Verified compilation
An introduction to CompCert

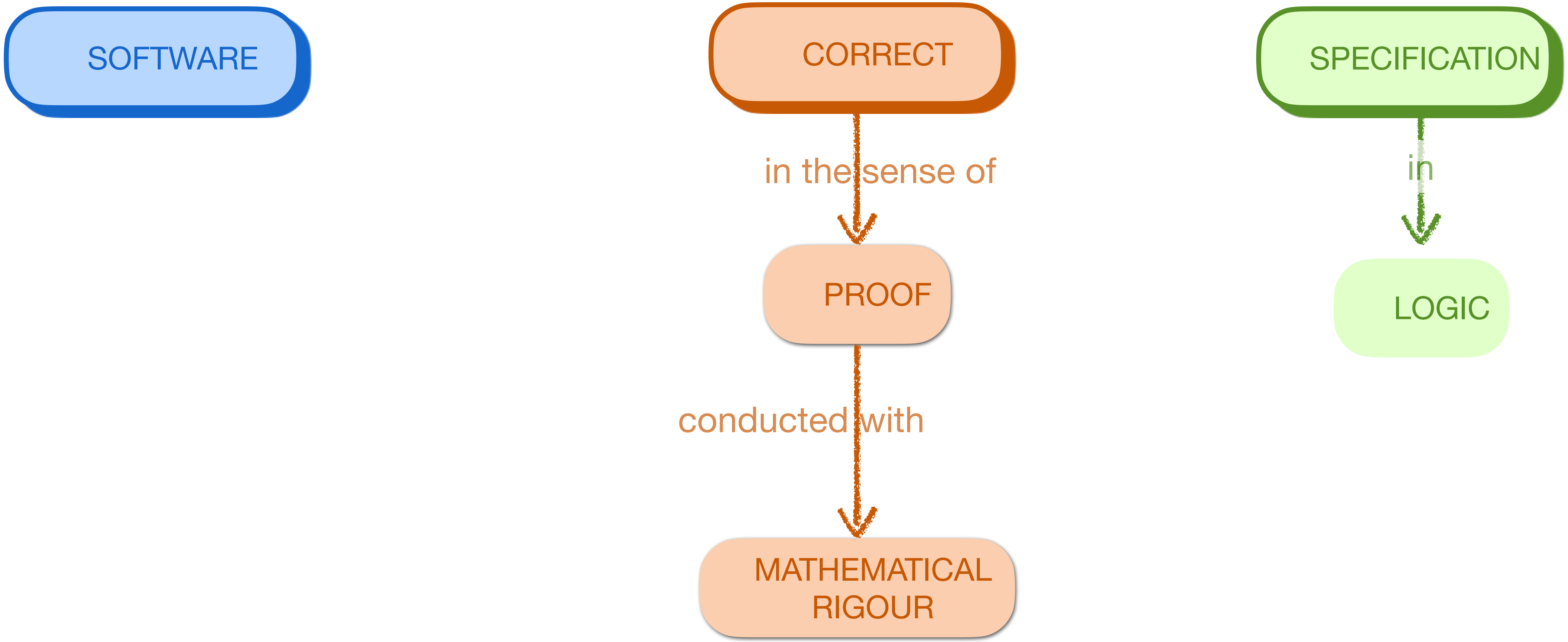
Sandrine Blazy

Université de Rennes

UMR IRISA

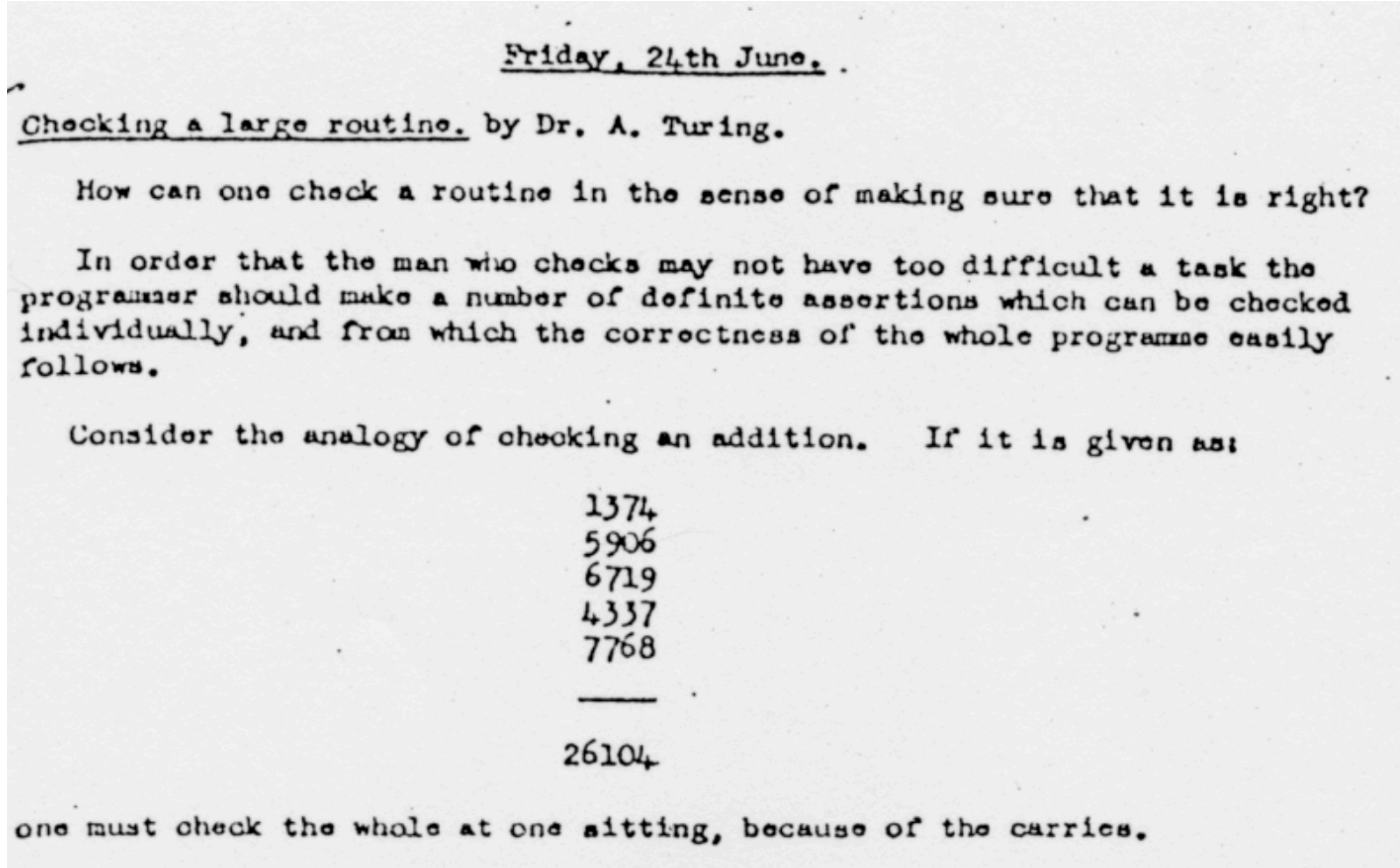
VTSA, Nancy, 2023-08-28

Deductive verification



From early intuitions …

A. M. Turing.
Checking a large routine. 1949.



Friday, 24th June.

Checking a large routine. by Dr. A. Turing.

How can one check a routine in the sense of making sure that it is right?

