pre}_M, \text{post}_M, \widetilde{\text{pre}}_M, \widetilde{\text{post}}_M : \wp(\Sigma) \rightarrow \wp(\Sigma)$ as follows:

$$\begin{aligned}
- \text{pre}_M(X) &\stackrel{\text{def}}{=} M \cap \text{pre}(M \cap X); \\
- \text{post}_M(X) &\stackrel{\text{def}}{=} M \cap \text{post}(M \cap X); \\
- \widetilde{\text{pre}}_M(X) &\stackrel{\text{def}}{=} \neg \text{pre}_M(\neg X) = \neg M \cup \widetilde{\text{pre}}(\neg M \cup X) \\
&= \{a \in \Sigma \mid \forall b. (a \rightarrow b \ \& \ a, b \in M) \Rightarrow b \in X\}; \\
- \widetilde{\text{post}}_M(X) &\stackrel{\text{def}}{=} \neg \text{post}_M(\neg X) = \neg M \cup \widetilde{\text{post}}(\neg M \cup X) \\
&= \{b \in \Sigma \mid \forall a. (a \rightarrow b \ \& \ a, b \in M) \Rightarrow a \in X\}.
\end{aligned}$$

Thus, M -restricted predicate transformers only consider states that belong to M . In fact, if $\text{pre}_M, \text{post}_M, \widetilde{\text{pre}}_M, \widetilde{\text{post}}_M$ are viewed as mappings $\wp(M) \rightarrow \wp(M)$ — i.e., both the argument and the image of the M -restricted transformers are taken as subsets of M — then they coincide with the corresponding standard predicate transformers on the M -restricted transition system $\mathcal{T}_M = \langle M, R/M \rangle$. Let us remark that, analogously to the unrestricted case, $\text{pre}_M, \text{post}_M$ are additive functions and $\widetilde{\text{pre}}_M, \widetilde{\text{post}}_M$ are co-additive functions. We also consider the following fixpoint definitions:

$$\begin{aligned}
- \text{pre}_M^*(X) &\stackrel{\text{def}}{=} \text{lfp}(\lambda Z. (X \cap M) \cup \text{pre}_M(Z)) \\
&= \{x \in \Sigma \mid \exists y \in X. x \rightarrow^* y \ \& \ \text{states}(x \rightarrow^* y) \subseteq M\}; \\
- \text{post}_M^*(X) &\stackrel{\text{def}}{=} \text{lfp}(\lambda Z. (X \cap M) \cup \text{post}_M(Z)) \\
&= \{y \in \Sigma \mid \exists x \in X. x \rightarrow^* y \ \& \ \text{states}(x \rightarrow^* y) \subseteq M\}; \\
- \widetilde{\text{pre}}_M^*(X) &\stackrel{\text{def}}{=} \text{gfp}(\lambda Z. (X \cup \neg M) \cap \widetilde{\text{pre}}_M(Z)) \\
&= \{x \in \Sigma \mid \forall y. (x \rightarrow^* y \ \& \ \text{states}(x \rightarrow^* y) \subseteq M) \Rightarrow y \in X\}; \\
- \widetilde{\text{post}}_M^*(X) &\stackrel{\text{def}}{=} \text{gfp}(\lambda Z. (X \cup \neg M) \cap \widetilde{\text{post}}_M(Z)) \\
&= \{y \in \Sigma \mid \forall x. (x \rightarrow^* y \ \& \ \text{states}(x \rightarrow^* y) \subseteq M) \Rightarrow x \in X\}.
\end{aligned}$$

Hence, we have that $x \in \text{pre}_M^*(X)$ iff x may reach X through a path inside M , while $x \in \widetilde{\text{pre}}_M^*(X)$ iff x inside M can only reach states in X . Let us note that, analogously to the unrestricted case, $\widetilde{\text{pre}}_M^*(\neg X) = \neg \text{pre}_M^*(X)$ and $\widetilde{\text{post}}_M^*(\neg X) = \neg \text{post}_M^*(X)$. Moreover, $\text{pre}_M^*(X) \subseteq \text{pre}^*(X)$ and $\text{post}_M^*(X) \subseteq \text{post}^*(X)$ while $\widetilde{\text{pre}}_M^*(X) \subseteq \widetilde{\text{pre}}^*(X)$ and $\text{post}^*(X) \subseteq \text{post}_M^*(X)$.

