3-Valued Abstraction-Refinement

Sharon Shoham

Academic College of Tel-Aviv Yaffo

Model Checking

An efficient procedure that receives:

- A finite-state model describing a system
- A temporal logic formula describing a property

It returns yes, if the system has the property no + Counterexample, otherwise

[EC81,QS82]

Model Checking

- Emerging as an industrial standard tool for verification of hardware designs: Intel, IBM, Cadence, ...
- Recently applied successfully also for software verification: SLAM (Microsoft), Java PathFinder and SPIN (NASA), BLAST (EPFL), CBMC (Oxford),...



Temporal Logics

Express properties of event orderings in time

- Linear Time
 - Every moment has a unique successor
 - Infinite sequences (words)
 - Linear Time Temporal Logic (LTL)



- Branching Time
 - Every moment has several successors
 - Infinite tree
 - Computation Tree Logic
 (CTL), CTL*, μ-calculus



Propositional Temporal Logic

AP - a set of atomic propositions



Path quantifiers: A for all paths E there exists a path

LTL / CTL / CTL*

- LTL of the form $A\psi$
 - ψ path formula, contains **no** path **quantifiers**
- interpreted over infinite computation paths

CTL - path quantifiers and temporal operators appear in pairs: AG, AU, AF, AX, EG, EU, EF, EX

interpreted over infinite computation trees

CTL* - Allows any combination of temporal operators and path quantifiers. Includes both LTL and CTL

ACTL / ACTL*

The universal fragments of the logics, with only universal path quantifiers

- Negation is allowed only on atomic propositions

Main Limitation of Model Checking

The state explosion problem: Model checking is efficient w.r.t. to the state space of the model. But...

The number of states in the system model grows exponentially with

- the number of variables
- the number of components in the system

"Solutions" to the State Explosion Problem

Symbolic model checking: The model is represented symbolically

- BDD-based model checking
- SAT-based Bounded/ Unbounded model checking

Small models replace the full, concrete model:

- Abstraction
- Compositional verification
- Partial order reduction
- Symmetry

Outline

- Background:
 - Model Checking



- Abstraction
- CEX guided abstraction-refinement (CEGAR)
- 3-Valued Abstraction Refinement (TVAR)
- Example: TVAR for CTL
- Investigation of abstract models used in TVAR
 - Monotonicity of Refinement
 - Completeness
 - Precision
 - Efficiency

Abstraction-Refinement

 Abstraction: removes or simplifies details that are irrelevant to the property under consideration

Can reduce the number of states

- from large to small
- from infinite to finite
- **Refinement** might be needed

Widely used Abstractions

- Localization reduction / invisible variables: each variable either keeps its concrete behavior or is fully abstracted (has free behavior) [Kurshan94]
 - Initially: unabstracted variables are those appearing in the checked property
- Predicate abstraction: concrete states are grouped together according to the set of predicates they satisfy [6597,5599]
 - Initially: predicates are extracted from the program's control flow and the checked property

Abstraction Example

Abstract states

Concrete states



Abstraction Example

- Abstraction (S_A , γ):
 - Finite set of
 abstract states S_A
 - Abstraction mapping $\gamma : S_A \rightarrow 2^{S_c}$

not necessarily disjoint sets



Abstraction Example

- Abstraction (S_A , γ):
 - Finite set of
 abstract states S_A
 - Abstraction mapping γ : $S_A \rightarrow 2^{S_c}$ not necessarily disjoint sets



• Concrete Kripke structure $M_c = (S_c, R_c, L_c)$ \implies Abstract model over S_A : labeling, transitions Need to be conservative

Why Conservative?

Goal:

- Model check M_A instead of M_C
- **Deduce** result over M_c from result over M_A

What can we deduce?

- true ?
- false ?
- Both ?

For which properties?

