Extending enumerative function synthesis via SMT-driven classification

Haniel Barbosa, Andrew Reynolds, Daniel Larraz, Cesare Tinelli

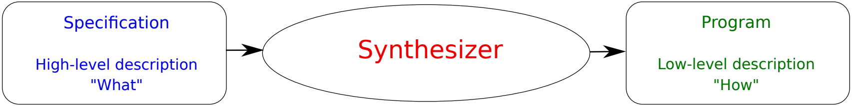
UFMG

The University of Iowa

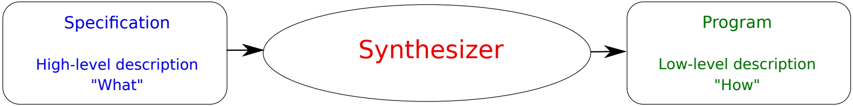
Lógicos em Quarentena

2020-04-30, The Internet

Program synthesis: "The Holy Grail of Computer Science"



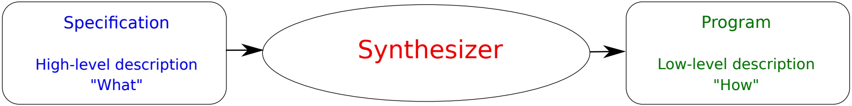
Program synthesis: “The Holy Grail of Computer Science”



In	Out
0	1
1	2
2	4
3	8

$$P(x) =$$

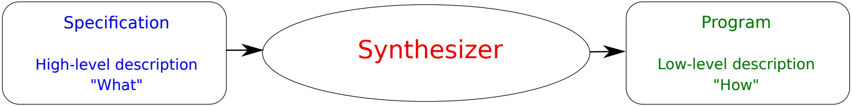
Program synthesis: “The Holy Grail of Computer Science”



In	Out
0	1
1	2
2	4
3	8

$$P(x) = \mathrm{ite}(x < 2,\, x + 1,\, \mathrm{ite}(x < 3,\, 2 * x,\, 2 * x + 2))$$

Program synthesis: "The Holy Grail of Computer Science"

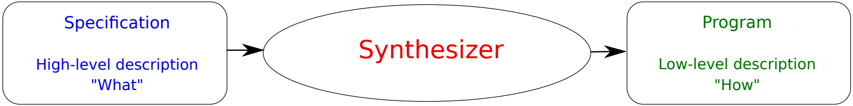


In	Out
0	1
1	2
2	4
3	8

$$P(x) = \mathrm{ite}(x < 2,\, x + 1,\, \mathrm{ite}(x < 3,\, 2 * x,\, 2 * x + 2))$$

$$P(x) = 2^x$$

Program synthesis: “The Holy Grail of Computer Science”



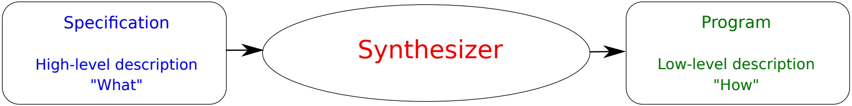
In	Out
0	1
1	2
2	4
3	8

$$P(x) = \mathrm{ite}(x < 2,\, x + 1,\, \mathrm{ite}(x < 3,\, 2 * x,\, 2 * x + 2))$$

$$P(x) = 2^x$$

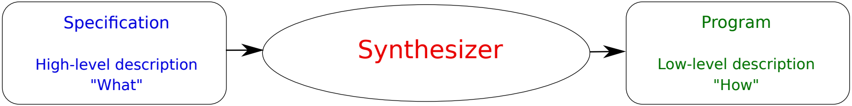
$$P(x) = \ldots$$

Program synthesis: “The Holy Grail of Computer Science”



In	Out
x=0,y=1	x=1,y=0
x=1,y=2	x=2,y=1
x=2,y=3	x=3,y=2
x=3,y=4	x=4,y=3

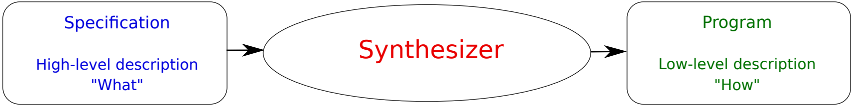
Program synthesis: “The Holy Grail of Computer Science”



In	Out
x=0,y=1	x=1,y=0
x=1,y=2	x=2,y=1
x=2,y=3	x=3,y=2
x=3,y=4	x=4,y=3

$$P(x, y) = \{x \leftarrow x + 1; y \leftarrow y - 1\}$$

Program synthesis: “The Holy Grail of Computer Science”

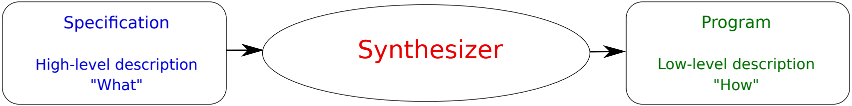


In	Out
x=0,y=1	x=1,y=0
x=1,y=2	x=2,y=1
x=2,y=3	x=3,y=2
x=3,y=4	x=4,y=3

$$P(x, y) = \{x \leftarrow x + 1; y \leftarrow y - 1\}$$

$$P(x, y) = \{z \leftarrow x; x \leftarrow y; y \leftarrow z\}$$

Program synthesis: "The Holy Grail of Computer Science"



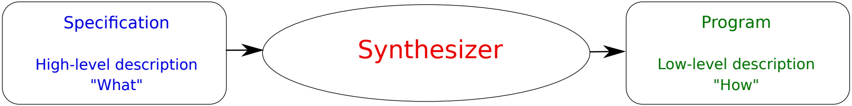
In	Out
x=0,y=1	x=1,y=0
x=1,y=2	x=2,y=1
x=2,y=3	x=3,y=2
x=3,y=4	x=4,y=3

$$P(x,y) = \{x \leftarrow x+1; y \leftarrow y-1\}$$

$$P(x,y) = \{z \leftarrow x; x \leftarrow y; y \leftarrow z\}$$

$$P(x,y) = \dots$$

Program synthesis: "The Holy Grail of Computer Science"



- Main challenges:
  - Exploring search space
  - Capturing intention

- Three main characteristics
  - How to write specification
  - How to constrain search space
  - How to guide the search

In	Out
x=0,y=1	x=1,y=0
x=1,y=2	x=2,y=1
x=2,y=3	x=3,y=2
x=3,y=4	x=4,y=3

$$P(x, y) = \{x \leftarrow x + 1; y \leftarrow y - 1\}$$

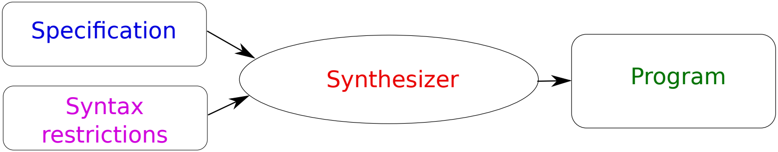
$$P(x, y) = \{z \leftarrow x; x \leftarrow y; y \leftarrow z\}$$

$$P(x, y) = \dots$$

Some applications of program synthesis

- ▷ Superoptimization [SSA13], [NRB+19], ...
- ▷ Program repair [NWK+17], [LCL+17], ...
- ▷ Programming by examples [Gul11], [FMG+17] ...
- ▷ Circuit synthesis [EWW16], ...
- ▷ Loop invariant synthesis [GLM+14], [PSM16], ...
- ▷ ...

Syntax-Guided Synthesis (SyGuS) [ABJ+13]



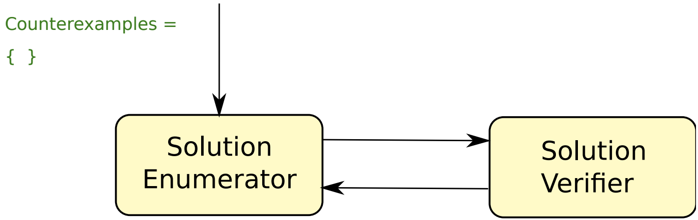
- ▷ Specification is given by (second-order) $\mathcal{T}$-formula: $\exists f.\, \forall \bar{x}.\, \varphi[f, \bar{x}]$
- ▷ Syntactic restrictions given by context-free grammar R

Enumerative CEGIS

Consider the example:

$$\varphi \quad = \quad f(x, x) \simeq x + 1 \wedge f(x, x+1) \simeq x$$

$$R \quad = \quad \begin{array}{l} A \to 0 \mid 1 \mid x \mid y \mid A + A \mid \mathrm{ite}(B, A, A) \\ B \to A \leq A \mid \neg B \end{array}$$



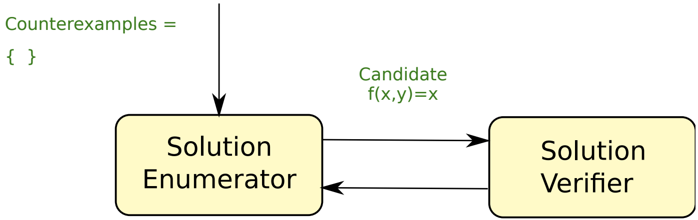
- De facto approach to SyGuS solving given its simplicity and efficacy

