

A photograph of a building with a green wall and a stone path leading to a doorway. The building is covered in dense green foliage, and a stone path leads to a doorway. The sky is blue, and there are trees in the background.

# Program analysis: from proving correctness to proving incorrectness

**Roberto Bruni, Roberta Gori  
(University of Pisa)  
Lecture #11**

**BISS 2024  
March 11-15, 2024**



# Separation SIL

# Ongoing work

## Sufficient Incorrectness Logic: SIL and Separation SIL

FLAVIO ASCARI, Università di Pisa, Italy  
ROBERTO BRUNI, Università di Pisa, Italy  
ROBERTA GORI, Università di Pisa, Italy  
FRANCESCO LOGOZZO, Meta Platforms, Inc., USA

Sound over-approximation methods have been proved effective for guaranteeing the absence of errors, but inevitably they produce false alarms that can hamper the programmers. Conversely, under-approximation methods are aimed at bug finding and are free from false alarms. We introduce Sufficient Incorrectness Logic (SIL), a new under-approximating, triple-based program logic to reason about program errors. SIL is designed to set apart the initial states leading to errors. We prove that SIL is correct and complete for a minimal set of rules, and we study additional rules that can facilitate program analyses. We formally compare SIL to existing triple-based program logics. Incorrectness Logic and SIL both perform under-approximations, but while the former exposes only true errors, the latter locates the set of initial states that lead to such errors. Hoare Logic performs over-approximations and as such cannot capture the set of initial states leading to errors in *non-deterministic* programs – for deterministic and terminating programs, Hoare Logic and SIL coincide. Finally, we instantiate SIL with Separation Logic formulae (Separation SIL) to handle pointers and dynamic allocation and we prove its correctness and, for loop-free programs, also its completeness. We argue that in some cases Separation SIL can yield more succinct postconditions and provide stronger guarantees than Incorrectness Separation Logic and can support effective backward reasoning.

CCS Concepts: • **Theory of computation** → **Logic and verification**; *Proof theory*; *Hoare logic*; **Separation logic**; *Programming logic*.

Additional Key Words and Phrases: Sufficient Incorrectness Logic, Incorrectness Logic, Necessary Conditions, Outcome Logic

### 1 INTRODUCTION

Formal methods aim to provide tools for reasoning and establishing program guarantees. Historically, research in formal reasoning progressed from manual correctness proofs to effective, automatic methods that improve program reliability and security. In the late 60s, Floyd [1967] and Hoare [1969] independently introduced formal systems to reason about programs. In the 70s/early 80s, the focus was on mechanization, with the introduction of numerous techniques such as predicate transformers [Dijkstra 1975], Abstract Interpretation [Cousot and Cousot 1977], model checking [Clarke and Emerson 1981], type inference [Damas and Milner 1982] and mechanized program proofs [Coquand and Huet 1985]. Those seminal works, in conjunction with the development of automatic and semi-automatic theorem provers (e.g., [de Moura 2007]) brought impressive wins in proving program correctness of real-world applications. For instance, the Astree abstract interpreter automatically proves the absence of runtime errors in millions of lines of safety-critical C [Blanchet et al. 2003], the SLAM model checker was used to check Windows drivers [Ball and Rajamani 2001], CompCert is a certified C compiler developed in Coq [Leroy 2009], and VCC uses the calculus of weakest precondition to verify safety properties of annotated Concurrent C programs [Cohen et al. 2009].

Despite the aforementioned successes, effective program correctness methods struggle to reach mainstream adoption. As program correctness is undecidable, all those methods *over-approximate* programs behaviours. Over-approximation guarantees soundness (if the program is proved to be

Authors' addresses: Flavio Ascari, Dipartimento di Informatica, Università di Pisa, Largo B. Pontecorvo 3, 56127, Pisa, Italy, flavio.ascari@phd.unipi.it; Roberto Bruni, Dipartimento di Informatica, Università di Pisa, Largo B. Pontecorvo 3, 56127, Pisa, Italy, roberto.bruni@unipi.it; Roberta Gori, Dipartimento di Informatica, Università di Pisa, Largo B. Pontecorvo 3, 56127, Pisa, Italy, roberta.gori@unipi.it; Francesco Logozzo, Meta Platforms, Inc., USA, logozzo@meta.com.