Depends on abstract model and abstract semantics

2-valued CounterExample-Guided Abstraction Refinement (CEGAR)

For ACTL*

[CGJLV00, JACM2003]

Abstraction preserving ACTL/ACTL*

Existential Abstraction:

every concrete transition is represented by an abstract transition



The abstract model is an **over-approximation** of the concrete model:

- The abstract model has more behaviors
- no concrete behavior is lost

Simulation Relation

 $H \subseteq S_C \times S_A$ is a simulation relation from M_C to M_A if whenever $(s_c, s_a) \in H$:

- $L_{\mathcal{C}}(s_c) \supseteq L_{\mathcal{A}}(s_a)$
- If $(s_c, s_c') \in R_c$ for some s_c' , then there exists $s_a' s.t. (s_a, s_a') \in R_A$ and $(s_c', s_a') \in H$

If there exists a simulation relation obeying the initial states, then $M_C \leq_{sim} M_A$

Existential Abstraction

- Abstract model is also a Kripke structure
- Same semantics is used for abstract and concrete models
- → Same model checking algorithms

Abstract model checking result is **true** or **false** (2-valued)

But... what can we **deduce**?

Logic Preservation Theorem

Theorem. Let $M_C \leq_{sim} M_A$. Then:

 Every ACTL/ACTL* property true in the abstract model is also true in the concrete model:

$$M_A \models \phi \Rightarrow M_C \models \phi$$

However, the reverse may not be valid:

- If $M_A \neq \phi$, need to check further
 - Check if abstract counterexample is spurious



Three-Valued Abstraction Refinement (TVAR)

for Full CTL*

[SG03,GLLS05]

•26

2-valued Approach is not Applicable

 Over-approximation (simulation) of the concrete model is not sound for verification of existential properties:

 $M_A \mid = E \psi$ does not imply $M_C \mid = E \psi$

→ More complex abstract models (and relations) are needed to ensure logic preservation

Abstract Models for CTL*

Branching-time temporal logics combining existential (E) and universal (A) quantifiers:

→ two transition relations [LT88]

- Rmay: an over-approximation
- **Rmust**: an **under-approximation**

Logic Preservation for CTL*

If M_A is an abstraction of M_C then for every CTL* formula φ ,

$$\begin{array}{c|c} \mathsf{M}_{\mathsf{A}} & | = \phi \Rightarrow \mathsf{M}_{\mathcal{C}} & | = \phi \\ \mathsf{M}_{\mathsf{A}} & | \neq \phi \Rightarrow \mathsf{M}_{\mathcal{C}} & | \neq \phi \end{array} \end{array}$$

• But sometimes $[M_A | = \phi] = don't know$

→ 3-Valued Semantics 3 possible values: True, False , ⊥ (indefinite)

Refinement

• Refinement is needed when result is \perp

Traditional abstraction-refinement for universal properties **not** applicable:

- Refinement needed when result is false
- Based on a counterexample

Three-Valued Abstraction-Refinement (TVAR)

The TVAR Methodology



Main Components

1. Abstract Models:

What formalism is suitable? How to construct an abstract model in a conservative way?

2. Model Checking:

How to evaluate **branching-time** formulas over abstract models based on the **3-valued semantics**?

3. Refinement:

How to refine the abstract model?

TVAR for CTL using Kripke Modal Transition Systems

[SG03]

Abstract Models

Kripke Modal Transition System (KMTS) [HJS01] • M = (AP, S, s⁰, Rmust, Rmay, L)

- **Rmust** \subseteq S x S: an **under-approximation**
- Rmay \subseteq S x S: an over-approximation
- Rmust \subseteq Rmay

For simplicity. In MixTS, no such requirement

Abstract Models

Labeling function :

- L: S \rightarrow 2^{Literals}
- Literals = $AP \cup \{\neg p \mid p \in AP\}$
- At most one of p and $\neg p$ is in L(s).
 - Concrete: exactly one of p and $\neg p$ is in L(s).
 - KMTS: possibly none of them is in L(s), meaning that the value of p in s is unknown

3-Valued Semantics [BG99]

 $[[lit]](s) = \text{tt if } lit \in L_A(s), \text{ ff if } -lit \in L_A(s), \perp o.w.$

 $[[AX\psi]] (s) = \begin{cases} \text{tt if forall } s', \text{ if } (s, s') \in \mathbf{Rmay}, \\ \text{ then } [[\psi]] (s') = \text{tt} \end{cases}$ $ff \text{ if exists } s' \text{ s.t. } (s, s') \in \mathbf{Rmust} \\ and [[\psi]] (s') = ff \end{cases}$ $\bot \text{ otherwise}$