Enumerative CEGIS [STB+06; URD+13]

Consider the example:

$$\varphi \quad = \quad f(x, x) \simeq x + 1 \wedge f(x, x+1) \simeq x$$

$$R \quad = \quad \begin{array}{l} A \to 0 \mid 1 \mid x \mid y \mid A + A \mid \mathrm{ite}(B, A, A) \\ B \to A \leq A \mid \neg B \end{array}$$



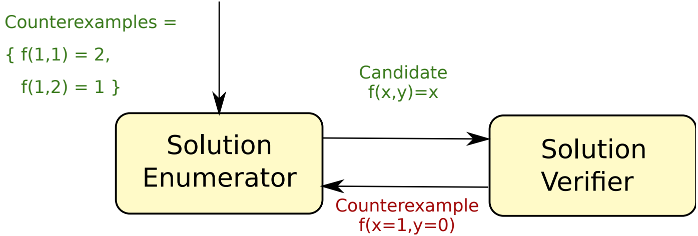
- ▷ De facto approach to SyGuS solving given its simplicity and efficacy

Enumerative CEGIS [STB+06; URD+13]

Consider the example:

$$\varphi \quad = \quad f(x, x) \simeq x + 1 \;\wedge\; f(x, x+1) \simeq x$$

$$R \quad = \quad \begin{array}{l} A \to 0 \mid 1 \mid x \mid y \mid A + A \mid \text{ite}(B, A, A) \\ B \to A \leq A \mid \neg B \end{array}$$



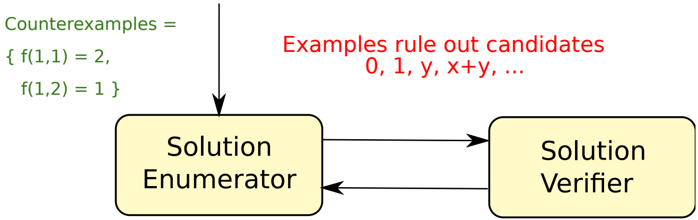
- ▷ De facto approach to SyGuS solving given its simplicity and efficacy

Enumerative CEGIS [STB+06; URD+13]

Consider the example:

$$\varphi \quad = \quad f(x, x) \simeq x + 1 \wedge f(x, x+1) \simeq x$$

$$R \quad = \quad \begin{array}{l} A \to 0 \mid 1 \mid x \mid y \mid A + A \mid \mathrm{ite}(B, A, A) \\ B \to A \leq A \mid \neg B \end{array}$$



▷ De facto approach to SyGuS solving given its simplicity and efficacy

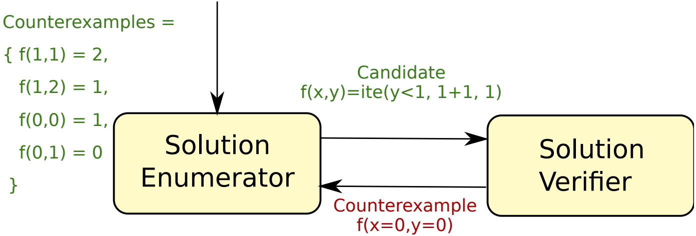
Enumerative CEGIS

[STB+06; URD+13]

Consider the example:

$$\varphi \quad = \quad f(x,\, x) \simeq x + 1 \,\wedge\, f(x,\, x+1) \simeq x$$

$$R \quad = \quad \begin{array}{l} A \to 0 \mid 1 \mid x \mid y \mid A + A \mid \mathrm{ite}(B,\, A,\, A) \\ B \to A \leq A \mid \neg B \end{array}$$



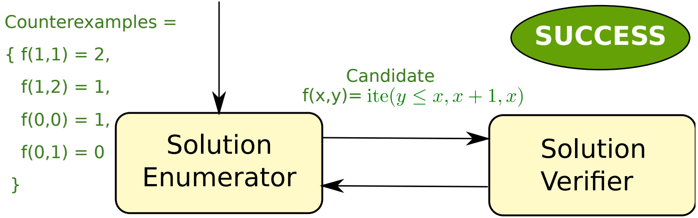
- ▷ De facto approach to SyGuS solving given its simplicity and efficacy

Enumerative CEGIS [STB+06; URD+13]

Consider the example:

$$\varphi \quad = \quad f(x,\, x) \simeq x + 1 \;\wedge\; f(x,\, x+1) \simeq x$$

$$R \quad = \quad \begin{array}{l} A \to 0 \mid 1 \mid x \mid y \mid A + A \mid \mathrm{ite}(B,\, A,\, A) \\ B \to A \leq A \mid \neg B \end{array}$$



▷ De facto approach to SyGuS solving given its simplicity and efficacy

Scalability issues

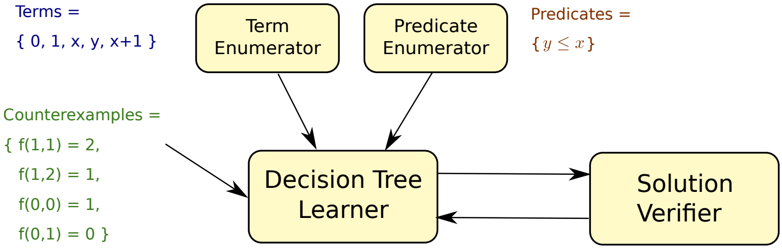
Enumerative techniques are effective but limited by the explosion of the enumeration space as term size increases

For this bit-vector grammar, enumerating

- Terms of size $= 1$: .05 seconds
- Terms of size $= 2$: .6 seconds
- Terms of size $= 3$: 48 seconds
- Terms of size $= 4$: 5.8 hours
- Terms of size $= 5$: ??? (100+ days)

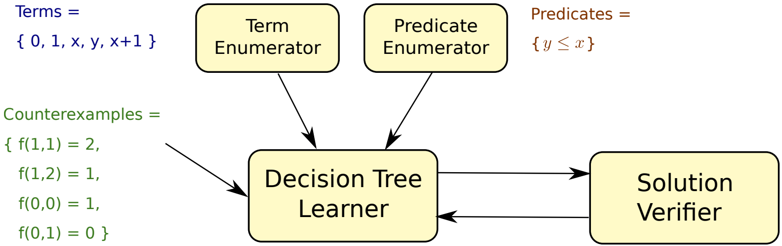
```
(synth-fun f ((s (BitVec 4))
              (t (BitVec 4)))
 (BitVec 4) (
 (Start (BitVec 4) (
  s t #x0
  (bvneg  Start)
  (bvnot  Start)
  (bvadd  Start Start)
  (bvmul  Start Start)
  (bvand  Start Start)
  (bvlshr Start Start)
  (bvor   Start Start)
  (bvshl  Start Start)))))
```

Divide and conquer (D&C) [ARU17; NSM18]

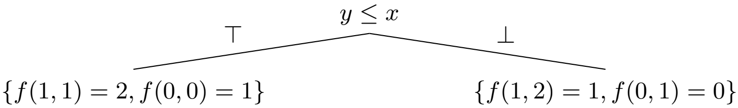


- ▷ Generate partial solutions correct on subset of input
- ▷ Unify partial solutions via decision tree learning

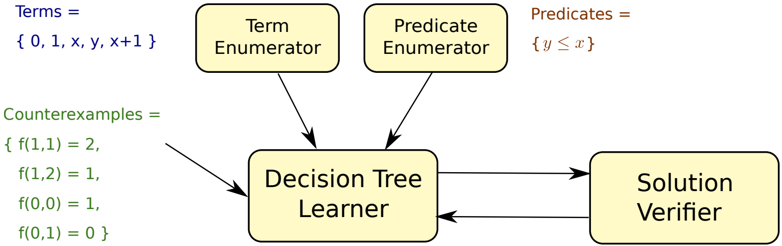
Divide and conquer (D&C) [ARU17; NSM18]



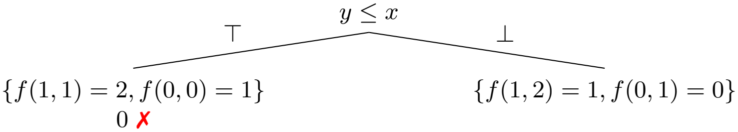
- ▷ Generate partial solutions correct on subset of input
- ▷ Unify partial solutions via decision tree learning



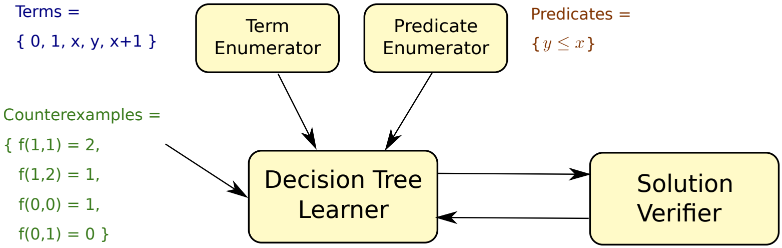
Divide and conquer (D&C) [ARU17; NSM18]