“Separation SIL can yield more succinct postconditions and provide stronger guarantees than ISL and can support effective backward reasoning”



# SepSIL = SIL + SL

SIL

$$\langle\langle P \rangle\rangle \mathit{r} \langle\langle Q \rangle\rangle$$

SL

$$\frac{\{P\} \mathit{r} \{Q\}}{\{P * R\} \mathit{r} \{Q * R\}}$$

SepSIL

$$\frac{\langle\langle P \rangle\rangle \mathit{r} \langle\langle Q \rangle\rangle}{\langle\langle P * R \rangle\rangle \mathit{r} \langle\langle Q * R \rangle\rangle}$$



# Regular commands

regular  
command

$r ::=$

$e$

|

$r_1; r_2$

|

$r_1 + r_2$

|

$r^*$

atomic  
command

choice

Kleene  
star

$e ::=$  skip

|  $b?$

|  $x := a$

|  $x := [y]$  // read

|  $[x] := y$  // write

|  $x := \text{alloc}()$

|  $\text{free}(x)$

simplified



# Assertion language

assertion

$P ::=$  true | false |  $a_1 < a_2$  |  $a_1 = a_2$  | ...  
|  $\neg P$  |  $P_1 \wedge P_2$  |  $\exists x. P$  | ...  
| emp  
|  $a_1 \mapsto a_2$   
|  $P_1 * P_2$   
|  $x \mapsto$

Boolean and  
classical  
assertions

structural  
assertions

track deallocated  
locations



# Local axioms: write

SL

$$\frac{}{\{x \mapsto \_ \} [x] := y \{x \mapsto y \}}$$

ISL

$$\frac{}{[x \mapsto v] [x] := y [ok : x \mapsto y]}$$

SepSIL

$$\frac{}{\llbracket x \mapsto \_ \rrbracket [x] := y \llbracket x \mapsto y \rrbracket}$$

weakest pre



# Local axioms: read

SL

$$\frac{}{\{y \mapsto v\} x := [y] \{x = v \wedge y \mapsto v\}}$$

ISL

$$\frac{}{[y \mapsto v] x := [y] [\text{ok} : x = v \wedge y \mapsto v]}$$

SepSIL

$$\frac{}{\langle\langle y \mapsto v \wedge (v = x') \rangle\rangle x := [y] \langle\langle y \mapsto v \wedge (x = x') \rangle\rangle}$$

Hoare style

applicable to any  
post



# Local axioms: allocation

SL

$$\frac{}{\{\text{emp}\} x := \text{alloc}() \{x \mapsto \_ \}}$$

ISL

$$\frac{}{[\text{emp}] x := \text{alloc}() [\text{ok} : x \mapsto \_]}$$

SepSL

$$\frac{}{\langle\langle \text{emp} \rangle\rangle x := \text{alloc}() \langle\langle x \mapsto \_ \rangle\rangle}$$



# Local axioms: dispose

SL

$$\{x \mapsto \_ \} \text{free}(x) \{ \text{emp} \}$$

ISL

$$\frac{}{[x \mapsto v] \text{free}(x) [\text{ok} : x \not\mapsto ]}$$

SepSL

$$\frac{}{\langle\langle x \mapsto \_ \rangle\rangle \text{free}(x) \langle\langle x \not\mapsto \rangle\rangle}$$

using cons can be strengthened to  $x \mapsto v$



# Different proofs of a real bug

# Use-after-lifetime bug

```
void deref_after_pb(std::vector<int> *v) {  
    int *x = &v->at(1);  
    v->push_back(42);  
    std::cout << *x << "\n"; }  
    from std::vector library, can deallocate and then reallocate v
```

```
push_back.cpp:7: error: VECTOR_INVALIDATION. accessing memory that was  
potentially invalidated by 'std::vector::push_back()' on line 6.
```

```
5.     int *x = &(v->at(1));  
6.     v->push_back(42);  
7. >   std::cout << *x << "\n"; }  
    if v is reallocated, x is invalidated
```

The C++ use-after-lifetime bug (above); the Pulse error message (below).