[[EXψ]] (s) - dual "exists succ. satisfying ψ"

Construction of Abstract Model

Labeling of abstract states



 $[\forall s_c \in \gamma(s_a) \ \text{lit} \in L_c(s_c)] \quad \Leftrightarrow \ \text{lit} \in L_A(s_a)$



 $[\forall s_c \in \gamma(s_a) \exists s_c' \in \gamma(s_a') \text{ s.t. } (s_c, s_c') \in \mathsf{R}_C] \Leftrightarrow (s_a, s_a') \in \mathsf{R}_{\mathsf{must}}$


 $[\exists s_c \in \gamma(s_a) \exists s_c' \in \gamma(s_a') \text{ s.t. } (s_c, s_c') \in \mathsf{R}_C] \Leftrightarrow (s_a, s_a') \in \mathsf{R}_{may}$

Mixed Simulation

 $H \subseteq S_C \times S_A$ is a mixed simulation relation from Kripke structure M_C to KMTS M_A if whenever $(s_c, s_a) \in H$:

- $L_{\mathcal{C}}(s_{c}) \supseteq L_{\mathcal{A}}(s_{a})$
- If $(s_c, s_c') \in R_c$, then there is $s_a' \ s.t. \ (s_a, s_a') \in Rmay$ and $(s_c', s_a') \in H$
- If $(s_a, s_a') \in \text{Rmust}$, then there is $s_c' \text{ s.t. } (s_c, s_c') \in R_C$ and $(s_c', s_a') \in H$

If there exists a mixed simulation relation obeying the initial states, then $M_C \leq_{mix} M_A$

Logic Preservation

Theorem. Let $M_C \leq_{mix} M_A$. Then: For every CTL* formula φ , $M_A \models \varphi \Rightarrow M_C \models \varphi$ $M_A \neq \varphi \Rightarrow M_C \neq \varphi$

But if $[M_A = \phi] = \bot$, the value in M_C is unknown

3-Valued Model Checking Example











MC graph



Terminal nodes: based on states labeling \land, \lor, AX, EX : according to sons, based on semantics

3-Valued Model Checking Results

- tt and ff are definite: hold in the concrete model as well.
- \bot is indefinite
 - \Rightarrow Refinement is needed.

Refinement

done by splitting abstract states
 (as for the case of 2-values)



Refinement

- Uses the colored MC graph
- Find a failure node n_f:
 - a node colored ⊥ whereas none of its sons was
 colored ⊥ at the time it got colored.
 - the point where certainty was lost
- purpose: change the \perp color of n_f .



- Failure reason is either:
 - A may-edge which is not a must-edge.
 - A *L*-terminal node
- Back in the model:
 - Either a transition (s, s') ∈ Rmay\Rmust:
 ⇒ Split s to get a must-transition or none.

- Failure reason is eit
 - A may-edge which i
 - A *L*-terminal node



- Back in the model:
 - Either a transition (s, s') \in Rmay\Rmust:
 - \rightarrow Split s to get a must-transition or none.

- Failure reason is either:
 - A may-edge which is not a must-edge.
 - A *L*-terminal node
- Back in the model:
 - Either a transition (s, s') \in Rmay\Rmust:
 - \rightarrow Split s to get a must-transition or none.
 - Or (s,lit) where lit ∉ L(s), ¬lit ∉ L(s)
 → Split s according to lit.

- Failure reason is eith
 - A may-edge which is
 A *L*-terminal node



- Back in the model:
 - Either a transition (s, s') \in Rmay\Rmust:
 - → Split s to get a must-transition or none.
 - Or (s,lit) where $lit \notin L(s)$, $\neg lit \notin L(s)$ \rightarrow Split s according to lit.

Split

- Refinement is reduced to separating sets of concrete states.
 - done by known techniques [CGJLV00,CGKS02]
 - → Refined abstraction mapping.
- Build refined abstract model and refined MC-graph accordingly.









Incremental Abstraction-Refinement

No reason to split states for which MC results are definite during refinement.

- After each iteration remember the nodes colored by definite colors.
- Prune the refined MC graph in sub-nodes of remembered nodes.
 [(s_a, φ) is a sub-node of (s_a', φ') if φ=φ' and γ(s_a)⊆γ'(s_a')]
- Color such nodes by their previous colors.





