Algorithmic Logic-Based Verification with SeaHorn

Software Engineering Institute
Carnegie Mellon University
Pittsburgh, PA 15213

Arie Gurfinkel with Teme Kahsai and Jorge A. Navas

based on joint work with Anvesh Komuravelli, and Nikolaj Bjørner
Automated Software Analysis

Program \[\rightarrow\] Automated Analysis

Correct \[\rightarrow\] Incorrect

- Software Model Checking with Predicate Abstraction
e.g., Microsoft’s SDV

- Abstract Interpretation with Numeric Abstraction
e.g., ASTREE, Polyspace
Turing, 1936: “undecidable”
How can one check a routine in the sense of making sure that it is right? The programmer should make a number of definite assertions which can be checked individually, and from which the correctness of the whole programme easily follows.
Three-Layers of a Program Verifier

Compiler
- compiles surface syntax to a target machine
- embodies syntax with semantics

Verification Condition Generator
- transforms a program and a property to a verification condition in logic
- employs different abstractions, refinements, proof-search strategies, etc.

Automated Theorem Prover / Reasoning Engine
- discharges verification conditions
- general purpose constraint solver
- SAT, SMT, Abstract Interpreter, Temporal Logic Model Checker, etc.
SeaHorn

A fully automated verification framework for LLVM-based languages.

http://seahorn.github.io
SeaHorn Verification Framework

Arie Gurfinkel  
Software Engineering Institute  
Carnegie Mellon University

Temesghen Kahsai  
Carnegie Mellon University  
NASA Ames

Jorge A. Navas  
SGT  
NASA Ames
The Plan

Introduction

Architecture and Usage

Demonstration

Constrained Horn Clauses as an Intermediate Representation

From Programs to Logic

• generating verification conditions

Program Transformations for Verification

Solving Constrained Horn Clauses

• synthesizing inductive invariants and procedure summaries

Conclusion
SeaHorn Verification Framework

Key Features
- LLVM front-end(s)
- Constrained Horn Clauses to represent Verification Conditions
- Comparable to state-of-the-art tools at SV-COMP’15

Goals
- be a state-of-the-art Software Model Checker
- be a framework for experimenting and developing CHC-based verification
Related Tools

CPAChecker
- Custom front-end for C
- Abstract Interpretation-inspired verification engine
- Predicate abstraction, invariant generation, BMC, k-induction

SMACK / Corral
- LLVM-based front-end
- Reduces C verification to Boogie
- Corral / Q verification back-end based on Bounded Model Checking with SMT

UFO
- LLVM-based front-end (partially reused in SeaHorn)
- Combines Abstract Interpretation with Interpolation-Based Model Checking
- (no longer actively developed)
SeaHorn Philosophy

Build a state-of-the-art Software Model Checker

• useful to “average” users
  – user-friendly, efficient, trusted, certificate-producing, …
• useful to researchers in verification
  – modular design, clean separation between syntax, semantics, and logic, …

Stand on the shoulders of giants

• reuse techniques from compiler community to reduce verification effort
  – SSA, loop restructuring, induction variables, alias analysis, …
  – static analysis and abstract interpretation
• reduce verification to logic
  – verification condition generation
  – Constrained Horn Clauses

Build reusable logic-based verification technology

• “SMT-LIB” for program verification
SeaHorn Usage

> sea pf FILE.c
Outputs sat for unsafe (has counterexample); unsat for safe

Additional options

- --cex=trace.xml outputs a counter-example in SV-COMP’15 format
- --show-invars displays computed invariants
- --track={reg,ptr,mem} track registers, pointers, memory content
- --step={large,small} verification condition step-semantics
  - small == basic block, large == loop-free control flow block
- --inline inline all functions in the front-end passes

Additional commands

- sea smt – generates CHC in extension of SMT-LIB2 format
- sea clp -- generates CHC in CLP format (under development)
- sea lfe-smt – generates CHC in SMT-LIB2 format using legacy front-end
Verification Pipeline

front-end

clang | pp | ms | opt | horn

compile  pre-process  optimize  mixed semantics  VC gen & solve
DEMO
From Programming to Modeling