abstracted from real  
occurrences at Facebook



# From C++ to regular commands

$$[v \mapsto a * a \mapsto -] \text{client}(v) \left[ \text{er}(\text{L}_{rx}) : \exists a'. v \mapsto a' * a' \mapsto - * a \not\mapsto \right]$$

```
void push_back(int **v)
{
  if (nondet()) {
    free(*v);
    *v = malloc(sizeof(int));
  }
}
```

```
void client(v) {
  int* x = *v;
  push_back(v);
  *x = 88; }
```

C version

```
push_back(v)  $\triangleq$ 
  local z, y in
    z := *;
    (assume(z  $\neq$  0); Lrv: y := [v];
     Lf: free(y);
     y := malloc(); [v] := y)
  + (assume(z = 0); skip)
```

```
client(v)  $\triangleq$ 
  local x in
    x := [v];
    push_back(v);
    Lrx: [x] := 88
```

ISL version

```
// client, inlining proc call
x := [v];
```

```
( // push_back
  y := [v];
  free(y);
  y := alloc();
  [v] := y
+
  skip
)
```

```
[x] := 88
```

SepISL version

stronger guarantee:  
any state in pre can  
lead to error

more succinct post

$$\langle\langle v \mapsto a * a \mapsto \_ * \text{true} \rangle\rangle \text{rclient} \langle\langle x \not\mapsto \_ * \text{true} \rangle\rangle$$

# ISL derivation

$[v \mapsto a * a \mapsto -]$

**local**  $y, z$  **in**

$z := *; // \text{HAVOC}$

$[ok: z=1 * v \mapsto a * a \mapsto -]$

$( \text{assume}(z \neq 0); // \text{ASSUME}$

$[ok: z=1 * z \neq 0 * v \mapsto a * a \mapsto -]$

$L_{rv}: y := [v]; // \text{LOAD}$

$[ok: z=1 * y=a * v \mapsto a * a \mapsto -]$

$L_f: \text{free}(y); // \text{FREE}$

$[ok: z=1 * y=a * v \mapsto a * a \not\mapsto ]$

$y := \text{malloc}(); // \text{ALLOC1, CHOICE}$

$[ok: z=1 * v \mapsto a * a \not\mapsto * y \mapsto -]$

$[v] := y; // \text{STORE}$

$[ok: z=1 * v \mapsto y * a \not\mapsto * y \mapsto -]$

$) + (\dots) // \text{CHOICE}$

$[ok: z=1 * v \mapsto y * a \not\mapsto * y \mapsto -]$

$// \text{LOCAL}$

$[ok: \exists a'. v \mapsto a' * a' \mapsto - * a \not\mapsto ]$

$[v \mapsto a * a \mapsto -]$

**local**  $x$  **in**

$x := [v]; // \text{LOAD}$

$[ok: x=a * v \mapsto a * a \mapsto -]$

$\text{push\_back}(v); // \text{PB-OK}$

$[ok: \exists a'. x=a * v \mapsto a' * a' \mapsto - * a \not\mapsto ] // \text{CONS}$

$[ok: \exists a'. x=a * v \mapsto a' * a' \mapsto - * x \not\mapsto ]$

$L_{rx}: [x] := 88; // \text{STOREER}$

$[er(L_{rx}): \exists a'. x=a * v \mapsto a' * a' \mapsto - * x \not\mapsto ]$

$// \text{LOCAL}$

$[er(L_{rx}): \exists a'. v \mapsto a' * a' \mapsto - * a \not\mapsto ]$



# SepSIL derivation

$\langle\langle v \mapsto a * a \mapsto \_ * \text{true} \rangle\rangle \Rightarrow \langle\langle v \mapsto a * a \mapsto \_ * (a = a \vee a \not\mapsto) * \text{true} \rangle\rangle$

$x := [v]; // \text{Load} + \text{Frame}$

$\langle\langle v \mapsto a * a \mapsto \_ * (x = a \vee x \not\mapsto) * \text{true} \rangle\rangle \Rightarrow \langle\langle (v \mapsto a * a \mapsto \_ * (x = a \vee x \not\mapsto) * \text{true}) \vee (x \not\mapsto * \text{true}) \rangle\rangle$