Example 4.1. Consider the safety verification problem depicted in Fig. 2 where the gray states determine the restricted space $M = 0134568$. It turns out that $\text{post}_M^*(\text{Init}) =$

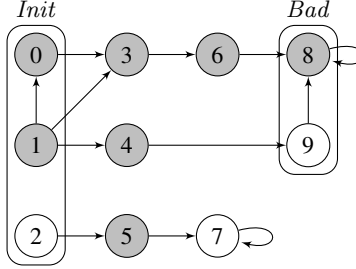


Fig. 2. Restricted predicated transformers.

013468, $\text{pre}_M^*(Bad) = 01368$, $\widetilde{\text{post}}_M^*(NInit) = 2579$ and $\widetilde{\text{pre}}_M^*(Safe) = 24579$. Note, for example, that $9 \in \text{post}^*(Init) = 0134689$ but $9 \notin \text{post}_M^*(Init)$. Also, $4 \in \widetilde{\text{pre}}_M^*(Safe)$ but $4 \notin \widetilde{\text{pre}}^*(Safe) = 257$ because there is no path that begins with 4 and remains inside M . \square

Thus, when using a M -restricted predicate transformer instead of a standard (i.e. unrestricted) predicate transformer it is enough to consider only the states belonging to M . It should be clear that when M is much smaller than the whole state space Σ such a restriction to states in M may induce space and time savings.

5 Minimal Counterexamples

One main idea of our abstraction refinement algorithm consists in overapproximating the set of states that belong to some safety counterexample, i.e. a finite path from an initial state to a bad state. However, a counterexample π may be redundant, namely π might contain a shorter sub-path that still is a safety counterexample. For example, in the transition system in Fig. 2, the path $\pi = 103688$ is a safety counterexample because it begins with an initial state and ends with a bad state although π is redundant because it contains a sub-path $\pi' = 0368$ that is a counterexample. Our algorithm will compute and maintain an overapproximation of the states that belong to counterexamples that are not reducible. Such counterexamples are called minimal counterexamples.