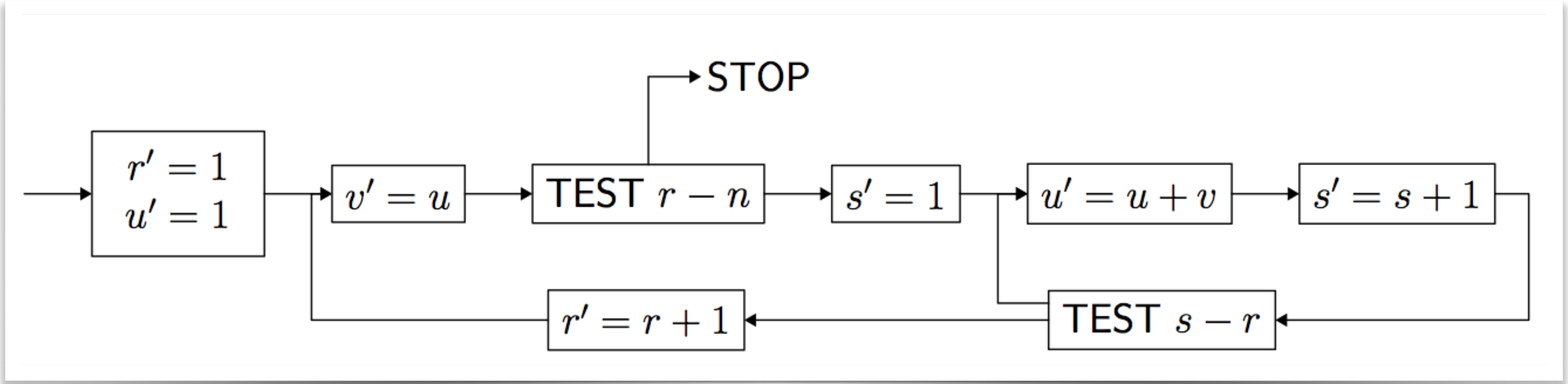
In order that the man who checks may not have too difficult a task the programmer should make a number of definite assertions which can be checked individually, and from which the correctness of the whole programme easily follows.

Consider the analogy of checking an addition. If it is given as:

1374
5906
6719
4337
7768

26104

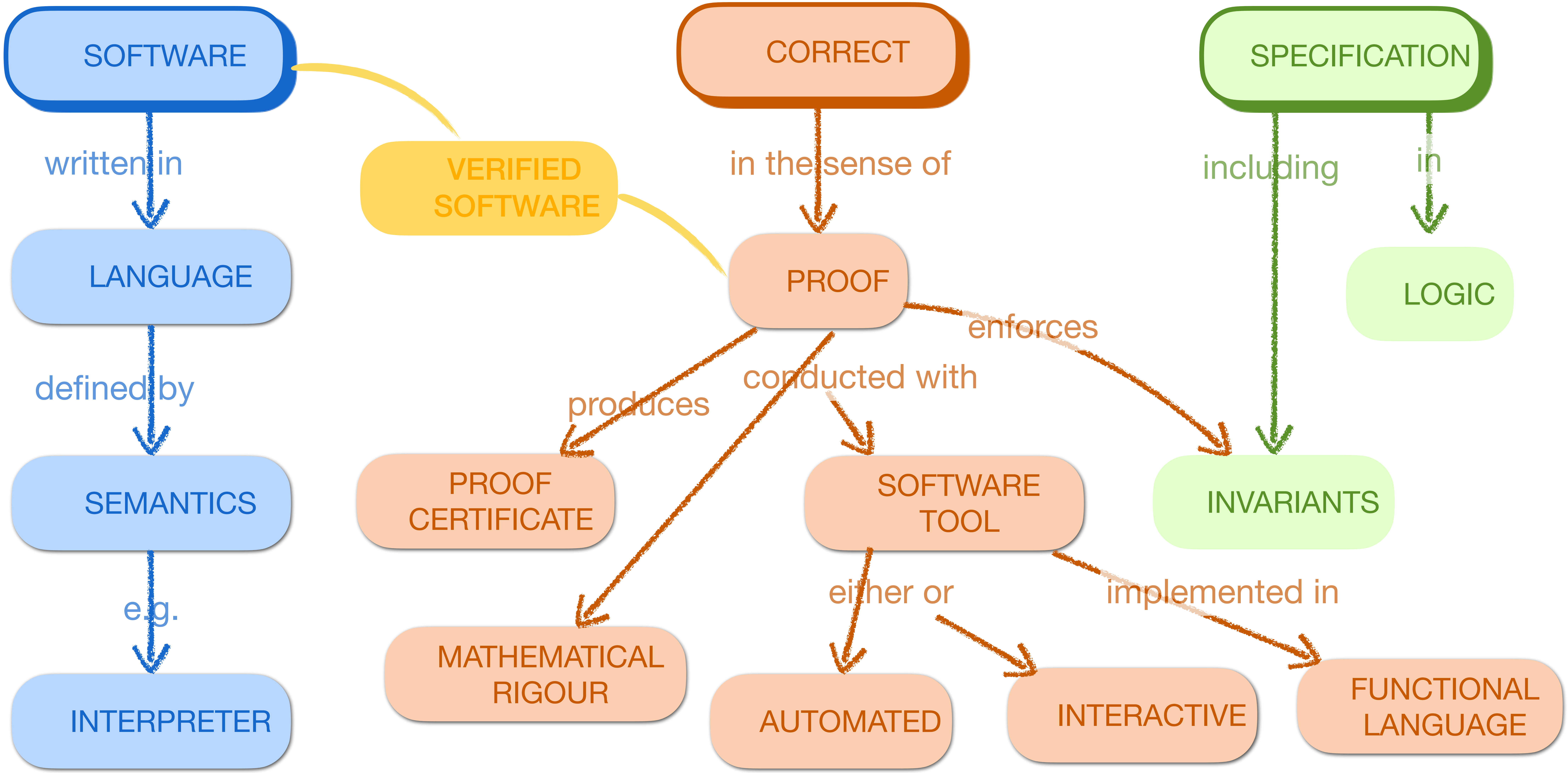
one must check the whole at one sitting, because of the carries.



$u \leftarrow 1$
for $r = 0$ **to** $n - 1$ **do**
  $v \leftarrow u$
  for $s = 1$ **to** r **do**
    $u \leftarrow u + v$

… to deductive-verification and automated tools

Floyd 1967, Hoare 1969



Another historical example

Boyer-Moore's majority. 1980

Given N votes, determine the majority if any

A	A	A	C	C	B	B	C	C	C	B	C	C

majority = A

cpt_delta = 3

MJRTY—A Fast Majority Vote Algorithm[1]

Robert S. Boyer and J Strother Moore

Computer Sciences Department
University of Texas at Austin
and
Computational Logic, Inc.
1717 West Sixth Street, Suite 290
Austin, Texas

Abstract

A new algorithm is presented for determining which, if any, of an arbitrary number of candidates has received a majority of the votes cast in an election. The number of comparisons required is at most twice the number of votes. Furthermore, the algorithm uses storage in a way that permits an efficient use of magnetic tape. A Fortran version of the algorithm is exhibited. The Fortran code has been proved correct by a mechanical verification system for Fortran. The system and the proof are discussed.

Another historical example

Boyer-Moore's majority. 1980

Given N votes, determine the majority if any

A	A	A	C	C	B	B	C	C	C	B	C	C

majority = A

cpt_delta = 3

A	~~A~~	~~A~~	~~C~~	~~C~~	B	B	C	C	C	B	C	C

majority = A

cpt_delta = 1

MJRTY—A Fast Majority Vote Algorithm[1]