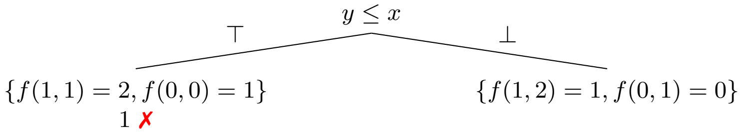
- ▷ Generate partial solutions correct on subset of input
- ▷ Unify partial solutions via decision tree learning



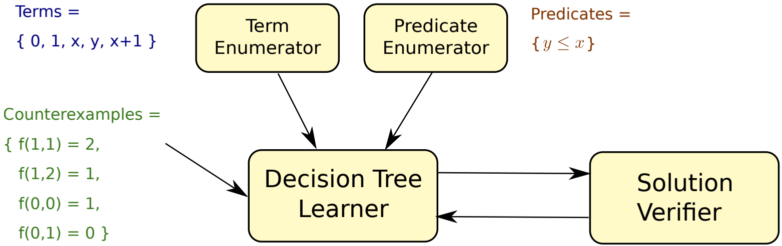
Divide and conquer (D&C) [ARU17; NSM18]



- Generate partial solutions correct on subset of input
- Unify partial solutions via decision tree learning



Divide and conquer (D&C) [ARU17; NSM18]



- ▷ Generate partial solutions correct on subset of input
- ▷ Unify partial solutions via decision tree learning

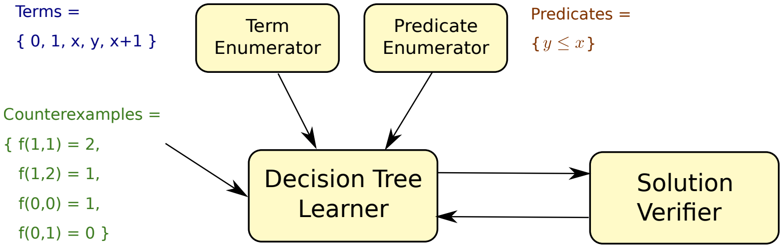
$$y \le x$$

$\top$ $\bot$

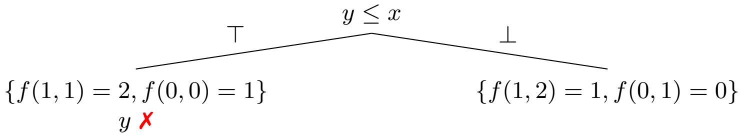
$\{f(1,1) = 2, f(0,0) = 1\}$ $\{f(1,2) = 1, f(0,1) = 0\}$

x ✗

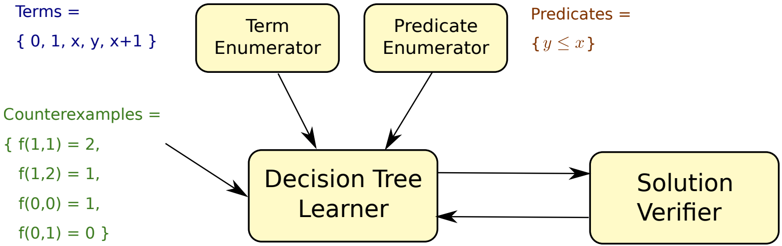
Divide and conquer (D&C) [ARU17; NSM18]



- Generate partial solutions correct on subset of input
- Unify partial solutions via decision tree learning



Divide and conquer (D&C) [ARU17; NSM18]



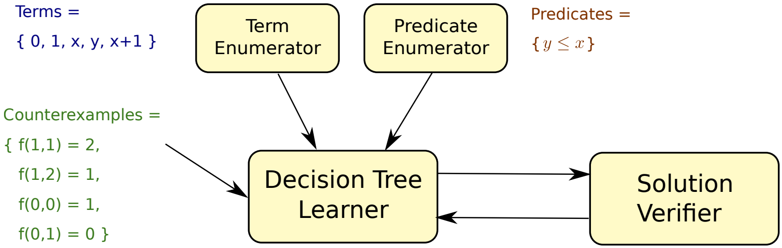
- Generate partial solutions correct on subset of input
- Unify partial solutions via decision tree learning

$$y \leq x$$

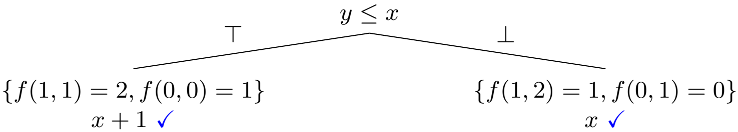
$\top$: $\{f(1,1) = 2, f(0,0) = 1\}$ — $x + 1$ ✓

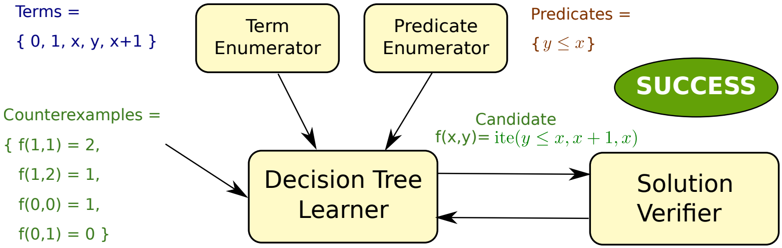
$\bot$: $\{f(1,2) = 1, f(0,1) = 0\}$

Divide and conquer (D&C) [ARU17; NSM18]

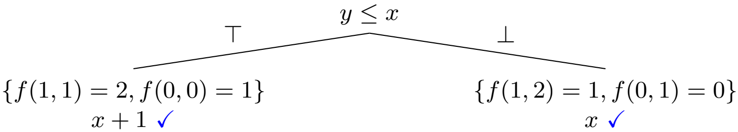


- ▷ Generate partial solutions correct on subset of input
- ▷ Unify partial solutions via decision tree learning





- Generate partial solutions correct on subset of input
- Unify partial solutions via decision tree learning



- D&C provides much better scalability

However...

- D&C can only be applied to point-wise specifications
  - Each input valuation is specified independently

However...

- D&C can only be applied to point-wise specifications
  - Each input valuation is specified independently

Consider augmenting the previous example:

$$\begin{aligned}\varphi = \quad & f(x, x) \simeq x+1 \land f(x, x+1) \simeq x \\ \land \quad & f(x, y) \simeq x+1 \Rightarrow f(x+2, y) \simeq x\end{aligned}$$

Counterexample $\{x \mapsto 1, y \mapsto 0\}$ yields the constraints:

$$f(1, 1) \simeq 2 \quad \land \quad f(1, 2) \simeq 1 \quad \land \quad f(1, 0) \simeq 2 \Rightarrow f(3, 0) \simeq 1$$

- A solution for $f(1, 0)$ restricts the solution for $f(3, 0)$
- Breaks assumption that partial solutions can be found *indepedently*

Challenges

- This limitation excludes interesting classes of synthesis problems
  - Invariants: $I(x) \wedge T(x, x') \Rightarrow I(x')$
  - Ranking functions: $rank(x') < rank(x)$
  - Modular arithmetic functions: $f(x) \simeq f(x+n)$
  - ...
- Extending D&C to arbitrary (non-point-wise) specifications:
  - Find a term assignment consistent with point dependencies
  - Correctly classify points according to term assignment

Challenges

- This limitation excludes interesting classes of synthesis problems
  - Invariants: $I(x) \wedge T(x, x') \Rightarrow I(x')$
  - Ranking functions: $rank(x') < rank(x)$
  - Modular arithmetic functions: $f(x) \simeq f(x + n)$
  - ...
- Extending D&C to arbitrary (non-point-wise) specifications:
  - Find a term assignment consistent with point dependencies

SMT solving

  - Correctly classify points according to term assignment

Challenges

- This limitation excludes interesting classes of synthesis problems
  - Invariants: $I(x) \wedge T(x, x') \Rightarrow I(x')$
  - Ranking functions: $rank(x') < rank(x)$
  - Modular arithmetic functions: $f(x) \simeq f(x + n)$
  - ...
- Extending D&C to arbitrary (non-point-wise) specifications:
  - Find a term assignment consistent with point dependencies

SMT solving

  - Correctly classify points according to term assignment

Decision tree learning

- SMT-based solution-complete strategy
- Heuristic strategy

SMT solving for SyGuS

Satisfiability Modulo Theories (SMT)

First-order formulas in CNF:

$$\begin{array}{rcl} t & ::= & x \mid f(t,\ldots,t) \\ \varphi & ::= & p(t,\ldots,t) \mid \neg\varphi \mid \varphi \vee \varphi \mid \forall x_1 \ldots x_n.\, \varphi \end{array}$$

Given a formula φ in FOL and background theories $\mathcal{T}_1, \ldots, \mathcal{T}_n$, finding a model $\mathcal{M}$ giving an *interpretation* to all terms and predicates such that $\mathcal{M} \models_{\mathcal{T}_1,\ldots,\mathcal{T}_n} \varphi$

Satisfiability Modulo Theories (SMT)

First-order formulas in CNF:

$$
\begin{array}{lcl}
t & ::= & x \mid f(t, \ldots, t) \\
\varphi & ::= & p(t, \ldots, t) \mid \neg\varphi \mid \varphi \vee \varphi \mid \forall x_1 \ldots x_n.\, \varphi
\end{array}
$$