Extend C programming language with 3 modeling features

Assertions

• `assert(e)` – aborts an execution when `e` is false, no-op otherwise

```c
void assert (_Bool b) { if (!b) abort(); }
```

Non-determinism

• `nondet_int()` – returns a non-deterministic integer value

```c
int nondet_int () { int x; return x; }
```

Assumptions

• `assume(e)` – “ignores” execution when `e` is false, no-op otherwise

```c
void assume (_Bool e) { while (!e); }
```
Constrained Horn Clauses

INTERMEDIATE REPRESENTATION
Constrained Horn Clauses (CHC)

A Constrained Horn Clause (CHC) is a FOL formula of the form

$$\forall V . (\phi \land p_1[X_1] \land \ldots \land p_n[X_n] \rightarrow h[X])$$,

where

- $A$ is a background theory (e.g., Linear Arithmetic, Arrays, Bit-Vectors, or combinations of the above)
- $\phi$ is a constrained in the background theory $A$
- $p_1, \ldots, p_n, h$ are n-ary predicates
- $p_i[X]$ is an application of a predicate to first-order terms
Example Horn Encoding

```plaintext
typedef X = int

int x = 1;
int y = 0;
while (*) {
    x = x + y;
    y = y + 1;
} assert(x >= y);
```

```plaintext
\[ l_0 : \]
\[ x = 1 \]
\[ y = 0 \]

\[ l_1 : b_1 = \text{nondet()} \]

\[ l_2 : \\
x = x + y \\
y = y + 1 \]

\[ l_3 : \\
b_2 = x \geq y \]

\[ l_4 : \\
\]

\[ l_\text{err} : \]

\[ 1 \] \( p_0 \).
\[ 2 \] \( p_1(x, y) \leftarrow p_0, x = 1, y = 0. \)
\[ 3 \] \( p_2(x, y) \leftarrow p_1(x, y). \)
\[ 4 \] \( p_3(x, y) \leftarrow p_1(x, y). \)
\[ 5 \] \( p_1(x', y') \leftarrow p_2(x, y), \]
\[ \quad x' = x + y, \]
\[ \quad y' = y + 1. \)
\[ 6 \] \( p_4 \leftarrow (x \geq y), p_3(x, y). \)
\[ 7 \] \( p_\text{err} \leftarrow (x < y), p_3(x, y). \)
\[ 8 \] \( p_4 \leftarrow p_4. \)
\[ 9 \] \( \bot \leftarrow p_\text{err}. \)
```
CHC Terminology

Rule

\[ h[X] \leftarrow p_1[X_1], \ldots, p_n[X_n], \phi. \]

Query

\[ \text{false} \leftarrow p_1[X_1], \ldots, p_n[X_n], \phi. \]

Fact

\[ h[X] \leftarrow \phi. \]

Linear CHC

\[ h[X] \leftarrow p[X_1], \phi. \]

Non-Linear CHC

\[ h[X] \leftarrow p_1[X_1], \ldots, p_n[X_n], \phi. \]

for \( n > 1 \)
CHC Satisfiability

A **model** of a set of clauses $\Pi$ is an interpretation of each predicate $p_i$ that makes all clauses in $\Pi$ valid.

A set of clauses is **satisfiable** if it has a model, and is unsatisfiable otherwise.

A model is **A-definable**, if each $p_i$ is definable by a formula $\psi_i$ in A.
Example Horn Encoding

\[
\begin{align*}
\text{l}_0 : & \quad x = 1 \quad y = 0 \\
\text{l}_1 : & \quad b_1 = \text{nondet()} \\
\text{l}_2 : & \quad x = x + y \quad y = y + 1 \\
\text{l}_3 : & \quad b_2 = x \geq y
\end{align*}
\]