( // push\_back: Choice

$\langle\langle v \mapsto a * a \mapsto \_ * (x = a \vee x \not\mapsto) * \text{true} \rangle\rangle$

$y := [v]; // \text{Load} + \text{Frame}$

$\langle\langle v \mapsto a * y \mapsto \_ * (x = y \vee x \not\mapsto) * \text{true} \rangle\rangle \Rightarrow \langle\langle v \mapsto \_ * y \mapsto \_ * (x = y \vee x \not\mapsto) * \text{true} \rangle\rangle$

$\text{free}(y); // \text{Free} + \text{Frame}$

$\langle\langle v \mapsto \_ * y \not\mapsto * (x = y \vee x \not\mapsto) * \text{true} \rangle\rangle \Rightarrow \langle\langle x \not\mapsto * v \mapsto \_ * \text{emp} * \text{true} \rangle\rangle$

$y := \text{alloc}(); // \text{Alloc} + \text{Frame}$

$\langle\langle x \not\mapsto * v \mapsto \_ * y \mapsto \_ * \text{true} \rangle\rangle \Rightarrow \langle\langle x \not\mapsto * v \mapsto \_ * \text{true} \rangle\rangle$

$[v] := y // \text{Write} + \text{Frame}$

$\langle\langle x \not\mapsto * v \mapsto y * \text{true} \rangle\rangle \Rightarrow \langle\langle x \not\mapsto * \text{true} \rangle\rangle$

+

$\langle\langle x \not\mapsto * \text{true} \rangle\rangle \text{ skip } \langle\langle x \not\mapsto * \text{true} \rangle\rangle // \text{Skip} + \text{Frame}$

)

$\langle\langle x \not\mapsto * \text{true} \rangle\rangle$

$[x] := 88$

# Correctness and completeness



# Relational semantics

$$\llbracket \text{skip} \rrbracket \triangleq \{(\sigma, \sigma)\}$$

$$\llbracket b? \rrbracket \triangleq \{(\sigma, \sigma) \mid \sigma = \langle s, h \rangle \wedge s \vDash b\}$$

$$\llbracket x := a \rrbracket \triangleq \{(\langle s, h \rangle, \langle s[x \mapsto \llbracket a \rrbracket s], h \rangle)\}$$

$$\llbracket x := [y] \rrbracket \triangleq \{(\langle s, h \rangle, \langle s[x \mapsto v], h \rangle) \mid v = h(s(y)) \in \mathbb{Z}\}$$

$$\llbracket [x] := y \rrbracket \triangleq \{(\langle s, h \rangle, \langle s, h[s(x) \mapsto s(y)] \rangle) \mid h(s(x)) \in \mathbb{Z}\}$$

$$\llbracket x := \text{alloc}(\ ) \rrbracket \triangleq \{(\langle s, h \rangle, \langle s[x \mapsto n], h[n \mapsto v] \rangle) \mid v \in \mathbb{Z} \wedge (n \notin \text{dom}(h) \vee h(n) = \perp)\}$$

$$\llbracket \text{free}(x) \rrbracket \triangleq \{(\langle s, h \rangle, \langle s, h[s(x) \mapsto \perp] \rangle) \mid h(s(x)) \in \mathbb{Z}\}$$

# Actual rules of SepSIL

$$\frac{}{\langle\langle \mathbf{emp} \rangle\rangle \text{ skip } \langle\langle \mathbf{emp} \rangle\rangle} \langle\langle \text{skip} \rangle\rangle \quad \frac{}{\langle\langle q[a/x] \rangle\rangle x := a \langle\langle q \rangle\rangle} \langle\langle \text{assign} \rangle\rangle$$

$$\frac{}{\langle\langle q \wedge b \rangle\rangle b? \langle\langle q \rangle\rangle} \langle\langle \text{assume} \rangle\rangle \quad \frac{}{\langle\langle x \mapsto - \rangle\rangle \text{ free}(x) \langle\langle x \not\mapsto - \rangle\rangle} \langle\langle \text{free} \rangle\rangle$$

$$\frac{}{\langle\langle \mathbf{emp} \rangle\rangle x := \text{alloc}() \langle\langle x \mapsto v \rangle\rangle} \langle\langle \text{alloc} \rangle\rangle$$

$$\frac{x \notin \text{fv}(a)}{\langle\langle y \mapsto a * q[a/x] \rangle\rangle x := [y] \langle\langle y \mapsto a * q \rangle\rangle} \langle\langle \text{load} \rangle\rangle$$

$$\frac{}{\langle\langle x \mapsto - \rangle\rangle [x] := y \langle\langle x \mapsto y \rangle\rangle} \langle\langle \text{store} \rangle\rangle$$