Example (cont.)



Refined MC-graph

Example (cont.)



Example (cont.)



Are KMTSs good enough for TVAR?

Investigation of Abstract Models

- Monotonicity of Refinement
- Precision
- Completeness
- Efficiency

(1) Monotonicity of Refinement

Is a **refined** abstract model **at least as precise** as the **unrefined** one?

Example

P :: input x > 0 pc=1: if x>5 then x := x+1 else x := x+2 pc=2: while true do if odd(x) then x := -1 else x := x+1

φ = EF (x ≤ 0)

P :: input x > 0An Abstract Model M pc=1: if x>5 then x := x+1 else x := x+2 pc=2: while true do pc=1 if odd(x) then x := -1S₀ x>0 else x := x+1S₂ **S**₁ pc=2 pc=2 x>0 x≤0

[EF (× ≤ 0)] (M) = ⊥



P :: input x > 0 pc=1: if x>5 then x := x+1 else x := x+2 pc=2: while true do if odd(x) then x := -1 else x := x+1





P :: input x > 0 pc=1: if x>5 then x := x+1 else x := x+2 pc=2: while true do if odd(x) then x := -1 else x := x+1



Problem



- When splitting states during refinement we may lose must transitions
- Existential formulas that were true before may become indefinite ! (also universal formulas that were false)
- Thus, the refined model is not necessarily more precise
- → refinement is not monotonic

Goal: define a refinement that adds underapproximated must transitions

[current refinements **remove** over-approximated **may** transitions]

Result: refined model will be **more precise**, i.e. more formulas will be definite (**tt** or **ff**) in it:

Monotonic Refinement

Notation: $M' \leq_{CTL} M : M'$ is more precise than M

Refinement M" of M, according to Godefroid et. al.



(2) Precision

Given a state abstraction (S_A, γ)

 "How many" formulas can be verified or falsified on the abstract model?

Refinement M" of M, according to Godefroid et. al.



Another Solution [56'04]

Use hyper-transitions as must transitions

Hyper-transition from a state $s \in S$ is

• (s, A) where and $A \subseteq S$ is nonempty
Generalized KMTS (GTS)

$M = (S, S_0, Rmay, Rmust, L)$

- S, S_0 , Rmay, L as before
- Rmust \subseteq S x 2^s

Constructing an Abstract GTS

Given M_C , S_A , and γ : $S_A \rightarrow 2^{S_C}$

• $(s_{\alpha}, A) \in \text{Rmust only if } \forall \exists \exists \text{-condition holds:}$



every state in $\gamma(s_a)$ has a corresponding transition

C
Reminder: in KMTS:
$$(s_a, s'_a) \in \text{Rmust only if } \forall \exists \text{-condition holds:}$$

 $\forall s_c \in \gamma(s_a) \exists s'_c \in \gamma(s'_a) : (s_c, s'_c) \in \mathbb{R}_c$
Given M

• $(s_{\alpha}, A) \in \text{Rmust only if } \forall \exists \exists \text{-condition holds}$:



every state in $\gamma(s_a)$ has a corresponding transition

3-Valued Semantics over GTS

 $[[lit]](s) = \text{tt if } lit \in L_A(s), \text{ ff if } -lit \in L_A(s), \perp o.w.$

 $[[AX\psi]] (s) = \begin{cases} \text{tt if forall } s', \text{ if } (s, s') \in \mathbf{Rmay}, \\ \text{then } [[\psi]] (s') = \text{tt} \end{cases}$ ff if exists $A \subseteq S_A \text{ s.t. } (s, A) \in \mathbf{Rmust}$ and $[[\psi]] (s') = \text{ff forall } s' \in A$ \bot otherwise

[[EXψ]] (s) - dual "exists succ. satisfying ψ"

Must Hyper-transition ($\forall \exists \exists$)



pc=1: if x>5 then x := x+1 else x := x+2 Pc=2: ...