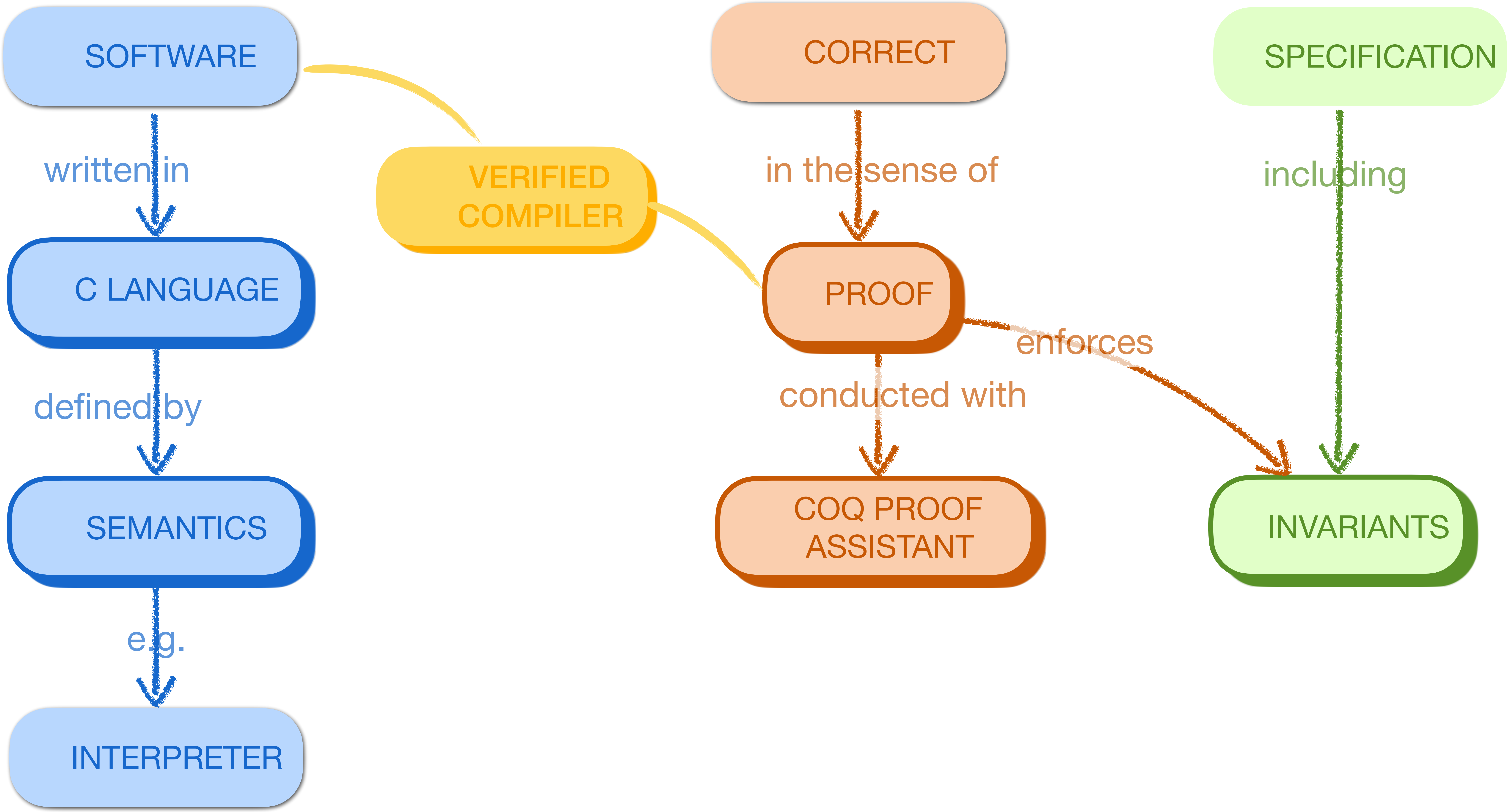
Robert S. Boyer and J Strother Moore

Computer Sciences Department
University of Texas at Austin
and
Computational Logic, Inc.
1717 West Sixth Street, Suite 290
Austin, Texas

Abstract

A new algorithm is presented for determining which, if any, of an arbitrary number of candidates has received a majority of the votes cast in an election. The number of comparisons required is at most twice the number of votes. Furthermore, the algorithm uses storage in a way that permits an efficient use of magnetic tape. A Fortran version of the algorithm is exhibited. The Fortran code has been proved correct by a mechanical verification system for Fortran. The system and the proof are discussed.

Part 1: summary



Lecture material

https://people.irisa.fr/Sandrine.Blazy/2023-VTSA

These slides
(with many slides borrowed from Xavier Leroy)

COLLÈGE
DE FRANCE
1530

Mechanized semantics, second lecture

***Traduttore, traditore*:**
formal verification of a compiler

Xavier Leroy

2019-12-12

Collège de France, chair of software sciences

Companion Coq development

Mechanized semantics: the Coq development

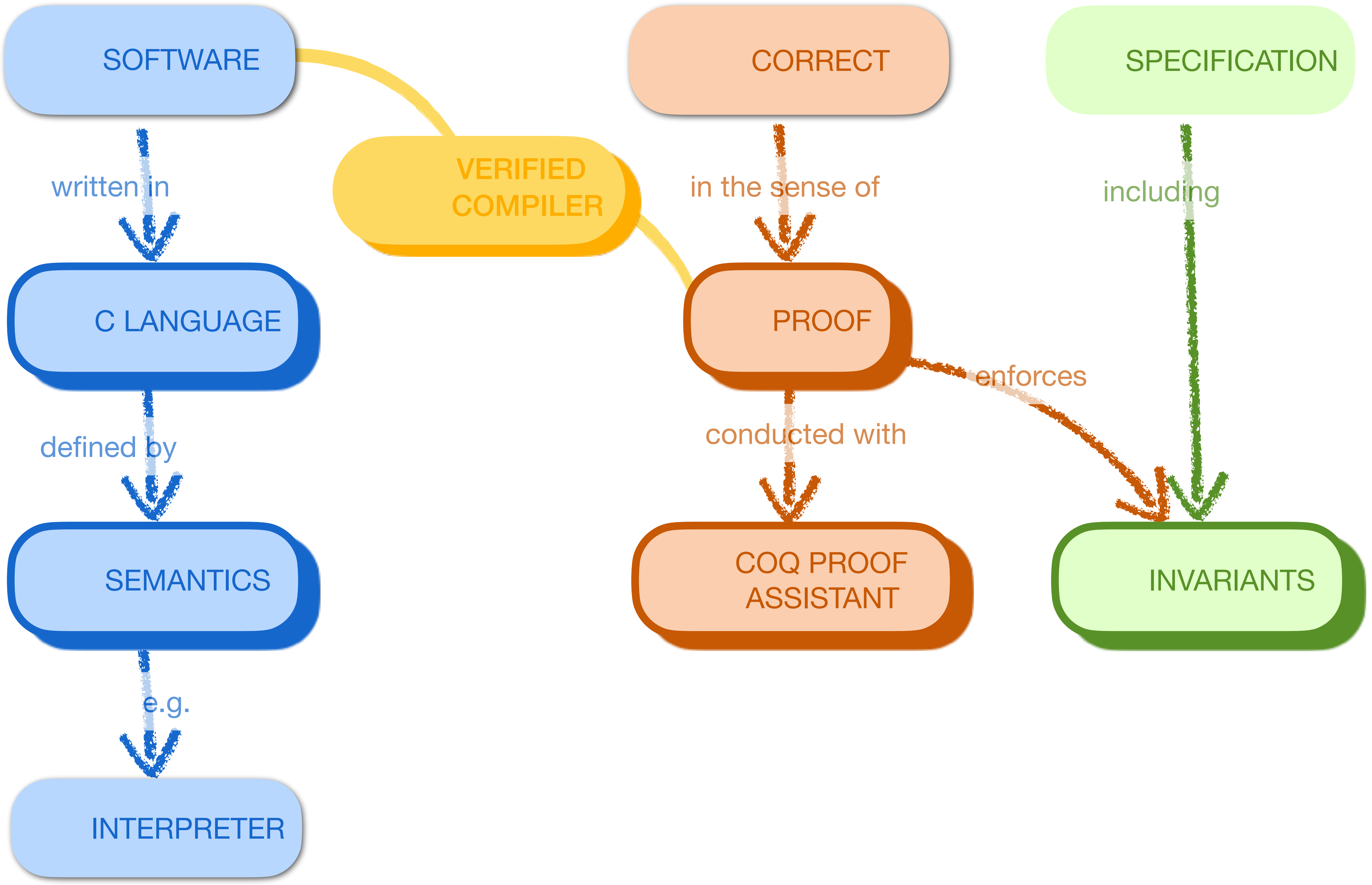
This repository contains the Coq sources for the course "Mechanized semantics" given by Xavier Leroy at Collège de France in 2019-2020.