\[
\begin{align*}
\langle 1 \rangle & : \quad p_0. \\
\langle 2 \rangle & : \quad p_1(x, y) \leftarrow p_0, x = 1, y = 0. \\
\langle 3 \rangle & : \quad p_2(x, y) \leftarrow p_1(x, y). \\
\langle 4 \rangle & : \quad p_3(x, y) \leftarrow p_1(x, y). \\
\langle 5 \rangle & : \quad p_1(x', y') \leftarrow p_2(x, y), \\
& \quad x' = x + y, \\
& \quad y' = y + 1. \\
\langle 6 \rangle & : \quad p_4 \leftarrow (x \geq y), p_3(x, y). \\
\langle 7 \rangle & : \quad p_{\text{err}} \leftarrow (x < y), p_3(x, y). \\
\langle 8 \rangle & : \quad p_4 \leftarrow p_4. \\
\langle 9 \rangle & : \quad \bot \leftarrow p_{\text{err}}.
\end{align*}
\]
Relationship between CHC and Verification

A program satisfies a property iff corresponding CHCs are satisfiable
• satisfiability-preserving transformations == safety preserving

Models for CHC correspond to verification certificates
• inductive invariants and procedure summaries

Unsatisfiability (or derivation of FALSE) corresponds to counterexample
• the resolution derivation (a path or a tree) is the counterexample

CAVEAT: In SeaHorn the terminology is reversed
• SAT means there exists a counterexample – a BMC at some depth is SAT
• UNSAT means the program is safe – BMC at all depths are UNSAT
FROM PROGRAMS TO CLAUSES
Hoare Triples

A Hoare triple \( \{\text{Pre}\} \ P \ {\text{Post}} \) is valid iff every terminating execution of \( P \) that starts in a state that satisfies \( \text{Pre} \) ends in a state that satisfies \( \text{Post} \).

**Inductive Loop Invariant**

\[
\begin{align*}
\text{Pre} & \Rightarrow \text{Inv} \\
\{\text{Inv} \land C\} \text{ Body} \{\text{Inv}\} & \quad \text{Inv} \land \lnot C \Rightarrow \text{Post} \\
\{\text{Pre}\} \textbf{ while } C \textbf{ do Body } \{\text{Post}\}
\end{align*}
\]

**Function Application**

\[
\begin{align*}
(\text{Pre} \land p=a) & \Rightarrow P \\
\{P\} \text{ Body}_F \{Q\} & \quad (Q \land p,r=a,b) \Rightarrow \text{Post} \\
\{\text{Pre}\} b = F(a) & \{\text{Post}\}
\end{align*}
\]

**Recursion**

\[
\begin{align*}
\{\text{Pre}\} b = F(a) & \{\text{Post}\} \vdash \{\text{Pre}\} \text{ Body}_F \{\text{Post}\} \\
\{\text{Pre}\} b = F(a) & \{\text{Post}\}
\end{align*}
\]
Weakest Liberal Pre-Condition

Validity of Hoare triples is reduced to FOL validity by applying a predicate transformer

Dijkstra’s weakest liberal pre-condition calculus [Dijkstra’75]

\[ \text{wlp} (P, \text{Post}) \]

weakest pre-condition ensuring that executing \( P \) ends in \( \text{Post} \)

\[ \{\text{Pre}\} \ P \{\text{Post}\} \text{ is valid} \iff \text{Pre} \Rightarrow \text{wlp} (P, \text{Post}) \]
A Simple Programming Language

Prog ::= def Main(x) { body_M }, ..., def P(x) { body_P }

body ::= stmt (; stmt)*

stmt ::= x = E | assert (E) | assume (E) |
          while E do S | y = P(E) |
          L:stmt | goto L (optional)