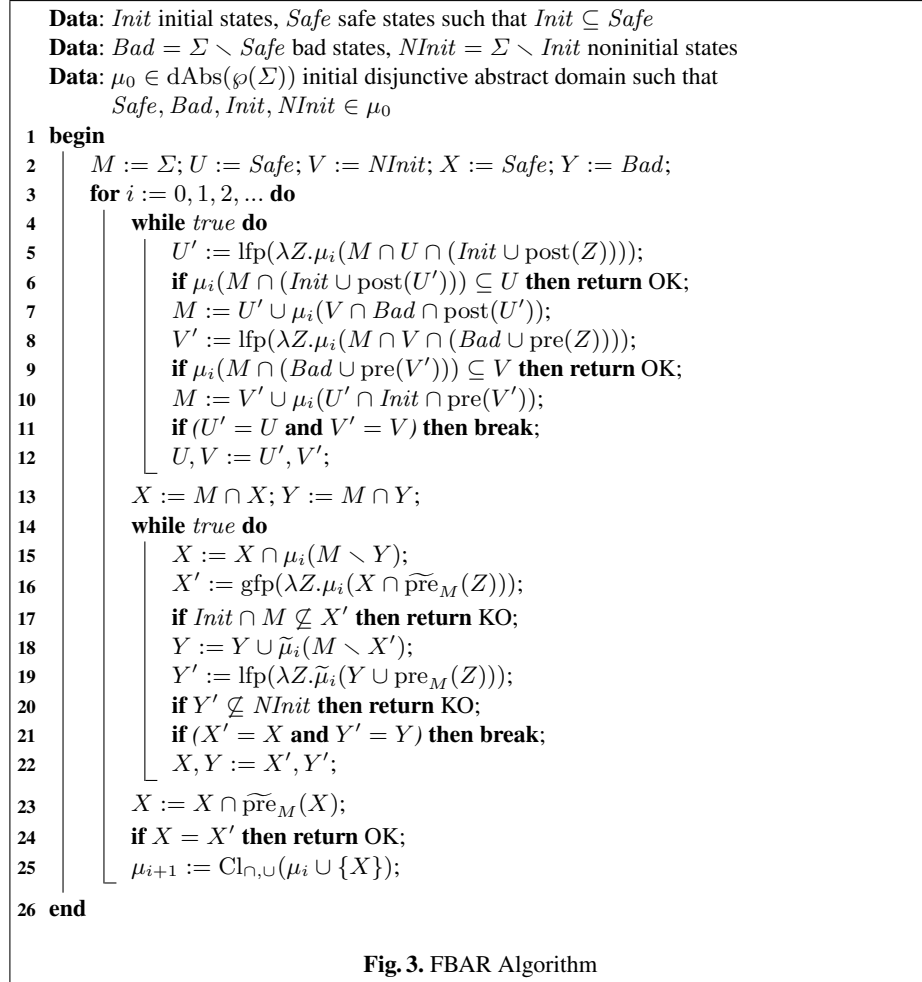
Let us formalize the above notions. Let $\mathcal{T} = \langle \Sigma, R, Init, Safe \rangle$ specify a safety problem. A (safety) counterexample is a finite path $\pi \in \text{Path}(\mathcal{T})$ such that $\text{first}(\pi) \in Init$ and $\text{last}(\pi) \in Bad$. A *minimal counterexample* is a counterexample $\pi \in \text{Path}(\mathcal{T})$ such that $\text{states}(\pi) \setminus \{\text{first}(\pi), \text{last}(\pi)\} \subseteq Safe \cap NInit$. We define $\text{MinCex} \stackrel{\text{def}}{=} \{\pi \in \text{Path}(\mathcal{T}) \mid \pi \text{ is a minimal counterexample}\}$.

Assume that M is an overapproximation of the states that occur in some minimal counterexample, i.e. $\text{states}(\text{MinCex}) \subseteq M$. Then, we provide a characterization of safe systems that only considers states in M : it turns out that a system is safe iff any state that is reachable from an initial state through a path inside M is safe.

Lemma 5.1. *If $\text{states}(\text{MinCex}) \subseteq M$ and $Init \subseteq Safe$ then the system \mathcal{T} is safe iff $\text{post}_M^*(Init) \subseteq M \cap Safe$ iff $\text{pre}_M^*(Bad) \subseteq M \cap NInit$.*

6 A Forward-Backward Abstraction Refinement Algorithm

The Forward-Backward Abstraction Refinement algorithm FBAR is defined in Fig. 3. FBAR takes as input a safety verification problem $\mathcal{T} = \langle \Sigma, R, Init, Safe \rangle$ and a disjunctive abstraction $\mu_0 \in \text{dAbs}(\wp(\Sigma))$. The main ideas and features of FBAR are summarized in the following list.



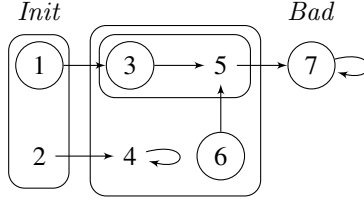
- (A) The loop at lines 4-12 computes and maintains an overapproximation M of the states that occur in some minimal counterexample by relying on a combined forward-backward abstract exploration of the state space. Such a combined forward-backward abstract computation was first described by Cousot [6] and further investigated and applied in [8, 9, 19]. The following procedure is iterated: we start from

$M \cap Init$ and abstractly go forward within M through post_M (line 5); then, we come back by starting from $M \cap Bad$ and abstractly going backward within M through pre_M (line 8). The abstract sets U and V are the results of these, resp., forward and backward abstract fixpoint computations. The following invariant properties hold (cf. Lemma 6.2 (1)): U is an overapproximation of the safe states that can be reached through a path within M , while V is an overapproximation of the noninitial states that can reach a bad state through a path within M . The combined forward-backward computation of U and V allows us to iteratively refine the overapproximation M of states (MinCex) (lines 7 and 10).