This is the English version of the Coq sources. La version commentée en français est disponible ici.

An HTML pretty-printing of the commented sources is also available:

1. The semantics of an imperative language
   - Module IMP: the imperative language IMP and its various semantics.
   - Library Sequences: definitions and properties of reduction sequences.
2. Formal verification of a compiler
   - Module Compil: compiling IMP to a virtual machine.
   - Library Simulation: simulation diagrams between two transition systems.

Part 2:
early intuitions



The miscompilation risk

Compilers still contain bugs!

> We found and reported **hundreds** of previously **unknown** bugs [...]. Many of the bugs we found cause a compiler to emit incorrect code **without any warning**. 25 of the bugs we reported against GCC were classified as **release-blocking**.

[Yang, Chen, Eide, Regehr. Finding and understanding bugs in C compilers. PLDI'11]

Verified compilation

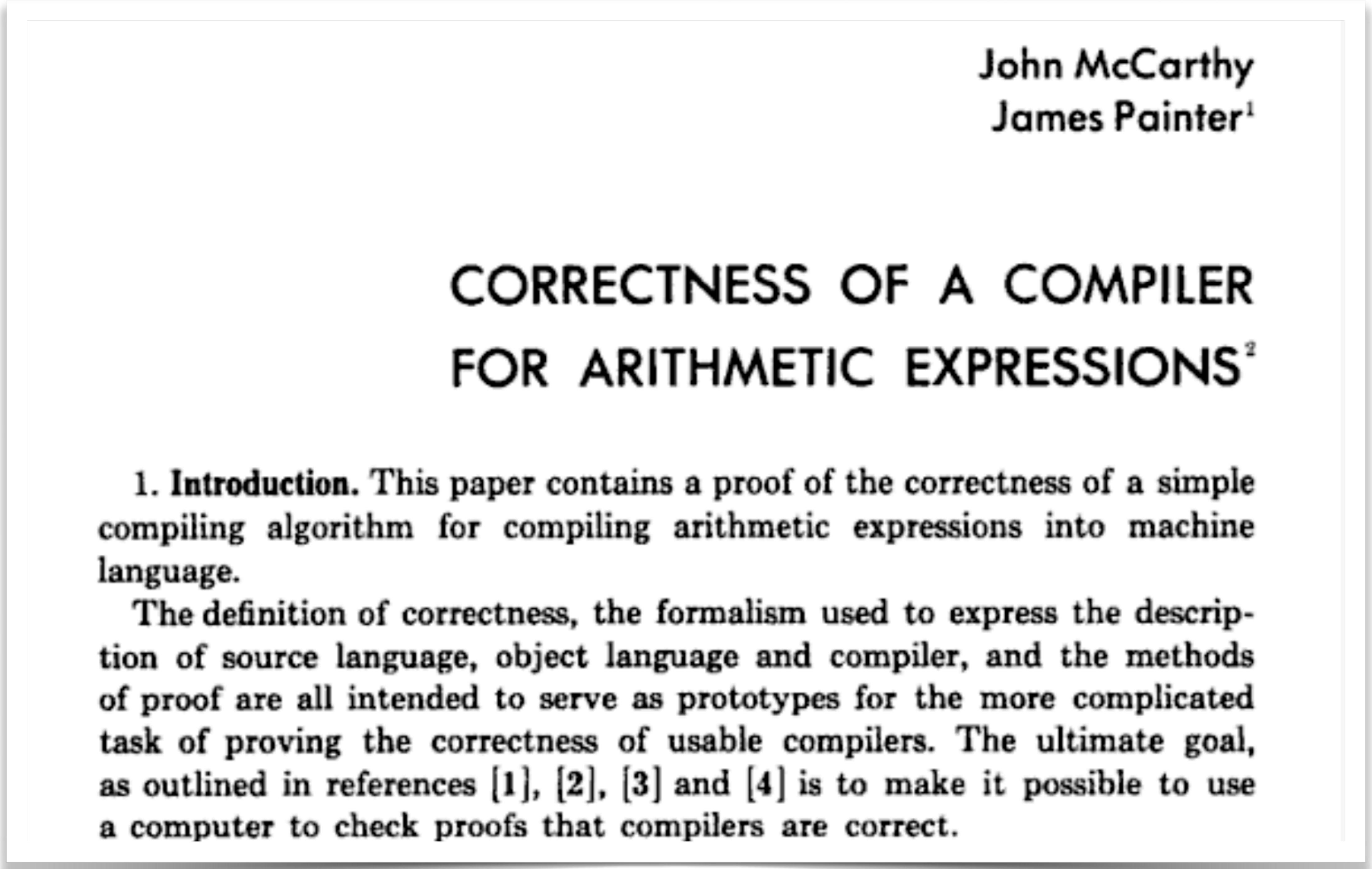
Compilers are complicated programs, but have a rather simple end-to-end specification:

> The generated code must behave as prescribed by the semantics of the source program.

This specification becomes mathematically precise as soon as we have formal semantics for the source language and the machine language.

Then, a formal verification of a compiler can be considered.

An old idea …



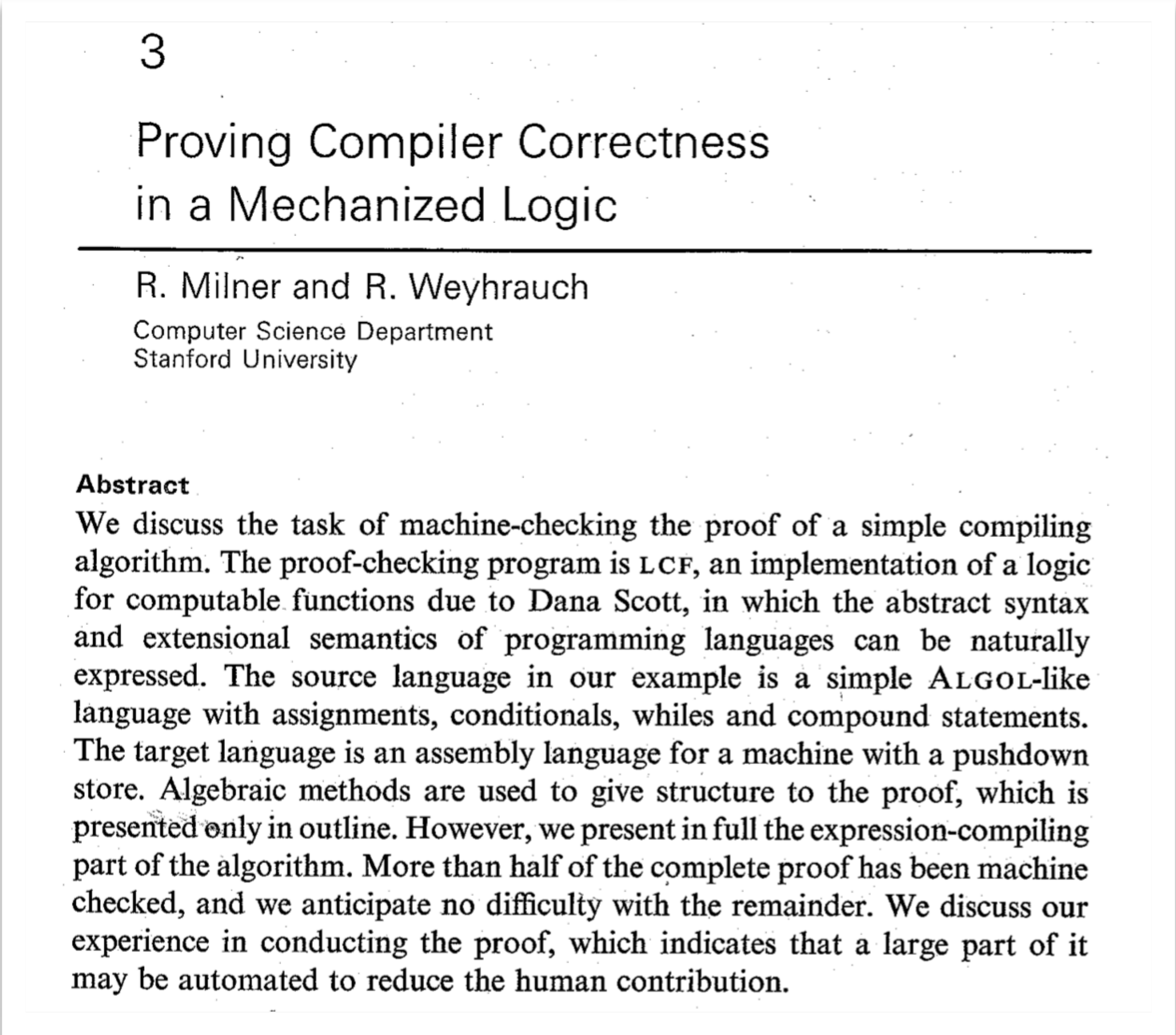
John McCarthy
James Painter[1]