Given a formula φ in FOL and background theories $\mathcal{T}_1, \ldots, \mathcal{T}_n$, finding a model $\mathcal{M}$ giving an *interpretation* to all terms and predicates such that $\mathcal{M} \models_{\mathcal{T}_1, \ldots, \mathcal{T}_n} \varphi$

Example

Is φ satisfiable modulo *equality* and *arithmetic*?

$$\varphi \;=\; (x_1 \geq 0) \wedge (x_1 < 1) \quad \wedge \quad (f(x_1) \not\simeq f(0) \quad \vee \quad x_3 + x_1 > x_3 + 1)$$

Satisfiability Modulo Theories (SMT)

First-order formulas in CNF:

$$\begin{array}{rcl} t & ::= & x \mid f(t, \ldots, t) \\ \varphi & ::= & p(t, \ldots, t) \mid \neg\varphi \mid \varphi \vee \varphi \mid \forall x_1 \ldots x_n .\, \varphi \end{array}$$

Given a formula φ in FOL and background theories $\mathcal{T}_1, \ldots, \mathcal{T}_n$, finding a model $\mathcal{M}$ giving an *interpretation* to all terms and predicates such that $\mathcal{M} \models_{\mathcal{T}_1, \ldots, \mathcal{T}_n} \varphi$

Example

Is φ satisfiable modulo *equality* and *arithmetic*?

$$\varphi \quad = \quad \underbrace{(x_1 \geq 0) \wedge (x_1 < 1)}_{\text{LIA}} \quad \wedge \quad (\underbrace{f(x_1) \not\simeq f(0)}_{\text{EUF}} \quad \vee \quad \underbrace{x_3 + x_1 > x_3 + 1}_{\text{LIA}})$$

Satisfiability Modulo Theories (SMT)

First-order formulas in CNF:

$$\begin{array}{lcl} t & ::= & x \mid f(t,\ldots,t) \\ \varphi & ::= & p(t,\ldots,t) \mid \neg\varphi \mid \varphi \vee \varphi \mid \forall x_1 \ldots x_n.\, \varphi \end{array}$$

Given a formula φ in FOL and background theories $\mathcal{T}_1, \ldots, \mathcal{T}_n$, finding a model $\mathcal{M}$ giving an *interpretation* to all terms and predicates such that $\mathcal{M} \models_{\mathcal{T}_1,\ldots,\mathcal{T}_n} \varphi$

Example

Is φ satisfiable modulo *equality* and *arithmetic*?

$$\varphi \quad = \quad \underbrace{(x_1 \geq 0) \wedge (x_1 < 1)}_{\text{LIA}} \quad \wedge \quad (\underbrace{f(x_1) \not\simeq f(0)}_{\text{EUF}}) \quad \vee \quad \underbrace{x_3 + x_1 > x_3 + 1}_{\text{LIA}})$$

$$\varphi \quad \models_{\text{LIA}} \quad x_1 \simeq 0$$

Satisfiability Modulo Theories (SMT)

First-order formulas in CNF:

$$\begin{array}{lcl} t & ::= & x \mid f(t, \ldots, t) \\ \varphi & ::= & p(t, \ldots, t) \mid \neg\varphi \mid \varphi \vee \varphi \mid \forall x_1 \ldots x_n.\, \varphi \end{array}$$

Given a formula φ in FOL and background theories $\mathcal{T}_1, \ldots, \mathcal{T}_n$, finding a model $\mathcal{M}$ giving an *interpretation* to all terms and predicates such that $\mathcal{M} \models_{\mathcal{T}_1, \ldots, \mathcal{T}_n} \varphi$

Example

Is φ satisfiable modulo *equality* and *arithmetic*?

$$\varphi \;=\; \underbrace{(x_1 \geq 0) \wedge (x_1 < 1)}_{\text{LIA}} \;\wedge\; (\underbrace{f(x_1) \not\simeq f(0)}_{\text{EUF}} \;\vee\; \underbrace{x_3 + x_1 > x_3 + 1}_{\text{LIA}})$$

$$\begin{array}{rcl} \varphi & \models_{\text{LIA}} & x_1 \simeq 0 \\ x_1 \simeq 0 & \models_{\text{EUF}} & f(x_1) \simeq f(0) \end{array}$$

Satisfiability Modulo Theories (SMT)

First-order formulas in CNF:

$$\begin{array}{lcl} t & ::= & x \mid f(t,\ldots,t) \\ \varphi & ::= & p(t,\ldots,t) \mid \neg\varphi \mid \varphi \vee \varphi \mid \forall x_1 \ldots x_n.\, \varphi \end{array}$$

Given a formula φ in FOL and background theories $\mathcal{T}_1, \ldots, \mathcal{T}_n$, finding a model $\mathcal{M}$ giving an *interpretation* to all terms and predicates such that $\mathcal{M} \models_{\mathcal{T}_1,\ldots,\mathcal{T}_n} \varphi$

Example

Is φ satisfiable modulo *equality* and *arithmetic*?

$$\varphi \quad = \quad \underbrace{(x_1 \geq 0) \wedge (x_1 < 1)}_{\text{LIA}} \quad \wedge \quad (\underbrace{f(x_1) \not\simeq f(0)}_{\text{EUF}} \quad \vee \quad \underbrace{x_3 + x_1 > x_3 + 1}_{\text{LIA}})$$

$$\begin{array}{rcl} \varphi & \models_{\text{LIA}} & x_1 \simeq 0 \\ x_1 \simeq 0 & \models_{\text{EUF}} & f(x_1) \simeq f(0) \\ x_1 \simeq 0 & \models_{\text{LIA}} & x_3 + x_1 \not> x_3 + 1 \end{array}$$

Satisfiability Modulo Theories (SMT)

First-order formulas in CNF:

$$\begin{array}{lcl} t & ::= & x \mid f(t,\ldots,t) \\ \varphi & ::= & p(t,\ldots,t) \mid \neg\varphi \mid \varphi \vee \varphi \mid \forall x_1 \ldots x_n.\ \varphi \end{array}$$

Given a formula φ in FOL and background theories $\mathcal{T}_1, \ldots, \mathcal{T}_n$, finding a model $\mathcal{M}$ giving an *interpretation* to all terms and predicates such that $\mathcal{M} \models_{\mathcal{T}_1,\ldots,\mathcal{T}_n} \varphi$

Example

Is φ satisfiable modulo *equality* and *arithmetic*?

$$\varphi \quad = \quad \underbrace{(x_1 \geq 0) \wedge (x_1 < 1)}_{\text{LIA}} \quad \wedge \quad (\underbrace{f(x_1) \not\simeq f(0)}_{\text{EUF}} \quad \vee \quad \underbrace{x_3 + x_1 > x_3 + 1}_{\text{LIA}})$$

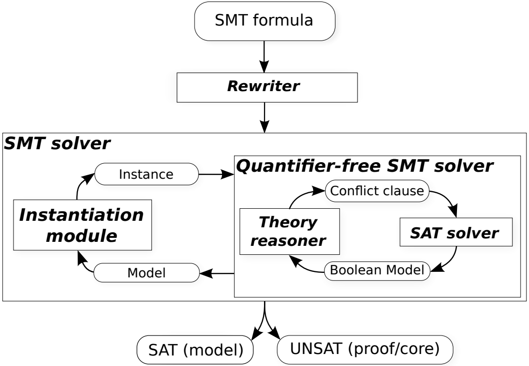
$$\begin{array}{rll} \varphi & \models_{\text{LIA}} & x_1 \simeq 0 \\ x_1 \simeq 0 & \models_{\text{EUF}} & f(x_1) \simeq f(0) \\ x_1 \simeq 0 & \models_{\text{LIA}} & x_3 + x_1 \not> x_3 + 1 \end{array}$$

Therefore $\models_{\text{EUF}\cup\text{LIA}} \neg\varphi$

SMT solving [BT18]

- Decidability depends on the theories being used
- Efficient decision procedures
  - Equality and uninterpreted functions (Congruence Closure (CC)) [NO80], [DST80]
  - Algebraic datatypes (CC + Injectivity, Distinctness, Exhaustiveness, Acyclicity) [BST07]
  - Linear arithmetic (Simplex) [DM06]
  - Bit-vectors (Bit-blasting)
  - Combination of theories (Nelson-Oppen)
  - ...
- Boolean search leverages SAT solvers
- Users may define their own theories
  - New operators as uninterpreted functions + Axioms

CDCL($\mathcal{T}$) architecture



- Rewriter simplifies terms

  $x + 0 \to x$ $\quad$ $\mathsf{a} \not\approx \mathsf{a} \to \bot$ $\quad$ $(\text{str.replace}\ x\ (\text{str.++}\ x\ x)\ y) \to x$
- SAT solver enumerates models for Boolean skeleton of formula
- Theory solvers check consistency in the theory
- Instantiation module selects relevant instances

Enumerative SyGuS in SMT [RKT+17], [RVB+18]