- (B) The OK condition at line 6 implies that $\text{post}_M^*(Init) \subseteq U$, so that $\text{post}_M^*(Init) \subseteq M \cap Safe$, and therefore, by Lemma 5.1, the system is safe. Analogously, for the “backward” OK condition at line 9.
- (C) The loop at lines 14-22 computes iteratively the abstract sets X and Y as follows. X is an overapproximation of the states in M that are globally safe and is computed at line 16 as a greatest fixpoint of the best correct approximation in μ_i of $\widetilde{\text{pre}}_M$. On the other hand, Y is an underapproximation of the states in M that can reach a bad state and is computed at line 19 as a least fixpoint of the best correct approximation of pre_M w.r.t. the underapproximating abstraction $\tilde{\mu}_i$. This is formally stated by Lemma 6.2 (3). While the sequence of computed X 's forms a descending chain of abstract sets, the Y 's give rise to an ascending chain of abstract sets. These abstract computations are iterated because the abstract set Y may help in refining X and, vice versa, X may help in refining Y . In fact, observe that the states in $M \setminus Y$ form a superset of the states in M that are globally safe, so that the overapproximation X can be refined by intersection with the abstract set $\mu_i(M \setminus Y)$. A dual reasoning holds for Y , where we exploit the fact that μ_i is a disjunctive abstraction and therefore $\tilde{\mu}_i$ is an underapproximating abstraction.
- (D) Since $X' \supseteq M \cap \widetilde{\text{pre}}^*(Safe)$, the KO condition at line 17 implies that $Init \cap M \not\subseteq M \cap \widetilde{\text{pre}}^*(Safe)$, namely that $Init \not\subseteq \widetilde{\text{pre}}^*(Safe)$, so that the system is unsafe. Analogously, for the “underapproximating” KO condition at line 20.
- (E) At the exit of the loop at lines 14-22, we perform a concrete step of computation at line 23 by calculating a refiner set $Ref = X \cap \widetilde{\text{pre}}_M(X)$. In contrast to the algorithm CGR where the refiner Z_{i+1} cannot already be in the abstraction μ_i , here it may happen that $X = Ref$. In this case (line 24), we obtain that $\text{post}_M^*(Init) \subseteq X$ and this allows us to conclude that the system is safe. Otherwise, $Ref \subsetneq X$ is used to refine μ_i to μ_{i+1} that is obtained by closing $\mu_i \cup \{Ref\}$ both under intersections — in order to have an abstraction — and unions — in order to have a disjunctive abstraction.

Let us now illustrate how FBAR works on a simple finite example.

Example 6.1. Let us consider the safety verification problem represented by the following diagram, where $\mu_0 = \text{Cl}_\cup(\{1, 12, 3, 35, 3456, 6, 7\})$.



Then, FBAR allows to derive that the system is unsafe with no refinement. In fact, FBAR gives rise to the following “execution trace”:

[line 2]:	$M = 1234567; U = 123456; V = 34567; X = 123456; Y = 7;$
[line 5]:	$U' = 123456;$
[line 6]:	$1234567 = \mu_0(M \cap (Init \cup \text{post}(U'))) \not\subseteq U = 123456;$
[line 7]:	$M = 1234567;$
[line 8]:	$V' = 3567;$
[line 9]:	$13567 = \mu_0(M \cap (Bad \cup \text{pre}(V'))) \not\subseteq V = 34567;$
[line 10]:	$M = 13567;$
[line 11]:	$U' = U \ \& \ V' \neq V;$
[line 12]:	$U = 123456; V = 3567;$
[line 5]:	$U' = 135;$
[line 6]:	$1357 = \mu_0(M \cap (Init \cup \text{post}(U'))) \not\subseteq U = 123456;$
[line 7]:	$M = 1357;$
[line 8]:	$V' = 357;$
[line 9]:	$1357 = \mu_0(M \cap (Bad \cup \text{pre}(V'))) \not\subseteq V = 3567;$
[line 10]:	$M = 1357;$
[line 11]:	$U' \neq U \ \& \ V' \neq V;$
[line 12]:	$U = 135; V = 357;$
[lines 5-10]:	a further iteration that does not change U', V' and M
[line 11]:	$U' = U \ \& \ V' = V;$
[line 13]:	$X = 135; Y = 7;$
[line 15]:	$X = 135;$
[line 16]:	$X' = \emptyset;$
[line 16]:	$Init \cap M \not\subseteq X' \Rightarrow \text{KO}$

Thus, FBAR needs no abstraction refinement and seven abstract fixpoint computations. On the other hand, CGR needs three abstraction refinements and eight abstract fixpoint computations in order to conclude that the system is unsafe. In fact, it computes the following sequence: $Z_0 = 123456, R_0 = 123456, S_0 = 123456, Z_1 = 12346, R_1 = 12346, S_1 = 12346, Z_2 = 124, R_2 = 124, S_2 = 124, Z_3 = 24, R_3 = 24, S_3 = 24$ and then concludes KO.

It can be also checked that the dual algorithm CGR^- needs one abstraction refinement in order to conclude that the system is unsafe while in this case CGR^+ performs just four abstract fixpoint computations. \square

The above described properties of the FBAR procedure are stated precisely as follows.

Lemma 6.2. *The following properties are invariant in the algorithm FBAR in Fig. 3:*

- (1) $\text{post}_{M \cap \text{Safe}}^*(\text{Init}) \subseteq U' \subseteq U \subseteq M$ & $\text{pre}_{M \cap \text{Init}}^*(\text{Bad}) \subseteq V' \subseteq V \subseteq M$.
- (2) $\text{states}(\text{MinCex}) \subseteq M$.
- (3) $M \cap \widetilde{\text{pre}}^*(\text{Safe}) \subseteq X' \subseteq X \subseteq M$ & $Y \subseteq Y' \subseteq M \cap \text{pre}^*(\text{Bad})$.

These invariant properties allows us to show that FBAR is a correct algorithm for safety checking.

Theorem 6.3 (Correctness). *If FBAR outputs OK/KO then \mathcal{T} is safe/unsafe.*

6.1 Termination

Termination of FBAR is similar to that of CGR.

Theorem 6.4 (Termination).

- (1) *If μ_0 is finite and there exists $\mathcal{X} \subseteq \wp(\Sigma)$ such that for all $i \geq 0$, $X_i \in \mathcal{X}$ and $\langle \mathcal{X}, \subseteq \rangle$ satisfies the descending chain condition then FBAR terminates.*
- (2) *If \mathcal{T} is unsafe then FBAR terminates.*

Hence, if the refiner sets X_i 's at line 23 all belong to a subset of the state space that satisfies the descending chain condition then FBAR terminates. This obviously implies termination when the state space Σ is finite. Ganty et al. [13, 14] show that this descending chain condition allows to show that the instantiation of the CGR algorithm to the class of well-structured transition systems (WSTSs) always terminates. This is an important result because WSTSs are a broad and relevant class of infinite-state transition systems that include, among others, Petri Nets, broadcast protocols and lossy-channel systems [11]. Since this termination condition works for FBAR exactly in the same way as for CGR, we conjecture that the descending chain condition should allow to show that FBAR terminates on WSTSs.

6.2 Relationship with CGR

We made a formal comparison between the FBAR and CGR algorithms and showed that when no abstraction refinement is needed, FBAR is better than CGR, i.e., if CGR terminates with a given abstraction μ with no abstraction refinement then this also happens for FBAR. As shown by the examples in Section 3, the converse is not true.