CORRECTNESS OF A COMPILER FOR ARITHMETIC EXPRESSIONS[2]

1. **Introduction.** This paper contains a proof of the correctness of a simple compiling algorithm for compiling arithmetic expressions into machine language.

The definition of correctness, the formalism used to express the description of source language, object language and compiler, and the methods of proof are all intended to serve as prototypes for the more complicated task of proving the correctness of usable compilers. The ultimate goal, as outlined in references [1], [2], [3] and [4] is to make it possible to use a computer to check proofs that compilers are correct.

Mathematical Aspects of Computer Science, 1967



3

Proving Compiler Correctness in a Mechanized Logic

R. Milner and R. Weyhrauch
Computer Science Department
Stanford University

Abstract

We discuss the task of machine-checking the proof of a simple compiling algorithm. The proof-checking program is LCF, an implementation of a logic for computable functions due to Dana Scott, in which the abstract syntax and extensional semantics of programming languages can be naturally expressed. The source language in our example is a simple ALGOL-like language with assignments, conditionals, whiles and compound statements. The target language is an assembly language for a machine with a pushdown store. Algebraic methods are used to give structure to the proof, which is presented only in outline. However, we present in full the expression-compiling part of the algorithm. More than half of the complete proof has been machine checked, and we anticipate no difficulty with the remainder. We discuss our experience in conducting the proof, which indicates that a large part of it may be automated to reduce the human contribution.

Machine Intelligence (7), 1972

Now taught as an exercise

(Mechanized semantics: when machines reason about their languages, X.Leroy)
(Software foundations, B.Pierce et al.: exercise stack compiler correct)

```
type exp = Nb of int | Id of string | Plus of exp * exp
```

semantics (aeval, s_execute)

compiler (compile)

```
type state = string → int
```

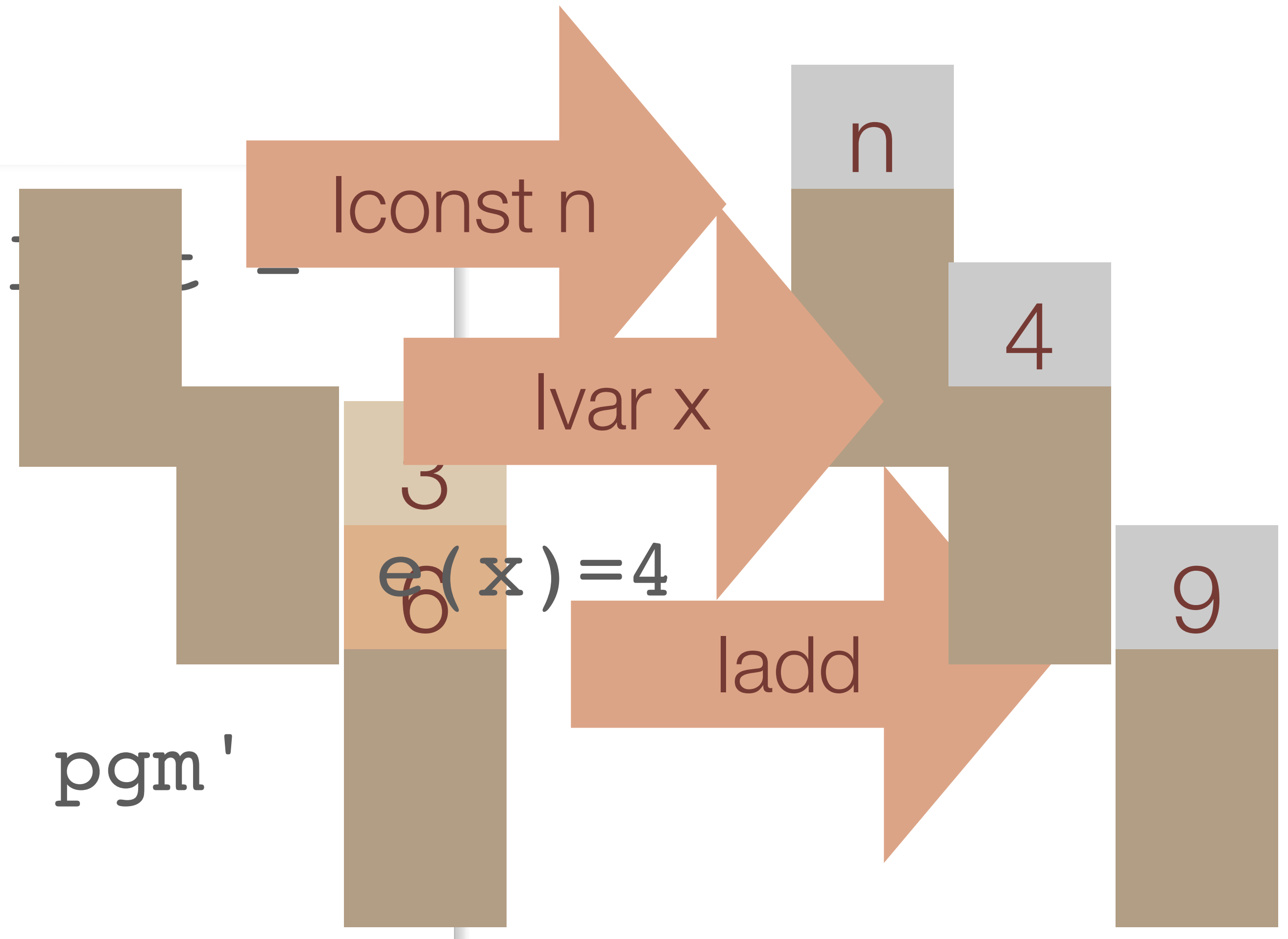
```
let rec aeval (s:state)(a:exp): int =
match a with
   | Nb n → n
   | Id x → s x
   | Plus (a1,a2) → (aeval s a1)+(aeval
```

compilation

```
let rec compile (a:exp): instr list =  match a with
   | Nb n →   [ Iconst n ]
   | Id x →   [ Ivar x ]
   | Plus (a1,a2) → (compile a1)@ (compile a2)@ [Iadd]
```

```
type instr = Iconst of int | Ivar of string  | Iadd
```

```
let rec s_execute (s:state)(stack: int list)(pgm: instr list): int list =
  match (pgm, stack) with
   | ([], _) → stack
   | (Iconst n :: pgm', _) → s_execute s (n :: stack) pgm'
   | (Ivar x :: pgm', _) → s_execute s (s x :: stack) pgm'
   | (Iadd :: pgm', n:: m :: stack') → s_execute s ((m+n) :: stack') pgm'
   | (_ :: pgm', _) → s_execute s stack pgm'
```