- Encode problem using a deep embedding into datatypes

$$\varphi \quad = \quad f(x,\, x) \simeq x + 1 \,\wedge\, f(x,\, x+1) \simeq x$$

$$R \quad = \quad \begin{array}{l} A \to 0 \mid 1 \mid x \mid y \mid A + A \mid \mathrm{ite}(B,\, A,\, A) \\ B \to A \leq A \mid \neg B \end{array}$$

  Becomes

$$\llbracket \varphi \rrbracket \quad = \quad \mathrm{eval}_{\mathrm{a}}(d, x, x) \simeq x + 1 \,\wedge\, \mathrm{eval}_{\mathrm{a}}(d, x, x+1) \simeq x$$

$$\llbracket R \rrbracket \quad = \quad \begin{array}{ll} \mathrm{a} & = \mathrm{Zero} \mid \mathrm{One} \mid \mathrm{X} \mid \mathrm{Y} \mid \mathrm{Plus(a, a)} \mid \mathrm{Ite(b,\, a,\, a)} \\ \mathrm{b} & = \mathrm{Leq(a,\, a)} \mid \mathrm{Neg(b)} \end{array}$$

- eval maps datatype terms to their corresponding theory terms
  - $\mathrm{eval}_{\mathrm{a}}(\mathrm{Plus}(\mathrm{X}, \mathrm{X}), 2, 3)$ is interpreted as $(x + x)\{x \mapsto 2, y \mapsto 3\} = 4$

Enumerative SyGuS in SMT [RKT+17], [RVB+18]

- Encode problem using a deep embedding into datatypes

$$\varphi \quad = \quad f(x,\, x) \simeq x + 1 \,\wedge\, f(x,\, x+1) \simeq x$$

$$R \quad = \quad \begin{array}{l} A \rightarrow 0 \mid 1 \mid x \mid y \mid A + A \mid \mathrm{ite}(B,\, A,\, A) \\ B \rightarrow A \leq A \mid \neg B \end{array}$$

Becomes

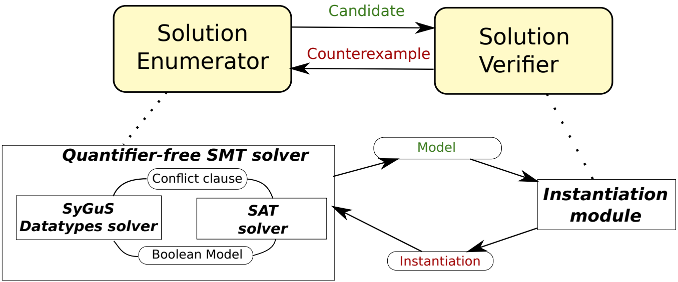
$$\llbracket\varphi\rrbracket \quad = \quad \mathrm{eval}_{\mathrm{a}}(d, x, x) \simeq x + 1 \,\wedge\, \mathrm{eval}_{\mathrm{a}}(d, x, x+1) \simeq x$$

$$\llbracket R\rrbracket \quad = \quad \begin{array}{ll} \mathrm{a} & = \mathrm{Zero} \mid \mathrm{One} \mid \mathrm{X} \mid \mathrm{Y} \mid \mathrm{Plus}(\mathrm{a}, \mathrm{a}) \mid \mathrm{Ite}(\mathrm{b},\, \mathrm{a},\, \mathrm{a}) \\ \mathrm{b} & = \mathrm{Leq}(\mathrm{a},\, \mathrm{a}) \mid \mathrm{Neg}(\mathrm{b}) \end{array}$$

- eval maps datatype terms to their corresponding theory terms
  - $\mathrm{eval}_{\mathrm{a}}(\mathrm{Plus}(\mathrm{X}, \mathrm{X}), 2, 3)$ is interpreted as $(x + x)\{x \mapsto 2, y \mapsto 3\} = 4$

- A solution is a model in which e.g.
  - $d = \mathrm{Ite}(\mathrm{Leq}(\mathrm{Y},\, \mathrm{X}),\, \mathrm{Plus}(\mathrm{X}, \mathrm{One}),\, \mathrm{X})$, corresponding to
  - $f = \lambda xy.\, \mathrm{ite}(y \leq x,\, x + 1,\, x)$

Enumerative SyGuS in SMT [RKT+17], [RVB+18]



- An instantiation module checks candidates against the specification
  - Generates lemmas witnessing why a candidate failed
- A specialized datatypes solver for SyGuS generates candidate solutions
  - Must satisfy all lemmas
  - Dedicated pruning
  - Parameterizable fairness criteria for enumeration

Unif+PI: a general divide-and-conquer framework for SyGuS solving

Recapping

- D&C can only be applied to point-wise specifications
  - Each input valuation is specified independently
- Extending D&C to arbitrary (non-point-wise) specifications requires:
  - Find a term assignment consistent with point dependencies

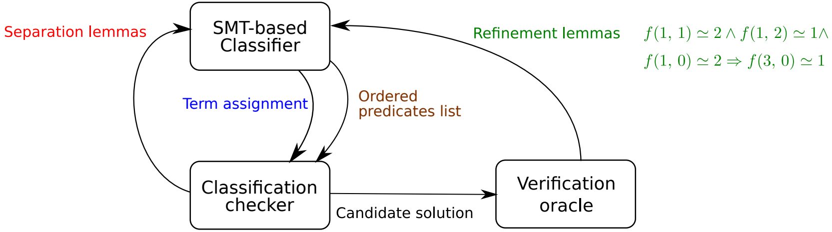
SMT solving

  - Correctly classify points according to term assignment

Decision tree learning

- SMT-based solution-complete strategy
- Heuristic strategy

Unif+PI: Synthesis via Pointwise-Indepentent unification

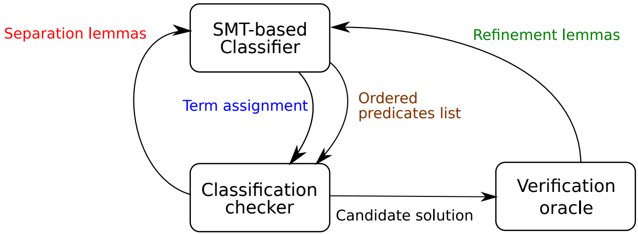


- *SMT-based classifier*
  - Assigns terms to points so that lemmas hold

$$f(1,\,1) \mapsto y + y, \quad \{f(1,\,0),\, f(3,\,0),\, f(1,\,2)\} \mapsto x$$

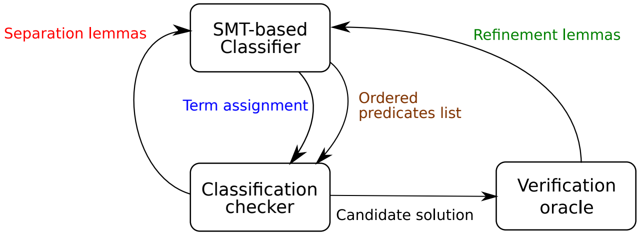
  - Generates ordered list of predicates to *separate* points: $P_1 \mapsto x \neq y$
- *Classification checker*: whether corresponding decision tree correctly classifies sample
  - Failures are encoded as *separation lemmas*

Unif+PI: Synthesis via Pointwise-Indepentent unification



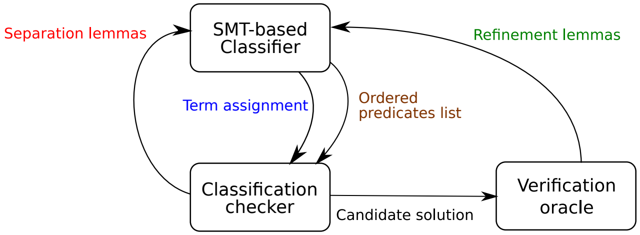
▷ Successful candidates that are not verified lead to refinement lemmas and the learning restarts

Unif+PI: Synthesis via Pointwise-Indepentent unification



- Successful candidates that are not verified lead to refinement lemmas and the learning restarts
- Bounded *solution-completeness* and *minimality* results due to exhaustive enumeration of possible classifiers according to
  - *size* and *number* of distinct terms to be assigned
  - *size* and *number* of distinct predicates

Unif+PI: Synthesis via Pointwise-Indepentent unification



- Successful candidates that are not verified lead to refinement lemmas and the learning restarts
- Bounded *solution-completeness* and *minimality* results due to exhaustive enumeration of possible classifiers according to
  - *size* and *number* of distinct terms to be assigned
  - *size* and *number* of distinct predicates
- Our fairness criteria are $size = log_2(\#terms)$, $\#pred = \#terms - 1$

Consider again:

$$\begin{aligned} \varphi = & \quad f(x, x) \simeq x + 1 \land f(x, x+1) \simeq x \\ \land & \quad f(x, y) \simeq x + 1 \Rightarrow f(x+2, y) \simeq x \end{aligned}$$

- Initially a single term of size 0 will be a trivial successful classifier

Consider again:

$$\begin{aligned}\varphi = \quad & f(x,\,x) \simeq x+1 \,\wedge\, f(x,\,x+1) \simeq x \\ \wedge \quad & f(x,\,y) \simeq x+1 \Rightarrow f(x+2,\,y) \simeq x\end{aligned}$$