Theorem 6.5. *If CGR for some disjunctive abstract domain μ outputs OK/KO with no abstraction refinement then FBAR for μ outputs OK/KO with no abstraction refinement.*

We did not succeed in comparing formally FBAR and CGR when abstraction refinements indeed happen. This does not appear to be an easy task mainly because FBAR and CGR use both different refiners — FBAR refines using $\widetilde{\text{pre}}_M$ while CGR uses $\widetilde{\text{pre}}$ — and different ways of refining the abstraction — FBAR needs a refined disjunctive abstraction while CGR needs a mere abstraction. We can only report that we were not able to find an example where CGR terminates while FBAR does not.

Let us also remark that Cousot-Ganty-Raskin [10] described how some acceleration techniques that compute underapproximations of the reflexive-transitive closure R^* of the transition relation R can be integrated into CGR. The correctness of this technique basically depends on the fact that replacing the transition relation R with a relation T such that $R \subseteq T \subseteq R^*$ is still correct in CGR. This also holds for FBAR so that these same techniques can be applied.

6.3 Implementation of Disjunctive Abstractions

An implementation of FBAR is subject for future work. However, let us mention that disjunctive abstractions can indeed be efficiently implemented through a state partition and a relation defined over it.

Let us recall from [20] the details of such a representation. Let $\mu \in \text{dAbs}(\wp(\Sigma))$. The partition $\text{par}(\mu) \in \text{Part}(\Sigma)$ induced by μ is defined by the following equivalence relation $\sim_\mu \subseteq \Sigma \times \Sigma$: $x \sim_\mu y$ iff $\mu(\{x\}) = \mu(\{y\})$. Moreover, let us define the following relation \preceq_μ on $\text{par}(\mu)$: $\forall B_1, B_2 \in \text{par}(\mu)$, $B_1 \preceq_\mu B_2$ iff $\mu(B_1) \subseteq \mu(B_2)$. It turns out that $(\text{par}(\mu), \preceq_\mu)$ is a poset. For example, consider the disjunctive abstraction μ depicted in Section 2, where $\mu = \text{Cl}_\cup(\{1, 12, 123, 4, 45, 6\})$. The poset $(\text{par}(\mu), \preceq_\mu)$ is then as follows:

$$\begin{array}{ccccc} & & & & 3 \\ & & & & | \\ & & & & 2 & 5 \\ & & & & | & | \\ & & & & 1 & 4 & 6 \\ & & & & | & | & | \\ & & & & 1 & 4 & 6 \end{array}$$

This allows us to represent the abstraction μ as follows: for any $S \subseteq \Sigma$, $\mu(S) = \cup\{B \in \text{par}(\mu) \mid \exists C \in \text{par}(\mu). C \cap S \neq \emptyset \ \& \ B \preceq_\mu C\}$.

As shown in [20], it turns out that this partition/relation-based representation provides an efficient way for representing and maintaining disjunctive abstractions. Moreover, [20] also shows that the abstraction refinement step $\mu_{i+1} = \text{Cl}_{\cap, \cup}(\mu_i \cup \{X\})$ at line 25 can be efficiently implemented by a procedure that is based on partition splitting and runs in $O(|\text{par}(\mu_i)|^2 + |X|)$ -time.

7 Future Work

A number of tasks are left for future research. Firstly, it would be interesting to complete the formal comparison between FBAR and CGR by investigating whether and how their refinements and final outputs can be related in the general case when the abstraction is refined. We also left open our conjecture that, analogously to CGR, the FBAR algorithm terminates when applied to well-structured transition systems. After the completion of such a comparison with CGR, it would be worth to develop a prototype of FBAR that can reuse the implementation of disjunctive abstraction refinements already available from [20].

Acknowledgments. This manuscript has been completed after the death of Olivia Rossi Doria on June 13th, 2006. Francesco Ranzato and Francesco Tapparo take this opportunity to acknowledge Olivia’s contribution to this paper as well as her personal friendship and professional skills.

References

1. T. Ball and S.K. Rajamani. The SLAM Project: Debugging system software via static analysis. In *Proc. 29th ACM POPL*, pp. 1-3, 2002.