Now taught as an exercise

(Mechanized semantics: when machines reason about their languages, X.Leroy)
(Software foundations, B.Pierce et al.: exercise stack_compiler_correct)

```
Fixpoint aeval(s:state)(a:aexp):nat := ...
```

compilation

```
Fixpoint compile(a:aexp): list sinstr := ...
```

```
Fixpoint s_execute(s:state)(stack:list nat)(prog:list sinstr):list nat := ...
```

semantics (aeval, s_execute)

compiler (compile)

interactive proof

```
Theorem s_compile_correct: ∀ s a,
 s_execute s [] (compile a) = [aeval s a].
Proof.
```

Now taught as an exercise

(Mechanized semantics: when machines reason about their languages, X.Leroy)
(Software foundations, B.Pierce et al.: exercise stack_compiler_correct)

semantics (aeval, cexec)

compiler (compile)

interactive proof

```
Fixpoint aeval(s:state)(a:aexp):nat := ...
```

compilation

```
Fixpoint compile(a:aexp): list sinstr := ...
```

```
Fixpoint s_execute(s:state)(stack:list nat)(prog:list sinstr):list nat := .
```

extraction

```
Theorem s_compile_correct_aux: ∀ s a stack,
 s_execute s stack (compile a) = aeval a :: stack.
Proof.
   induction a; (* … *)
Qed.
```

proof by induction on the structure of a

toy-compiler.ml

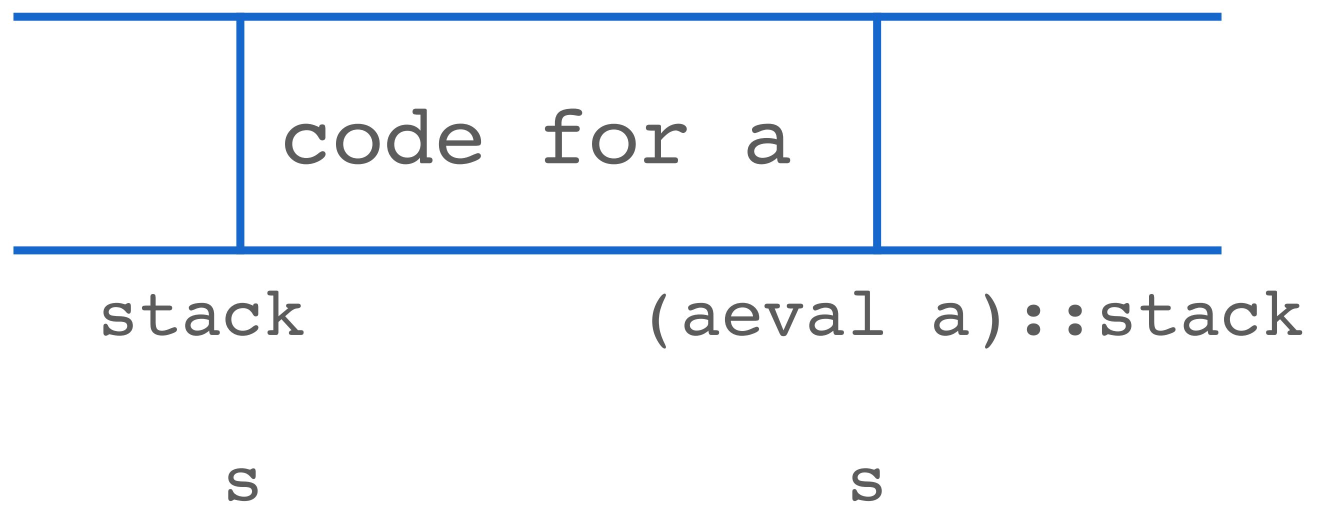
OCaml

```
Theorem s_compile_correct: ∀ s a,
 s_execute s [] (compile a) = [aeval s a].
Proof.
   intros. apply s_compile_correct_aux.
Qed.
```

```
Extraction compile.
```

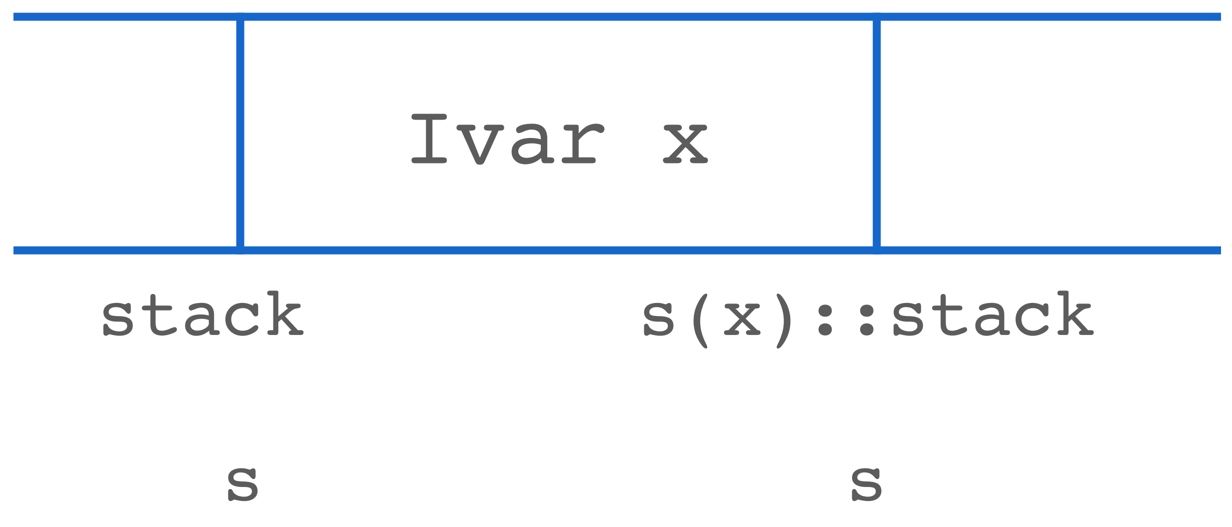
Proof by induction on the structure of expressions

```
Theorem s_compile_correct_aux: ∀ s a stack,
 s_execute s stack (compile a) = aeval a :: stack.
Proof.
   induction a; (* … *)
Qed.
```



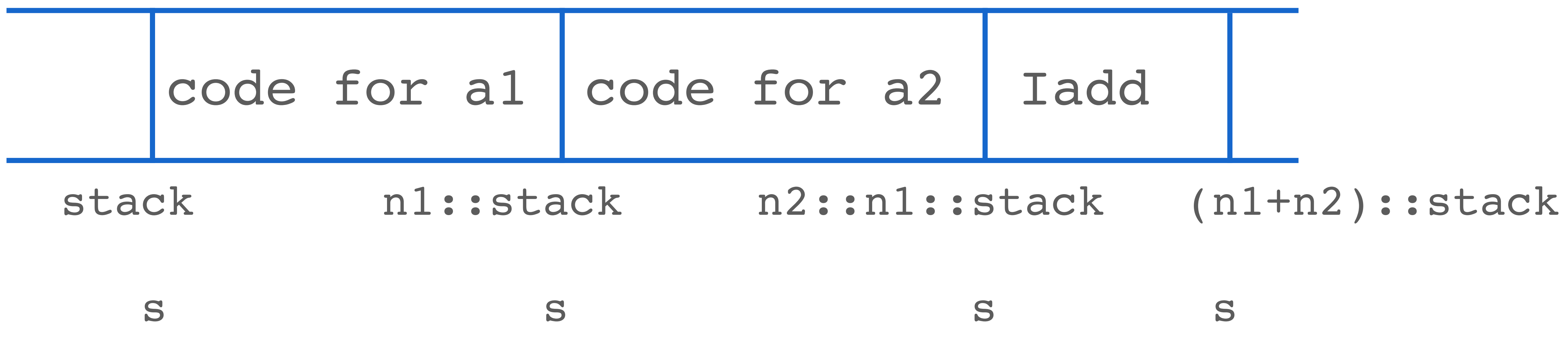
Base case: a = Id x

```
Theorem s_compile_correct_aux: ∀ s a stack,
 s_execute s stack (compile a) = aeval a :: stack.
Proof.
   induction a; (* … *)
Qed.
```