- Initially a single term of size 0 will be a trivial successful classifier
- Refinement lemma:

$$f(1,\,1) \simeq 2 \quad \wedge \quad f(1,\,0) \simeq 2 \Rightarrow f(3,\,0) \simeq 1 \quad \wedge \quad f(1,\,2) \simeq 1$$

Consider again:

$$\begin{aligned} \varphi = \quad & f(x,\, x) \simeq x + 1 \wedge f(x,\, x+1) \simeq x \\ \wedge \quad & f(x,\, y) \simeq x + 1 \Rightarrow f(x+2,\, y) \simeq x \end{aligned}$$

- Initially a single term of size 0 will be a trivial successful classifier
- Refinement lemma:

$$f(1,\, 1) \simeq 2 \quad \wedge \quad f(1,\, 0) \simeq 2 \Rightarrow f(3,\, 0) \simeq 1 \quad \wedge \quad f(1,\, 2) \simeq 1$$

- Since no assignment with *a single* term suffices, the threshold is increased to consider two distinct terms
  - Maximum size increases to 1 and up to 1 predicate can be used

Consider again:

$$\begin{aligned}\varphi = \quad & f(x,\, x) \simeq x+1 \wedge f(x,\, x+1) \simeq x \\ \wedge \quad & f(x,\, y) \simeq x+1 \Rightarrow f(x+2,\, y) \simeq x\end{aligned}$$

- ▷ Initially a single term of size 0 will be a trivial successful classifier
- ▷ Refinement lemma:

$$f(1,\,1) \simeq 2 \quad \wedge \quad f(1,\,0) \simeq 2 \Rightarrow f(3,\,0) \simeq 1 \quad \wedge \quad f(1,\,2) \simeq 1$$

- ▷ Since no assignment with *a single* term suffices, the threshold is increased to consider two distinct terms
  - ▶ Maximum size increases to 1 and up to 1 predicate can be used
- ▷ A candidate classifier is

$$f(1,\,1) \mapsto y+y, \quad \{f(1,\,0),\, f(3,\,0),\, f(1,\,2)\} \mapsto x$$
$$P_1 \mapsto \top$$

Consider again:

$$\begin{aligned}\varphi = \quad & f(x, x) \simeq x + 1 \wedge f(x, x+1) \simeq x \\ \wedge \quad & f(x, y) \simeq x + 1 \Rightarrow f(x+2, y) \simeq x\end{aligned}$$

- Initially a single term of size 0 will be a trivial successful classifier
- Refinement lemma:

$$f(1, 1) \simeq 2 \quad \wedge \quad f(1, 0) \simeq 2 \Rightarrow f(3, 0) \simeq 1 \quad \wedge \quad f(1, 2) \simeq 1$$

- Since no assignment with *a single* term suffices, the threshold is increased to consider two distinct terms
  - Maximum size increases to 1 and up to 1 predicate can be used
- A candidate classifier is

$$f(1, 1) \mapsto y + y, \quad \{f(1, 0), f(3, 0), f(1, 2)\} \mapsto x$$
$$P_1 \mapsto \top$$

- This classifier fails on the sample, yielding a separation lemma

$$P_1 \simeq \top \Rightarrow f(1, 1) \simeq f(1, 0)$$

$$\begin{array}{rcl} \varphi_R & = & f(1,\,1) \simeq 2 \quad \wedge \quad f(1,\,0) \simeq 2 \Rightarrow f(3,\,0) \simeq 1 \quad \wedge \quad f(1,\,2) \simeq 1 \\ \varphi_S & = & P_1 \simeq \top \Rightarrow f(1,\,1) \simeq f(1,\,0) \end{array}$$

- Given this constraints and current threshold the next candidate classifier produced is:

$$\{f(1,\,1),\, f(1,\,0),\, f(3,\,0)\} \mapsto y+1, \quad f(1,\,2) \mapsto 1$$
$$P_1 \mapsto y \leq x$$

$$\varphi_R \;=\; f(1,\,1) \simeq 2 \quad \wedge \quad f(1,\,0) \simeq 2 \Rightarrow f(3,\,0) \simeq 1 \quad \wedge \quad f(1,\,2) \simeq 1$$
$$\varphi_S \;=\; P_1 \simeq \top \Rightarrow f(1,\,1) \simeq f(1,\,0)$$

- Given this constraints and current threshold the next candidate classifier produced is:

$$\{f(1,\,1),\, f(1,\,0),\, f(3,\,0)\} \mapsto y+1, \quad f(1,\,2) \mapsto 1$$
$$P_1 \mapsto y \le x$$

- Running the classification checker:

$f(1,1)$

$$\begin{array}{lcl}\varphi_R & = & f(1,\,1) \simeq 2 \quad \wedge \quad f(1,\,0) \simeq 2 \Rightarrow f(3,\,0) \simeq 1 \quad \wedge \quad f(1,\,2) \simeq 1 \\ \varphi_S & = & P_1 \simeq \top \Rightarrow f(1,\,1) \simeq f(1,\,0)\end{array}$$

- Given this constraints and current threshold the next candidate classifier produced is:

$$\{f(1,\,1),\,f(1,\,0),\,f(3,\,0)\} \mapsto y+1, \quad f(1,\,2) \mapsto 1$$
$$P_1 \mapsto y \leq x$$

- Running the classification checker:

$f(1,1),\,f(1,0)$

$$\begin{array}{rcl} \varphi_R & = & f(1,\,1) \simeq 2 \quad \wedge \quad f(1,\,0) \simeq 2 \Rightarrow f(3,\,0) \simeq 1 \quad \wedge \quad f(1,\,2) \simeq 1 \\ \varphi_S & = & P_1 \simeq \top \Rightarrow f(1,\,1) \simeq f(1,\,0) \end{array}$$

- Given this constraints and current threshold the next candidate classifier produced is:

$$\{f(1,\,1),\, f(1,\,0),\, f(3,\,0)\} \mapsto y+1, \quad f(1,\,2) \mapsto 1$$
$$P_1 \mapsto y \leq x$$

- Running the classification checker:

$f(1,1),\, f(1,0),\, f(3,0)$

$$\begin{aligned}\varphi_R &= f(1,\,1) \simeq 2 \quad \wedge \quad f(1,\,0) \simeq 2 \Rightarrow f(3,\,0) \simeq 1 \quad \wedge \quad f(1,\,2) \simeq 1 \\ \varphi_S &= P_1 \simeq \top \Rightarrow f(1,\,1) \simeq f(1,\,0)\end{aligned}$$

- Given this constraints and current threshold the next candidate classifier produced is:

$$\{f(1,\,1),\, f(1,\,0),\, f(3,\,0)\} \mapsto y+1, \quad f(1,\,2) \mapsto 1$$
$$P_1 \mapsto y \leq x$$

- Running the classification checker:

$f(1,1),\, f(1,0),\, f(3,0) \diamond f(1,2)$

$$\begin{aligned} \varphi_R &= f(1,\,1) \simeq 2 \quad \wedge \quad f(1,\,0) \simeq 2 \Rightarrow f(3,\,0) \simeq 1 \quad \wedge \quad f(1,\,2) \simeq 1 \\ \varphi_S &= P_1 \simeq \top \Rightarrow f(1,\,1) \simeq f(1,\,0) \end{aligned}$$

- Given this constraints and current threshold the next candidate classifier produced is:

$$\{f(1,\,1),\, f(1,\,0),\, f(3,\,0)\} \mapsto y+1, \quad f(1,\,2) \mapsto 1$$
$$P_1 \mapsto y \leq x$$

- Running the classification checker:

$$f(1,1),\, f(1,0),\, f(3,0) \diamond f(1,2) \;\to\; \begin{array}{ccc} & y \leq x & \\ \top \swarrow & & \searrow \bot \\ f(1,1),\, f(1,0),\, f(3,0) & & f(1,2) \end{array}$$

$$\begin{aligned}\varphi_R &= f(1,\,1) \simeq 2 \quad \wedge \quad f(1,\,0) \simeq 2 \Rightarrow f(3,\,0) \simeq 1 \quad \wedge \quad f(1,\,2) \simeq 1 \\ \varphi_S &= P_1 \simeq \top \Rightarrow f(1,\,1) \simeq f(1,\,0)\end{aligned}$$

- Given this constraints and current threshold the next candidate classifier produced is:

$$\{f(1,\,1),\, f(1,\,0),\, f(3,\,0)\} \mapsto y+1, \quad f(1,\,2) \mapsto 1$$
$$P_1 \mapsto y \leq x$$

- Running the classification checker:

$$f(1,1),\, f(1,0),\, f(3,0) \diamond f(1,2) \;\rightarrow$$

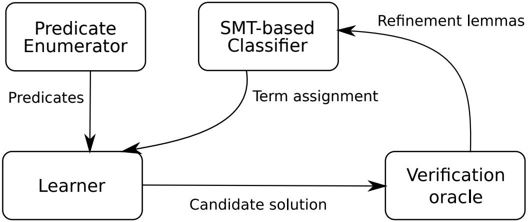
$$y \leq x$$

$\top$ $\qquad$ $\bot$

$$f(1,1),\, f(1,0),\, f(3,0) \qquad\qquad f(1,2)$$

- As the classification succeeds, a candidate is generated
- The candidate fails, so the process restarts with new refinement lemmas
- Eventually finds solution $f = \lambda xy.\, \mathrm{ite}(x \leq y,\, \mathrm{ite}(y \leq x,\, x+1,\, x),\, y)$