2. S. Chaki, E.M. Clarke, A. Groce, S. Jha and H. Veith. Modular verification of software components in C. *IEEE Transactions on Software Engineering*, 30(6):388-402, 2004.
3. E.M. Clarke, O. Grumberg and D.A. Peled. *Model checking*. The MIT Press, 1999.
4. E.M. Clarke, O. Grumberg, S. Jha, Y. Lu and H. Veith. Progress on the state explosion problem in model checking In *Informatics - 10 Years Back, 10 Years Ahead*, LNCS 2000:176-194, 2001
5. 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.
6. P. Cousot. *Méthodes itératives de construction et d'approximation de points fixes d'opérateurs monotones sur un treillis, analyse sémantique de programmes*. PhD Thesis, Univ. de Grenoble, France, 1978.
7. P. Cousot and R. Cousot. Systematic design of program analysis frameworks. In *Proc. 6th ACM POPL*, pp. 269-282, 1979.
8. P. Cousot and R. Cousot. Abstract interpretation and application to logic programs. *J. Logic Programming*, 13(2-3):103-179, 1992.
9. P. Cousot and R. Cousot. Refining model checking by abstract interpretation. *Automated Software Engineering*, 6(1):69-95, 1999.
10. P. Cousot, P. Ganty and J.-F. Raskin. Fixpoint-guided abstraction refinements. In *Proc. 14th Int. Symp. on Static Analysis (SAS'07)*, LNCS 4634:333-348, 2007.
11. A. Finkel and P. Schnoebelen. Well-structured transition systems everywhere! *Theoretical Computer Science*, 256(1-2):63-92, 2001.
12. P. Ganty. *The Fixpoint Checking Problem: An Abstraction Refinement Perspective*. PhD Thesis, Université Libre de Bruxelles, 2007.
13. P. Ganty, J.F. Raskin and L. Van Begin. A complete abstract interpretation framework for coverability properties of WSTS. In *Proc. Int. Conf. on Verification, Model Checking and Abstract Interpretation (VMCAI'06)*. LNCS 3855:49-64, 2006.
14. P. Ganty, J.F. Raskin and L. Van Begin. From many places to few: Automatic abstraction refinement for Petri nets. In *Proc. 28th Int. Conf. on Applications and Theory of Petri Nets (ICATPN'07)*. LNCS 4546:124-143, 2007.
15. B.S. Gulavani and S.K. Rajamani. Counterexample driven refinement for abstract interpretation. In *Proc. 12th Int. Conf. on Tools and Algorithms for the Construction and Analysis of Systems (TACAS'06)*, LNCS 3920:474-488, 2006.
16. T.A. Henzinger, R. Jhala, R. Majumdar and G. Sutre. Lazy abstraction. In *Proc. 29th ACM POPL*, pp. 58-70, 2002.
17. T.A. Henzinger, R. Jhala, R. Majumdar and G. Sutre. Software verification with BLAST. In *Proc. of the 10th SPIN Workshop on Model Checking Software (SPIN'03)*, LNCS 2648:235-239, 2003.
18. R. Manevich, J. Field, T.A. Henzinger, G. Ramalingam and M. Sagiv. Abstract counterexample-based refinement for powerset domains. In *Program Analysis and Compilation, Theory and Practice: Essays Dedicated to Reinhard Wilhelm*, LNCS 4444:273-292, 2006.
19. D. Massé. Combining forward and backward analyses of temporal properties. In *Proc. 2nd Int. Symp. on Program as Data Objects (PADO'01)*, LNCS 2053:103-116, 2001.
20. F. Ranzato and F. Tapparo. An abstract interpretation-based refinement algorithm for strong preservation. In *Proc. 11th Intern. Conf. on Tools and Algorithms for the Construction and Analysis of Systems (TACAS'05)*, LNCS 3440:140-156, 2005. Extended version available on arXiv at: <http://arxiv.org/abs/cs/0612120>.