E ::= expression over program variables
Horn Clauses by Weakest Liberal Precondition

\[ \text{Prog ::= def Main(x) \{ body}_M \}, ..., \text{def P(x) \{ body}_P \} \]

\[
\begin{align*}
\text{wlp (x=E, Q) & = let x=E in Q} \\
\text{wlp (assert(E), Q) & = E \land Q} \\
\text{wlp (assume(E), Q) & = E \rightarrow Q} \\
\text{wlp (while E do S, Q) & = I(w) \land} \\
& \quad \forall w . ((I(w) \land E) \rightarrow \text{wlp (S, I(w)))) \land ((I(w) \land \neg E) \rightarrow Q)) \\
\text{wlp (y = P(E), Q) & = p}_{\text{pre}}(E) \land (\forall r. p(E, r) \rightarrow Q[r/y])
\end{align*}
\]

\[
\begin{align*}
\text{ToHorn (def P(x) \{S\}) & = wlp (x0=x; assume(p}_{\text{pre}}(x)); S, p(x0, ret))} \\
\text{ToHorn (Prog) & = wlp (Main(), true) \land \forall \{P \in \text{Prog}\} . \text{ToHorn (P)}}
\end{align*}
\]
Example of a WLP Horn Encoding

{Pre: $y \geq 0$}
$\quad x_o = x$;
$\quad y_o = y$;
$\quad \text{while } y > 0 \text{ do}$
$\quad \quad x = x+1;$
$\quad \quad y = y-1;$
{Post: $x=x_o+y_o$}

ToHorn

C1: $I(x,y,x,y) \leftarrow y \geq 0$.
C2: $I(x+1,y-1,x_o,y_o) \leftarrow I(x,y,x_o,y_o), y>0$.
C3: false $\leftarrow I(x,y,x_o,y_o), y \leq 0, x \neq x_o+y_o$

{y $\geq 0$} P {x = x$_{old}$+y$_{old}$} is true iff the query C$_3$ is satisfiable
Example Horn Encoding

\[
\begin{align*}
\text{int } & x = 1; \\
\text{int } & y = 0; \\
\text{while } & (* ) \{ \\
& x = x + y; \\
& y = y + 1; \\
\} \\
\text{assert} ( x \geq y );
\end{align*}
\]

\[
\begin{align*}
l_0 : & \\
& x = 1 \\
& y = 0
\end{align*}
\]

\[
\begin{align*}
l_1 : b_1 = \text{non}d et ()
\end{align*}
\]

\[
\begin{align*}
l_2 : & \\
& x = x + y \\
& y = y + 1
\end{align*}
\]

\[
\begin{align*}
l_3 : & \\
& b_2 = x \geq y
\end{align*}
\]

\[
\begin{align*}
l_4 : & \\
& F
\end{align*}
\]

\[
\begin{align*}
l_{\text{err}} : & \\
& T
\end{align*}
\]

\[
\begin{align*}
\langle 1 \rangle & p_0. \\
\langle 2 \rangle & p_1 ( x, y ) \leftarrow \\
& p_0, x = 1, y = 0. \\
\langle 3 \rangle & p_2 ( x, y ) \leftarrow p_1 ( x, y ). \\
\langle 4 \rangle & p_3 ( x, y ) \leftarrow p_1 ( x, y ). \\
\langle 5 \rangle & p_1 ( x', y' ) \leftarrow \\
& p_2 ( x, y ), \\
& x' = x + y, \\
& y' = y + 1. \\
\langle 6 \rangle & p_4 \leftarrow ( x \geq y ) \land p_3 ( x, y ). \\
\langle 7 \rangle & p_{\text{err}} \leftarrow ( x < y ) \land p_3 ( x, y ). \\
\langle 8 \rangle & p_4 \leftarrow p_4. \\
\langle 9 \rangle & \perp \leftarrow p_{\text{err}}.
\end{align*}
\]
From CFG to Cut Point Graph

A Cut Point Graph hides (summarizes) fragments of a control flow graph by (summary) edges

Vertices (called, cut points) correspond to some basic blocks

An edge between cut-points $c$ and $d$ summarizes all finite (loop-free) executions from $c$ to $d$ that do not pass through any other cut-points
Cut Point Graph Example

CFG

CPG
Deeply nested assertions
Deeply nested assertions

Counter-examples are long
Hard to determine (from main) what is relevant
Mixed Semantics

Stack-free program semantics combining:

• operational (or small-step) semantics
  – i.e., usual execution semantics
• natural (or big-step) semantics: function summary [Sharir-Pnueli 81]
  – \((\sigma, \sigma') \in \|f\|\) iff the execution of \(f\) on input state \(\sigma\) terminates and results in state \(\sigma'\)
• some execution steps are big, some are small

Non-deterministic executions of function calls

• update top activation record using function summary, or
• enter function body, forgetting history records (i.e., no return!)