Unif+PI with unconstrained predicate enumeration



- Unif+PI+E uses SMT solver only to produce term assignments
  - Relies on standard decision tree learning to classify a labeled sample
  - Predicates chosen from enumerated pool with information-gain heuristic
  - Separation conflicts solved when new predicates are enumerated

- Often sacrificing completeness and minimality allows problems to be solved more efficiently

Experimental results

Setup

- Benchmarks (all over LIA)
  - 127 invariant synthesis benchmarks from SyGuS-COMP'18
  - 440 invariant synthesis benchmarks from test suite of Kind 2
- Three configurations of CVC4SY

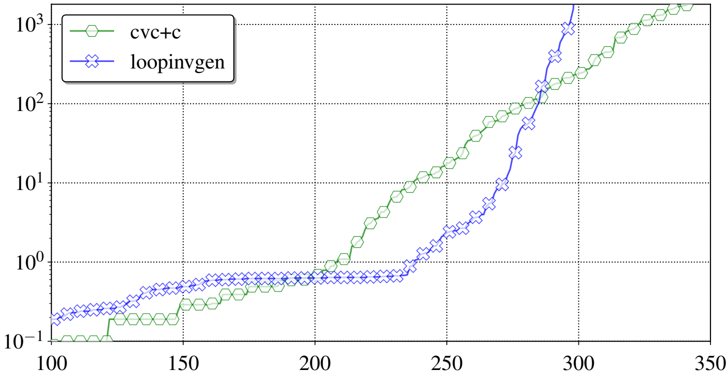
CVC+C	enumerative CEGIS [RBN+19]
CVC+UPI	Unif+PI
CVC+UPI+E	Unif+PI+E

- LOOPINVGEN [PM17] and CVC+C as baselines
- 1800s timeout, 8gb RAM

Full data at `http://cvc4.cs.stanford.edu/papers/FMCAD2019-UnifPI/`

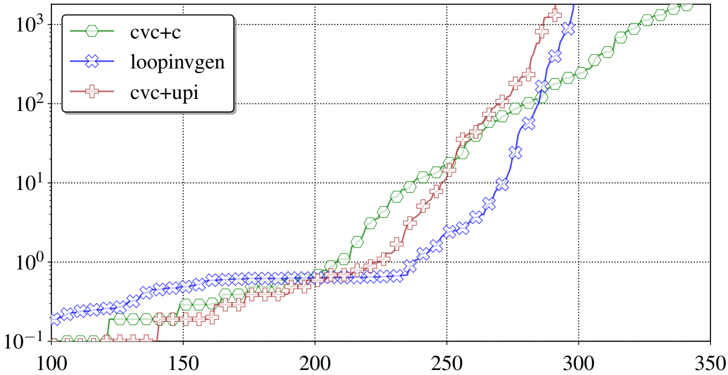
Summary

	Solved	Unique	Total time	Fastest	Shortest
CVC+C	**341**	30	436251s	245	259
LOOPINVGEN	298	7	433273s	261	**289**



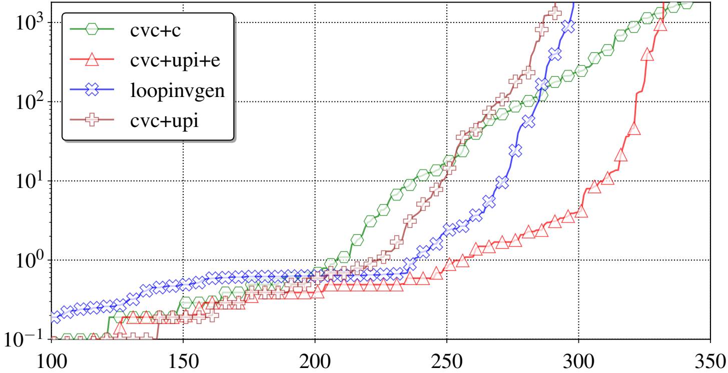
Summary

	Solved	Unique	Total time	Fastest	Shortest
CVC+C	**341**	30	436251s	245	259
CVC+UPI	291	3	494534s	236	231
LOOPINVGEN	298	7	433273s	261	**289**



Summary

	Solved	Unique	Total time	Fastest	Shortest
CVC+C	**341**	30	436251s	245	259
CVC+UPI+E	332	**47**	**414356**s	**306**	222
CVC+UPI	291	3	494534s	236	231
LOOPINVGEN	298	7	433273s	261	**289**
CVC-PORT	400	-	31476s	379	306



Advantages and disadvantages of Unif+PI

- CVC+UPI and CVC+UPI+E thrive when invariants can be built from combination of small literals
- CVC+C is superior when invariant is a single complex literal
  - 29 of its 30 unique solves are such cases
- CVC+UPI and CVC+UPI+E also suffer from dependence on samples
  - Sometimes search is biased towards simple classifiers when only a more complex one would suffice

Inv Track (829)

Solver	Solved	Fastest	Smallest	Score
CVC4-su	**592**	423	264	**4493**
LoopInvGen	512	**442**	**364**	4250
LoopInvGen-gplearn	511	411	349	4137
CVC4-Fast	522	319	243	3810
CVC4-Smart	539	283	260	3804
OASIS	538	20	317	3067
DryadSynth	277	161	39	1907

- ▷ 829 benchmarks from the literature in loop invariant synthesis
- ▷ 3600s timeout

Injecting some welcome realism

- Kind 2 employs in cooperation:
  - IC3 [Bra11]
  - k-induction [SSS00]
  - Generation of auxiliary invariants [KGT11]
- Kind 2 solves all the 480 benchmarks it its test suite in less than 120s
- Considering k-induction in isolation, CVC-PORT is competitive

	Solved	Unique	Time (commonly solved)
CVC-PORT	**323**	**82**	109.6
Kind 2 (k-induction)	313	72	**9.6**

Injecting some welcome realism

- Kind 2 employs in cooperation:
  - IC3 [Bra11]
  - k-induction [SSS00]
  - Generation of auxiliary invariants [KGT11]
- Kind 2 solves all the 480 benchmarks it its test suite in less than 120s
- Considering k-induction in isolation, CVC-PORT is competitive

	Solved	Unique	Time (commonly solved)
CVC-PORT	**323**	**82**	109.6
Kind 2 (k-induction)	313	72	**9.6**

- We consider this encouraging given our framework is
  - not theory-specific
  - single-threaded
  - not optimized for reachability

Conclusions

Conclusions

- New enumerative function synthesis framework via divide and conquer
  - No dependence on point-wise specifications
  - Powered by SMT-driven classification algorithms
  - Implemented in CVC4SY
- Experimental evaluation shows significant gains w.r.t. previous SyGuS techniques for invariant synthesis

Future work

- Improving classification
  - Using constraint solving for synthesizing term assignments
  - Only considering relevant arguments when synthesizing predicates

$$f(0, 0, 0, 1, 2, 1, 0) \diamond f(1, 0, 0, 5, 2, 1, 3)$$

    - Can drastically reduce search space

- Improving sample
  - Reducing noise: make points as similar as possible

$$f(1, 0, 0, 1, 2, 1, 0) \diamond f(1, 0, 0, 5, 2, 1, 0)$$

  - Improve diversity via clustering analysis: only add new points to sample that are sufficiently different

Extending enumerative function synthesis via SMT-driven classification

Haniel Barbosa, Andrew Reynolds, Daniel Larraz, Cesare Tinelli

UFMG

The University of Iowa

Lógicos em Quarentena

2020-04-30, The Internet

Extra slides

Invariant Synthesis

```
Add(Int x, y) {
  z := x; i := 0;
  assume(y > 0);
  while (i < y) {
    z := z + 1;
    i := i + 1;
  }
  return z;
}
```

Post-condition:
$\forall x, y:\ z = x + y$

Result is the sum of the inputs

Invariant Synthesis

```
Add(Int x, y) {
  z := x; i := 0;
  assume(y > 0);
  while (i < y) {
    z := z + 1;
    i := i + 1;
  }
  return z;
}
```

Invariant?