Every concrete state in $\gamma(s_{00})$ has a transition to a concrete state in either $\gamma(s_{10})$ or $\gamma(s_{11})$



 $M_G \leq_{CTL} M'' \leq_{CTL} M$ and $[EF(x \leq 0)](M_G) = tt$

Monotonicity Theorem:

Let M_{A} and M'_{A} be two abstract GTSs of M_{c} such that

- M'_A is obtained from M_A by splitting states
- Both M_A and M'_A are exact

Then M'_A is more precise than M_A



To complete the picture...

Extension of the game-based 3-Valued
Model Checking and Failure Analysis to
GTSs

Investigation of Abstract Models

- Monotonicity of Refinement
- Precision
- Completeness
- Efficiency

(3) Completeness

- Suppose $M_c \mid = \varphi$
- Does there exist a finite abstraction (S_A, γ) such that $[M_A \mid = \varphi] = tt$?

Monotonicity vs. Completeness vs. Precision

- Monotonicity of refinement: Given two abstractions, where one is a split of the other, is refined abstraction more precise than unrefined one?
- Precision:

How many formulas can be verified on the abstract model, with a given abstraction (S_A, γ) ?

• Completeness:

Does there exist an abstraction (S_A , γ) for which we can verify the formula on the abstract model?

Are KMTSs complete?

No fairness constraints
incomplete for liveness properties

What about **Safety**? (no least fixpoint)

No [Dams & Namjoshi, 2004] But GTSs are! [de Alfaro et al, 2004]

Investigation of Abstract Models

- Monotonicity of Refinement
- Precision
- Completeness
- Efficiency

(4) Efficiency

Cost:

- Size of the abstract model w.r.t. $|S_A|$
- Efficiency of Model Checking

Drawback of GTS

The number of must hyper transitions might be exponential in the number of abstract states $|S_A|$

Optimization:

including only (s, A) such that A is minimal

Does not change precision of the abstract model

But, might still be too large

In Practice

 Not all hyper-transitions are relevant for specific model checking problem



→ Need to find designated hyper-transitions

Alternative Approach [5606]

 Compute hyper-transitions during Model Checking, by need

→ Game-based Model Checking

Our Algorithm

Ordinary transitions

- Compute over approximation of concrete transition relation
 (s_a, s'_a)∈R_A iff
 ∃s_c∈γ(s_a) ∃s'_c∈γ(s'_a): (s_c, s'_c) ∈ R_c
 All reachable states are considered
- Construct MC graph based on R_A
- Apply bottom up coloring



(s, A_{tt}) meets $\forall \exists \exists$ -condition [must]? yes: [[$EX\psi$]](s₀) = tt



(s, A_{tt}) meets $\forall \exists \exists$ -condition [must]? yes: $[[EX\psi]](s_0) = tt$ All may transitions reach A_{ff} ? yes: $[[EX\psi]](s_0) = ff$ otherwise: $[[EX\psi]](s_0) = \bot$

Abstract Model Checking

Loops: slight complication

Comparable to the complexity without hypertransitions

In the paper [SG06]:

- Abstract MC for the alternation-free μ -calculus
- Complexity: $O(|S_A|^2 \times |\varphi|)$
- In particular: num of ∀∃∃ checks, num of hyper transitions

As precise as constructing the full GTS

Abstraction-Refinement

- If $[[\phi]](s_0) = \bot$, apply refinement by splitting abstract states, as in [5603]
- Refinement is monotonic:

refined model is more precise, i.e. more μ -calculus formulas are definite (tt or ff) in it

Abstraction-refinement loop

Summary

We presented the TVAR framework for <u>3-valued abstraction-refinement</u> in model checking:

- Properties preserved:
 - CEGAR: truth of ACTL*
 - TVAR: both truth and falsity of Full CTL*
- Refinement eliminates
 - CEGAR: Counterexamples
 - TVAR: indefinite results (\perp)

Summary

The TVAR framework requires

- 1. Different abstract models (Rmust, Rmay)
 - Rmust is harder to compute, and problematic in terms of monotonicity, precision, completeness, and efficiency
 - KMTS, GTS, HTS
- 2. Adapted Model checking for new models:
 - 3-valued Coloring of MC-graph

Summary

The TVAR framework requires

- 3. Refinement eliminating indefinite results
 - Identify failure state and cause
 - Incremental abstraction-refinement (similar to lazy abstraction in 2-valued MC)

Gives benefits in preciseness and in the properties preserved