Inductive case: a = Plus (a1,a2)

```
Theorem s_compile_correct_aux: ∀ s a stack,
 s_execute s stack (compile a) = aeval a :: stack.
Proof.
   induction a; (* … *)
Qed.
```



Course outline

Mechanized semantics of imperative languages

Formal verification in Coq of a non-optimizing compiler for a simple imperative language (from IMP to VM language)

- study of proof techniques for semantic preservation

Extension of these ideas to CompCert, a realistic C compiler

The CompCert formally verified compiler

(X.Leroy, S.Blazy et al.) https://compcert.org

A moderately optimizing C compiler

Targets several architectures (PowerPC, ARM, RISC-V and x86)

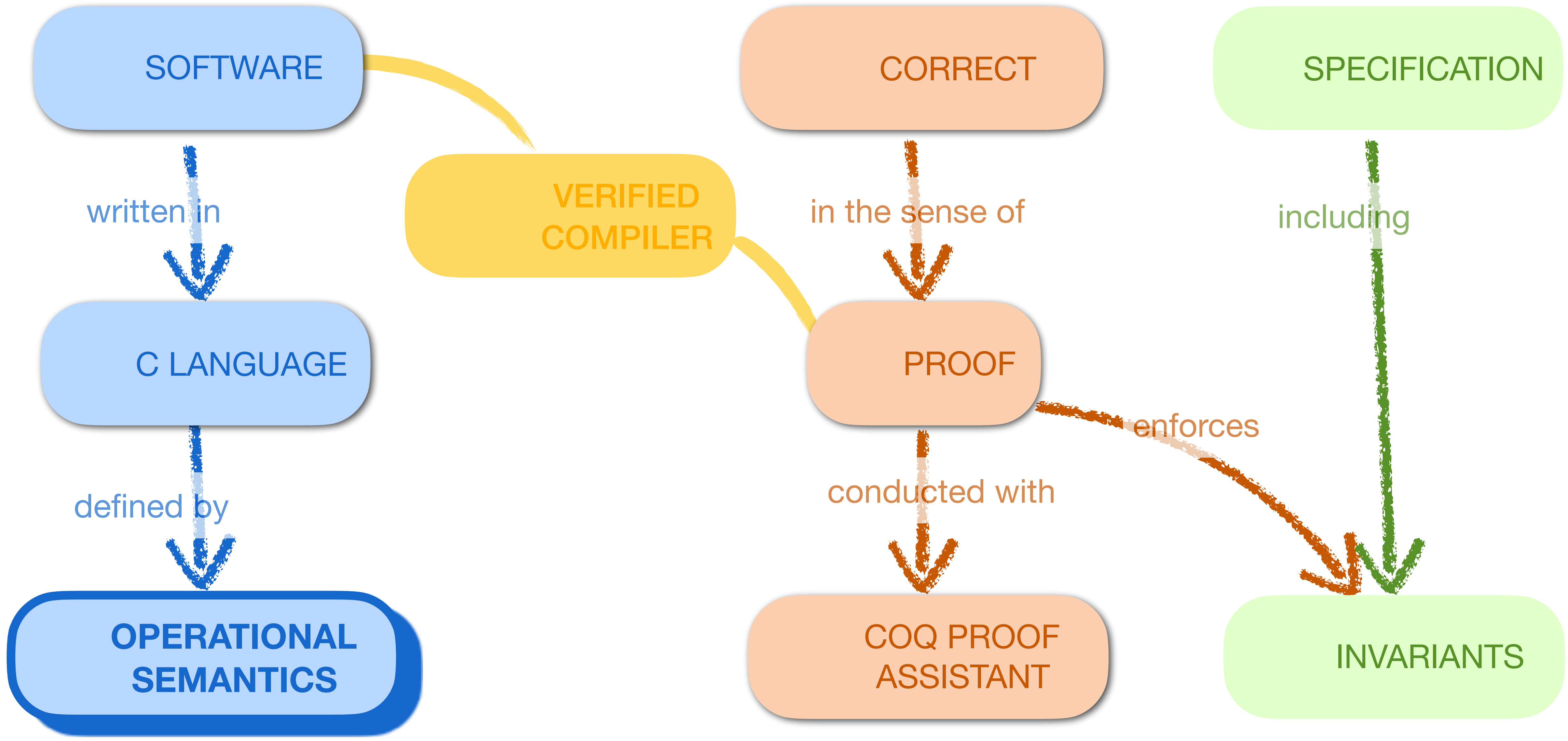
Programmed and verified using the Coq proof assistant

Shared infrastructure for ongoing research

Used in commercial settings (for emergency power generators and flight control navigation algorithms) and for software certification - AbsInt company
Improved performances of the generated code while providing proven traceability information

ACM Software System award 2021
ACM SIGPLAN Programming Languages Software award 2022

Part 3:
basics of
verified compilation



Compiling IMP commands

```
a:aexp := n | x | a1 + a2 | a1 - a2
```

```
b:bexp := true | false | a1 = a2 | a1 ≤ a2
        | not b | b1 and b2
```

boolean expressions

```
c:com := skip
    | x := a
    | c1 ; c2
    | if b then c1 else c2
    | while b do c done
```

semantics (aeval, beval, ceval)

Big-step style for the semantics of IMP expressions, with a functional definition

```
aeval (s:state) (a:aexp)
```

Big-step style for commands, using a relational definition c / s ⇒ s'

For IMP commands, a relational definition is better than a functional definition.

A big-step semantics for IMP

Relation c / s ⇒ s’

$$\text{skip} / s \Rightarrow s$$

$$x := a / s \Rightarrow s\,[x \leftarrow (\text{aeval } a\ s)]$$

$$\frac{c1 / s1 \Rightarrow s \qquad c2 / s \Rightarrow s2}{(c1\ ;\ c2) / s1 \Rightarrow s2}$$

$$\frac{\text{eval } s\ b = \text{true} \qquad c1 / s \Rightarrow s1}{(\text{if } b \text{ then } c1 \text{ else } c2) / s \Rightarrow s1}$$

$$\frac{\text{eval } s\ b = \text{false} \qquad c2 / s \Rightarrow s2}{(\text{if } b \text{ then } c1 \text{ else } c2) / s \Rightarrow s2}$$

$$\frac{\text{eval } s\ b = \text{false}}{(\text{while } b \text{ do } c \text{ end}) / s \Rightarrow s}$$

$$\frac{\text{eval } s1\ b = \text{true} \qquad c / s1 \Rightarrow s \qquad \text{while } b \text{ do } c \text{ end} / s \Rightarrow s2}{(\text{while } b \text{ do } c \text{ end}) / s1 \Rightarrow s2}$$

This rule can not be defined by a terminating Coq function.

In Coq: an inductive predicate (`cexec s c s')`)

Extending the VM language: components of the machine

The **code** C: a fixed list of instructions

The **program counter** pc: an integer giving the position of the currently executing instruction in C

The **store** and the **stack**, as before

Inspiration: old HP pocket calculators, the Java Virtual Machine

Extending the VM language: instruction set

compil.v

```
i  := Iconst(n)
    | Ivar(x)
    | Iadd
    | Isetvar(x)          pop an integer and assign it to variable
    | Ibranch(d)          skip forward or backward d instructions
    | Iopp                pop one integer, push its opposite
    | Ibeq(d1,d0)         pop 2 integers, skip d1 instructions if =, d0 if ≠
    | Ible (d1 d0: Z)     pop 2 integers, skip d1 instructions if ≤, if >
    | Ihalt               stop execution
```

branch offset relative to the next instruction

By default, each instruction increments pc by 1.
Exception: branch instructions increment it by 1+d.

x := x + 1

```
Definition ex_code1:code := Ivar "x" :: Iconst 1 :: Iadd :: Isetvar "x" :: nil.
Definition ex_code2:code := Ivar "x" :: Iconst 1 :: Iadd :: Isetvar "x" :: Ibranch (-5) :: nil.
```

VM semantics

compil.v

Small-step semantics, given by a transition relation $s \to s'$ representing the execution of one instruction