---

$$\frac{\langle\langle p \rangle\rangle r \langle\langle q \rangle\rangle \quad x \notin \text{fv}(r)}{\langle\langle \exists x.p \rangle\rangle r \langle\langle \exists x.q \rangle\rangle} \langle\langle \text{exists} \rangle\rangle$$

$$\frac{\langle\langle p \rangle\rangle r \langle\langle q \rangle\rangle \quad \text{fv}(t) \cap \text{mod}(r) = \emptyset}{\langle\langle p * t \rangle\rangle r \langle\langle q * t \rangle\rangle} \langle\langle \text{frame} \rangle\rangle$$



# Correctness

**Th.** [*correctness*]

If  $\langle\langle P \rangle\rangle r \langle\langle Q \rangle\rangle$  then  $P \subseteq \llbracket \overleftarrow{r} \rrbracket Q$

**Proof.** By induction on the derivation.

# (Relative) completeness

**Th.** *[completeness]*

Any valid triple  $\langle\langle P \rangle\rangle r \langle\langle Q \rangle\rangle$  can be derived

**Proof.** See ArXiV draft.

## 6.7 Relative completeness of Separation SIL

The proof system in Section 6.3 is not complete. To move towards completeness, we first limit the assertion language to the existential fragment of first-order logic:

$$\text{Asl } \exists p, q, t ::= \text{false} \mid \text{true} \mid p \wedge q \mid p \vee q \mid \exists x. p \mid a \times a \\ \mid \text{emp} \mid x \mapsto a \mid x \not\mapsto \mid p * q$$

We remove negation, so we don't include universal quantifiers and heap assertions must be positive. However, we argue that this is sufficient to find bugs: for instance, in the example in Section 6.5, we only used assertions from this subset.

With this limited assertion language, the proof system in Section 6.3 is complete for all atomic commands except `alloc`. To deal with `alloc`, we need the ability to refer to the specific memory location that was allocated. However, the naive solution to add a constraint  $x = \alpha$  in the post of `alloc` makes the frame rule unsound: for instance, the following triple is not valid:

$$\langle\text{emp} * \alpha \mapsto \neg\rangle x := \text{alloc}() \langle(x \mapsto \neg \wedge x = \alpha) * \alpha \mapsto \neg\rangle.$$

To recover the frame rule, just like ISL needs the deallocated assertion in the post [Raad et al. 2020, §3], we need a "will be allocated" assertion in the pre. To this end we use the  $\not\mapsto$  assertion, and change the semantic model to only allocate a memory location that is explicitly  $\perp$  instead of one not in the domain of the heap. We formalize this by letting  $\text{avail}(l) \triangleq h(l) = \perp$  in Figure 13a, and

replacing the axiom `alloc` with

$$\langle\beta \not\mapsto\rangle x := \text{alloc}() \langle x = \beta \wedge x \mapsto \perp \rangle \langle\text{alloc}\rangle$$

Soundness still holds for this different semantics. Moreover, we can prove relative completeness [Apt and Olderog 2019, §4.3] for loop-free programs:

**THEOREM 6.3 (RELATIVE COMPLETENESS FOR LOOP-FREE PROGRAMS).** *Suppose to have an oracle to prove implications between formulas in Asl. Let  $r \in \text{HRCmd}$  be a regular command without  $*$  and  $p, q \in \text{Asl}$  such that  $\llbracket \tilde{r} \rrbracket \{q\} \supseteq \{p\}$ . Then the triple  $\langle p \rangle r \langle q \rangle$  is provable.*


The proof relies on the possibility to rewrite any  $q$  in an equivalent assertion of the form  $\exists x_1. \dots \exists x_n. \bigvee_{1 \leq i \leq k} q_i$  where all  $q_i$  are assertions involving atoms composed with  $\wedge$  and  $*$  only. This way, completeness is proved for such  $q_i$  first and then extended to the entire  $q$  thanks to rules `disj` and `exists`. Notably, we show that the weakest (possible) precondition  $\llbracket \tilde{r} \rrbracket \{q\}$  of loop-free programs is always expressible as an assertion  $t \in \text{Asl}$ , namely  $\{t\} = \llbracket \tilde{r} \rrbracket \{q\}$ , and prove that the triple  $\langle t \rangle r \langle q \rangle$  can be derived. Then, by `cons`, the theorem follows for any  $p$  that implies  $t$ .