Preserves reachability and non-termination

**Theorem:** Let \(K\) be the operational semantics, \(K^m\) the stack-free semantics, and \(L\) a program location. Then,

\[ K \models EF (pc=L) \iff K^m \models EF (pc=L) \quad \text{and} \quad K \models EG (pc \neq L) \iff K^m \models EG (pc \neq L) \]
def main():
1: int x = nd();
2: x = x+1;
3: while(x>=0)
4: x=f(x);
5: if(x<0)
6: Error;
7:
8: END;

def f(int y): ret y
9: if(y,10){
10: y=y+1;
11: y=f(y);
12: else if(y>0)
13: y=y+1;
14: y=y-1
15:

Summary of f(y)
(1≤y≤9 ∧ y’=y) ∨ (y≤0 ∧ y’=y-1)

Verification with SeaHorn
Gurfinkel, 2015
© 2015 Carnegie Mellon University
Mixed Semantics as Program Transformation

```
main ()
    p1 (); p1 ();
    assert (c1);

p1 ()
    p2 ();
    assert (c2);

p2 ()
    assert (c3);
```

```
main_new ()
    if (*) goto p1_entry;
    else p1_new ();
    if (*) goto p1_entry;
    else p1_new ();
    if (¬c1) goto error;
    assume (false);

p1_entry :
    if (*) goto p2_entry;
    else p2_new ();
    if (¬c2) goto error;
    p2_entry :
    if (¬c3) goto error;
    assume (false);

error : assert (false);
```

Verification with SeaHorn
Gurfinkel, 2015
© 2015 Carnegie Mellon University
Mixed Semantics: Summary

Every procedure is inlined at most once
- in the worst case, doubles the size of the program
- can be restricted to only inline functions that directly or indirectly call `error()` function

Easy to implement at compiler level
- create “failing” and “passing” versions of each function
- reduce “passing” functions to returning paths
- in main(), introduce new basic block `bb.F` for every failing function `F()`, and call `failing.F` in `bb.F`
- inline all failing calls
- replace every call to `F` to non-deterministic jump to `bb.F` or call to passing `F`

Increases context-sensitivity of context-insensitive analyses
- context of failing paths is explicit in main (because of inlining)
- enables / improves many traditional analyses
SOLVING CHC WITH SMT
Programs, Cexs, Invariants

A program $P = (V, \text{Init}, \rho, \text{Bad})$

- Notation: $\mathcal{F}(X) = \exists \; u \cdot (X \land \rho) \lor \text{Init}$

$P$ is UNSAFE if and only if there exists a number $N$ s.t.

$$
\text{Init}(v_0) \land \left( \bigwedge_{i=0}^{N-1} \rho(v_i, v_{i+1}) \right) \land \text{Bad}(v_N) \not\rightarrow \bot
$$

$P$ is SAFE if and only if there exists a safe inductive invariant $\text{Inv}$ s.t.

$$
\begin{align*}
\text{Init}(u) & \Rightarrow \text{Inv}(u) \\
\text{Inv}(u) \land \rho(u, v) & \Rightarrow \text{Inv}(v) \\
\text{Inv}(u) & \Rightarrow \neg \text{Bad}(u)
\end{align*}
$$
Verification by Evolving Approximations

Approx. 1

Solver

Inductive Invariant

Safe?

No

Approx. 2

Solver

Inductive Invariant

Safe?

No

Approx. 3

Solver

Inductive Invariant

Safe?