Post-condition:
$\forall x, y : \; z = x + y$

Result is the sum of the inputs

Verification:

$$
\begin{array}{lcl}
z = x \wedge i = 0 \wedge y > 0 & \rightarrow & Inv(x,\, y,\, z,\, i) \\
Inv(x,\, y,\, z,\, i) \wedge i < y \wedge z' = z + 1 \wedge i' = i + 1 & \rightarrow & Inv(x,\, y,\, z',\, i') \\
Inv(x,\, y,\, z,\, i) \wedge i \geq y & \rightarrow & z = x + y
\end{array}
$$

Invariant Synthesis

```
Add(Int x, y) {
  z := x; i := 0;
  assume(y > 0);
  while (i < y) {
    z := z + 1;
    i := i + 1;
  }
  return z;
}
```

$$Inv(x,\, y,\, z,\, i)$$
$$z = x + i$$
$$z \le x + y$$

Post-condition:
$\forall x, y:\ z = x + y$

Result is the sum of the inputs

Verification:

$$\begin{array}{lcl}
z = x \wedge i = 0 \wedge y > 0 & \rightarrow & Inv(x,\, y,\, z,\, i) \\
Inv(x,\, y,\, z,\, i) \wedge i < y \wedge z' = z + 1 \wedge i' = i + 1 & \rightarrow & Inv(x,\, y,\, z',\, i') \\
Inv(x,\, y,\, z,\, i) \wedge i \ge y & \rightarrow & z = x + y
\end{array}$$

Invariant Synthesis in SyGuS

- State-of-the-art: LoopInvGen [PM17]: *data-driven* loop invariant inference with automatic feature synthesis
  - Precondition inference from sets of "good" and "bad" states
    - Feature synthesis for solving conflicts
  - PAC (*probably approximately correct*) algorithm for building candidate invariants
- "Bad" states are dependent on model of initial condition (no guaranteed convergence)
- No support for implication counterexamples

Invariant Synthesis with Unif+PI

- Refinement lemmas allows derivation of three kinds on data points:
  - "good points" (invariant must always hold)
  - "bad points" (invariant can never hold)
  - "implication points" (if invariant holds in first point it must hold in second)
- Native support for implication counterexamples
- Straightforward usage of classic information gain heuristic to build candidate solutions with decision tree learning
  - SMT solver "resolves" implication counterexample points as "good" and "bad"
  - Out-of-the-box Shannon entropy

References

References

Rajeev Alur, Rastislav Bodík, Garvit Juniwal, et al. "Syntax-guided synthesis". In: Formal Methods In Computer-Aided Design (FMCAD). IEEE, 2013, pp. 1–8.

Rajeev Alur, Arjun Radhakrishna, and Abhishek Udupa. "Scaling Enumerative Program Synthesis via Divide and Conquer". In: Tools and Algorithms for Construction and Analysis of Systems (TACAS). Ed. by Axel Legay and Tiziana Margaria. Vol. 10205. Lecture Notes in Computer Science. 2017, pp. 319–336.

Aaron R. Bradley. "SAT-Based Model Checking without Unrolling". In: Verification, Model Checking, and Abstract Interpretation (VMCAI). Ed. by Ranjit Jhala and David Schmidt. Berlin, Heidelberg: Springer Berlin Heidelberg, 2011, pp. 70–87.

Clark Barrett, Igor Shikanian, and Cesare Tinelli. "An Abstract Decision Procedure for a Theory of Inductive Data Types". In: JSAT 3.1-2 (2007), pp. 21–46.

Clark W. Barrett and Cesare Tinelli. "Satisfiability Modulo Theories". In: Handbook of Model Checking. Ed. by Edmund M. Clarke, Thomas A. Henzinger, Helmut Veith, et al. Springer, 2018, pp. 305–343.

References

Bruno Dutertre and Leonardo de Moura. "A Fast Linear-Arithmetic Solver for DPLL(T)". English. In: Computer Aided Verification (CAV). Ed. by Thomas Ball and Robert B. Jones. Vol. 4144. Lecture Notes in Computer Science. Springer Berlin Heidelberg, 2006, pp. 81–94.

Peter J. Downey, Ravi Sethi, and Robert Endre Tarjan. "Variations on the Common Subexpression Problem". In: J. ACM 27.4 (Oct. 1980), pp. 758–771.

Hassan Eldib, Meng Wu, and Chao Wang. "Synthesis of Fault-Attack Countermeasures for Cryptographic Circuits". In: Computer Aided Verification (CAV), Part II. Ed. by Swarat Chaudhuri and Azadeh Farzan. Vol. 9780. Lecture Notes in Computer Science. Springer, 2016, pp. 343–363.

Yu Feng, Ruben Martins, Jacob Van Geffen, et al. "Component-based synthesis of table consolidation and transformation tasks from examples". In: Conference on Programming Language Design and Implementation (PLDI). Ed. by Albert Cohen and Martin T. Vechev. ACM, 2017, pp. 422–436.

Pranav Garg, Christof Löding, P. Madhusudan, et al. "ICE: A Robust Framework for Learning Invariants". In: Computer Aided Verification (CAV). Ed. by Armin Biere and Roderick Bloem. Vol. 8559. Lecture Notes in Computer Science. Springer, 2014, pp. 69–87.

References

Sumit Gulwani. "Automating string processing in spreadsheets using input-output examples". In: Symposium on Principles of Programming Languages (POPL). Ed. by Thomas Ball and Mooly Sagiv. ACM, 2011, pp. 317–330.

Temesghen Kahsai, Yeting Ge, and Cesare Tinelli. "Instantiation-Based Invariant Discovery". In: NASA Formal Methods. Ed. by Mihaela Gheorghiu Bobaru, Klaus Havelund, Gerard J. Holzmann, et al. Vol. 6617. Lecture Notes in Computer Science. Springer, 2011, pp. 192–206.

Xuan-Bach D. Le, Duc-Hiep Chu, David Lo, et al. "S3: syntax- and semantic-guided repair synthesis via programming by examples". In: Joint Meeting on Foundations of Software Engineering (ESEC/FSE). Ed. by Eric Bodden, Wilhelm Schäfer, Arie van Deursen, et al. ACM, 2017, pp. 593–604.

Greg Nelson and Derek C. Oppen. "Fast Decision Procedures Based on Congruence Closure". In: J. ACM 27.2 (1980), pp. 356–364.

References

Andres Nötzli, Andrew Reynolds, Haniel Barbosa, et al. "Syntax-Guided Rewrite Rule Enumeration for SMT Solvers". In: Theory and Applications of Satisfiability Testing (SAT). Ed. by Mikolás Janota and Inês Lynce. Vol. 11628. Lecture Notes in Computer Science. Springer, 2019, pp. 279–297.

Daniel Neider, Shambwaditya Saha, and P. Madhusudan. "Compositional Synthesis of Piece-Wise Functions by Learning Classifiers". In: ACM Trans. Comput. Log. 19.2 (2018), 10:1–10:23.

ThanhVu Nguyen, Westley Weimer, Deepak Kapur, et al. "Connecting Program Synthesis and Reachability: Automatic Program Repair Using Test-Input Generation". In: Tools and Algorithms for Construction and Analysis of Systems (TACAS). Ed. by Axel Legay and Tiziana Margaria. Berlin, Heidelberg: Springer Berlin Heidelberg, 2017, pp. 301–318.

Saswat Padhi and Todd D. Millstein. "Data-Driven Loop Invariant Inference with Automatic Feature Synthesis". In: CoRR abs/1707.02029 (2017). arXiv: 1707.02029.

References

Saswat Padhi, Rahul Sharma, and Todd D. Millstein. “Data-driven precondition inference with learned features”. In: Conference on Programming Language Design and Implementation (PLDI). Ed. by Chandra Krintz and Emery Berger. ACM, 2016, pp. 42–56.

Andrew Reynolds, Haniel Barbosa, Andres Nötzli, et al. “cvc4sy: Smart and Fast Term Enumeration for Syntax-Guided Synthesis”. In: Computer Aided Verification (CAV), Part II. Ed. by Isil Dillig and Serdar Tasiran. Vol. 11562. Lecture Notes in Computer Science. Cham: Springer International Publishing, 2019, pp. 74–83.

Andrew Reynolds, Viktor Kuncak, Cesare Tinelli, et al. “Refutation-based synthesis in SMT”. In: Formal Methods in System Design (2017).

Andrew Reynolds, Arjun Viswanathan, Haniel Barbosa, et al. “Datatypes with Shared Selectors”. In: International Joint Conference on Automated Reasoning (IJCAR). Ed. by Didier Galmiche, Stephan Schulz, and Roberto Sebastiani. Vol. 10900. Lecture Notes in Computer Science. Springer, 2018, pp. 591–608.

Eric Schkufza, Rahul Sharma, and Alex Aiken. “Stochastic superoptimization”. In: Architectural Support for Programming Languages and Operating Systems (ASPLOS
Ed. by Vivek Sarkar and Rastislav Bodík. ACM, 2013, pp. 305–316.

References

Mary Sheeran, Satnam Singh, and Gunnar Stålmarck. “Checking Safety Properties Using Induction and a SAT-Solver”. In: Formal Methods In Computer-Aided Design (FMCAD). Ed. by Warren A. Hunt Jr. and Steven D. Johnson. Vol. 1954. Lecture Notes in Computer Science. Springer, 2000, pp. 108–125.

Armando Solar-Lezama, Liviu Tancau, Rastislav Bodík, et al. “Combinatorial sketching for finite programs”. In: Architectural Support for Programming Languages and Operating Systems (ASPLOS Ed. by John Paul Shen and Margaret Martonosi. ACM, 2006, pp. 404–415.

Abhishek Udupa, Arun Raghavan, Jyotirmoy V. Deshmukh, et al. “TRANSIT: specifying protocols with concolic snippets”. In: Conference on Programming Language Design and Implementation (PLDI). Ed. by Hans-Juergen Boehm and Cormac Flanagan. ACM, 2013, pp. 287–296.