# Questions


# Question 1

Which SepSIIL triples are valid ?

$\langle\langle \text{emp} \rangle\rangle \text{ free}(x) \langle\langle x \not\mapsto \rangle\rangle$  

$\langle\langle x \not\mapsto \rangle\rangle \text{ free}(x) \langle\langle x \not\mapsto \rangle\rangle$  

$\langle\langle \text{false} \rangle\rangle \text{ free}(x) \langle\langle \text{emp} \rangle\rangle$  

$\langle\langle x \mapsto \_ \rangle\rangle \text{ free}(x) \langle\langle \text{emp} \rangle\rangle$  

# Question 2

Transform the following C-like code in the syntax of SepSI

$i := 0 ; q := * p ; \text{while } (q \neq \text{nil}) \text{ do } \{ q := * q ; i := i + 1 \}$

$i := 0 ;$

$q := [p] ;$

$( (q \neq \text{nil}?) ; q := [q] ; i := i + 1 )^*$ ;

$(q = \text{nil}?)$



# \* Exam 16

Prove the SepSIL triple  $\langle\langle p \mapsto \text{nil} * \text{true} \rangle\rangle c \langle\langle i = 0 \rangle\rangle$  where

$c \triangleq i := 0 ; q := * p ; \text{while } (q \neq \text{nil}) \text{ do } \{ q := * q ; i := i + 1 \}$

# Exam registration form

Registration to the exam



\* Obbligatoria

1. First name  
\*

Inserisci la risposta

2. Last name  
\*

Inserisci la risposta

3. University \*

Inserisci la risposta

4. email \*

Inserisci la risposta

5. Are you willing to give the exam?  
\*  Yes  
 No  
 Maybe

6. Which deadline should we set for receiving your solutions?  
\*

Immetti la data (dd/MM/yyyy)



# Some references



Roberto Bruni, Roberto Giacobazzi, Roberta Gori, Isabel Garcia-Contreras , Dusko Pavlovic:  
**Abstract extensionality: on the properties of incomplete abstract interpretations.**  
Proc. ACM Program. Lang. 4(POPL): 28:1-28:28 (2020)



Roberto Bruni, Roberto Giacobazzi, Roberta Gori, Francesco Ranzato:  
**A Logic for Locally Complete Abstract Interpretations.** LICS 2021: 1-13

Roberto Bruni, Roberto Giacobazzi, Roberta Gori, Francesco Ranzato:  
**A Correctness and Incorrectness Program Logic.** J. ACM 70(2): 15:1-15:45 (2023)



Flavio Ascari, Roberto Bruni, Roberta Gori:  
**Limits and difficulties in the design of under-approximation abstract domains.**  
FoSSaCS 2022: 21-39



Roberto Bruni, Roberto Giacobazzi, Roberta Gori, Francesco Ranzato:  
**Abstract interpretation repair.** PLDI 2022: 426-441



Roberto Bruni, Roberta Gori, Nicolas Manini:  
**Deciding Program Properties via Complete Abstractions on Bounded Domains.**  
SAS 2022: 175-200



Flavio Ascari, Roberto Bruni, Roberta Gori:  
**Logics for Extensional, Locally Complete Analysis via Domain Refinements.**  
ESOP 2023: 1-27



Flavio Ascari, Roberto Bruni, Roberta Gori, Francesco Logozzo:  
**Sufficient Incorrectness Logic: SIL and Separation SIL.**  
CoRR abs/2310.18156 (2023)



**{many}** (HL ; IL ; AI ; LCL ; NC ; SIL  
+  
SL ; ISL ; SepSIL)\* **[thanks]**