No
IC3/PDR Algorithm Overview

Input: Safety problem \( \langle \text{Init}(X), \text{Tr}(X, X'), \text{Bad}(X) \rangle \)

\[
F_0 \leftarrow \text{Init}; N \leftarrow 0 \text{ repeat}
\]

\[
\text{G} \leftarrow \text{PdrMkSafe}([F_0, \ldots, F_N], \text{Bad})
\]

\[
\text{if } G = [] \text{ then return } \text{Reachable;}
\]

\[
\forall 0 \leq i \leq N \cdot F_i \leftarrow G[i]
\]

\[
F_0, \ldots, F_N \leftarrow \text{PdrPush}([F_0, \ldots, F_N])
\]

\[
\text{if } \exists 0 \leq i < N \cdot F_i = F_{i+1} \text{ then return } \text{Unreachable;}
\]

\[
N \leftarrow N + 1; F_N \leftarrow \emptyset
\]

until \( \infty \);
IC3/PDR in Pictures

Frame $R_0$  Frame $R_1$  lemma

Cex Queue

PdrMkSafe

Trace
IC3/PDR in Pictures
IC3/PDR in Pictures

PDR Invariants

\[ R_i \rightarrow \neg \text{Bad} \quad \text{Init} \rightarrow R_i \]

\[ R_i \rightarrow R_{i+1} \quad R_i \land \rho \rightarrow R_{i+1} \]

Inductive
Spacer: Solving CHC in Z3

Spacer: solver for SMT-constrained Horn Clauses
  - stand-alone implementation in a fork of Z3
  - http://bitbucket.org/spacer/code

Support for Non-Linear CHC
  - model procedure summaries in inter-procedural verification conditions
  - model assume-guarantee reasoning
  - uses MBP to under-approximate models for finite unfoldings of predicates
  - uses MAX-SAT to decide on an unfolding strategy

Supported SMT-Theories
  - Best-effort support for arbitrary SMT-theories
    - data-structures, bit-vectors, non-linear arithmetic
  - Full support for Linear arithmetic (rational and integer)
  - Quantifier-free theory of arrays
    - only quantifier free models with limited applications of array equality
CRAB: Cornucopia of Abstractions

A library of abstract domains build on top of NASA Ikos (Inference Kernel for Open Static Analyzers)

A language-independent intermediate representation

Many abstract domains

• intervals (with congruences) (with uninterpreted functions)
• zones, dbms, octagons
• pointer analysis with offsets
• array analysis with smashing

Fixpoint iteration library

• precise interleaving between widening and narrowing
• extensible with thresholds

Efficient reusable data-structure

• simple API for integrating new abstract domains
RESULTS
SV-COMP 2015

4\textsuperscript{th} Competition on Software Verification held (here!) at TACAS 2015

Goals

• Provide a snapshot of the state-of-the-art in software verification to the community.
• Increase the visibility and credits that tool developers receive.
• Establish a set of benchmarks for software verification in the community.

Participants:

• Over 22 participants, including most popular Software Model Checkers and Bounded Model Checkers

Benchmarks:

• C programs with error location (programs include pointers, structures, etc.)
• Over 6,000 files, each 2K – 100K LOC
• Linux Device Drivers, Product Lines, Regressions/Tricky examples
• http://sv-comp.sosy-lab.org/2015/benchmarks.php
Results for DeviceDriver category

Time in s

Accumulated score
Detecting Buffer Overflow in Auto-pilot software

Show absence of Buffer Overflows in
- paparazzi and mnav autopilots

Automatically instrument buffer accesses with runtime checks
Use SeaHorn to validate that run-time checks never fail
- somewhat slower than pure abstract interpretation
- much more precise!

LLVM Pass to insert BO checks
Conclusion

SeaHorn ([http://seahorn.github.io](http://seahorn.github.io))

- a state-of-the-art Software Model Checker
- LLVM-based front-end
- CHC-based verification engine
- a framework for research in logic-based verification

The future

- making SeaHorn useful to users of verification technology
  - counterexamples, build integration, property specification, proofs, etc.
- targeting many existing CHC engines
  - specialize encoding and transformations to specific engines
  - communicate results between engines
- richer properties
  - termination, liveness, synthesis
Available postdoctoral positions

What: development and application of SeaHorn

Where: CMU/NASA Silicon Valley Campus

Contact:
Temegshen Kahsaiteimesghehnh.kausaizene@nasa.gov
Arie Gurfinkel arie@cmu.edu