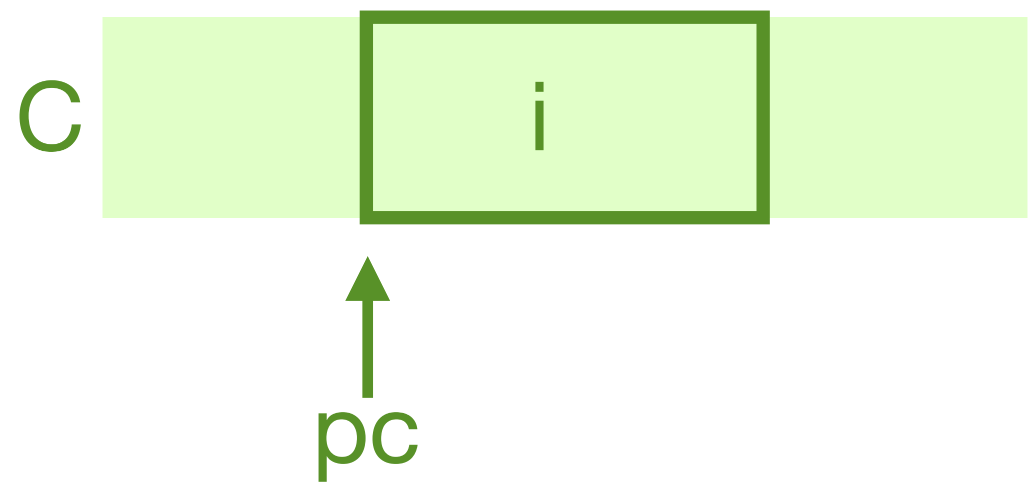
```
Definition code := list instr.
Definition stack := list Z.
Definition store := ident → Z.
Definition config := (Z * stack * store).
```

pc, position of
the currently executing
instruction

$S \to S'$

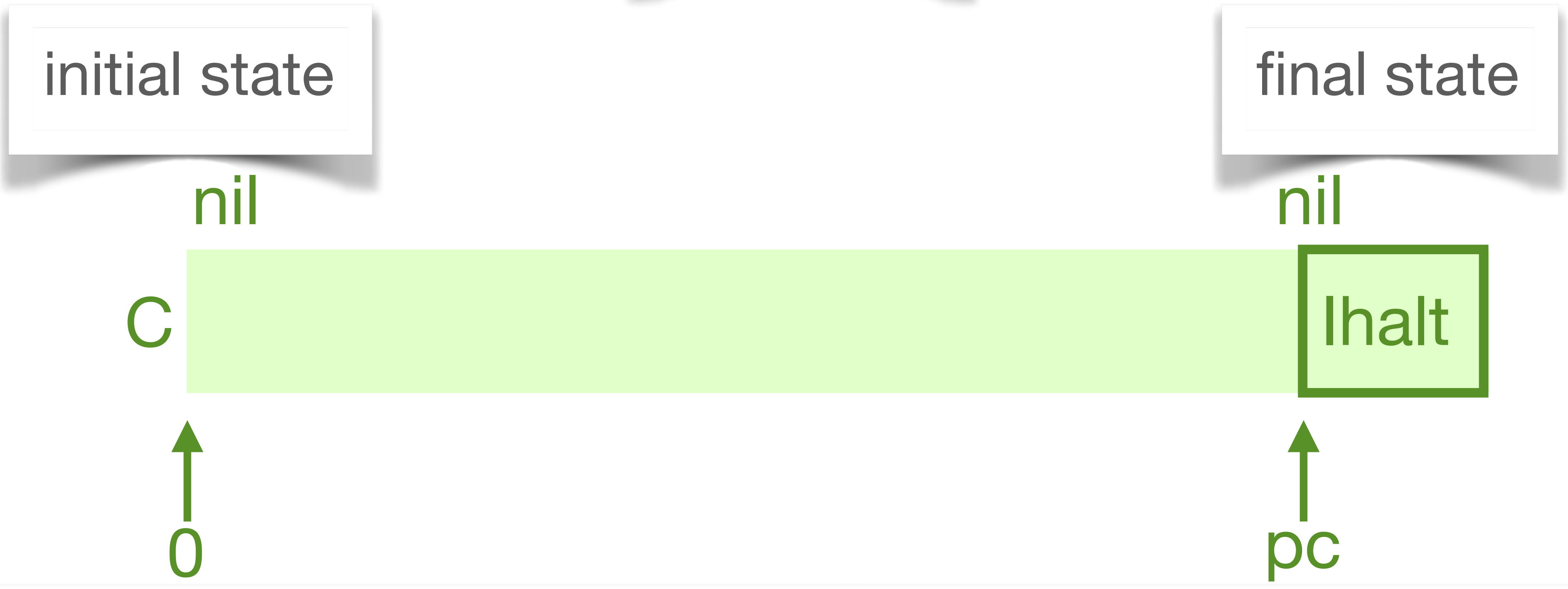
```
Inductive transition (C:code): config → config → Prop := …
```

instr_at C pc = Some i



Executing machine programs

By iterating the transition relation $S \to S'$




```
Definition transitions (C: code): config → config → Prop :=
  star (transition C).
```

reflexive transitive closure

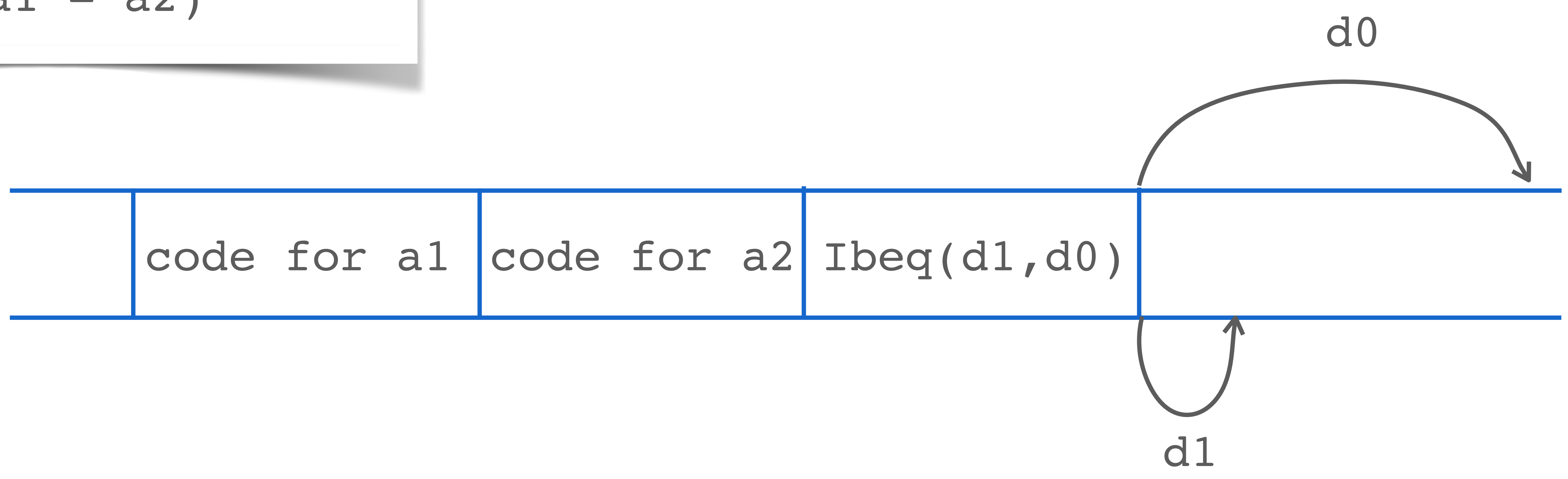
initial stores

final stores

```
Definition machine_terminates (C: code) (s_init s_final: store) :=
  ∃ pc, transitions C (0, nil, s_init) (pc, nil, s_final)
          ∧ instr_at C pc = Some Ihalt.
```

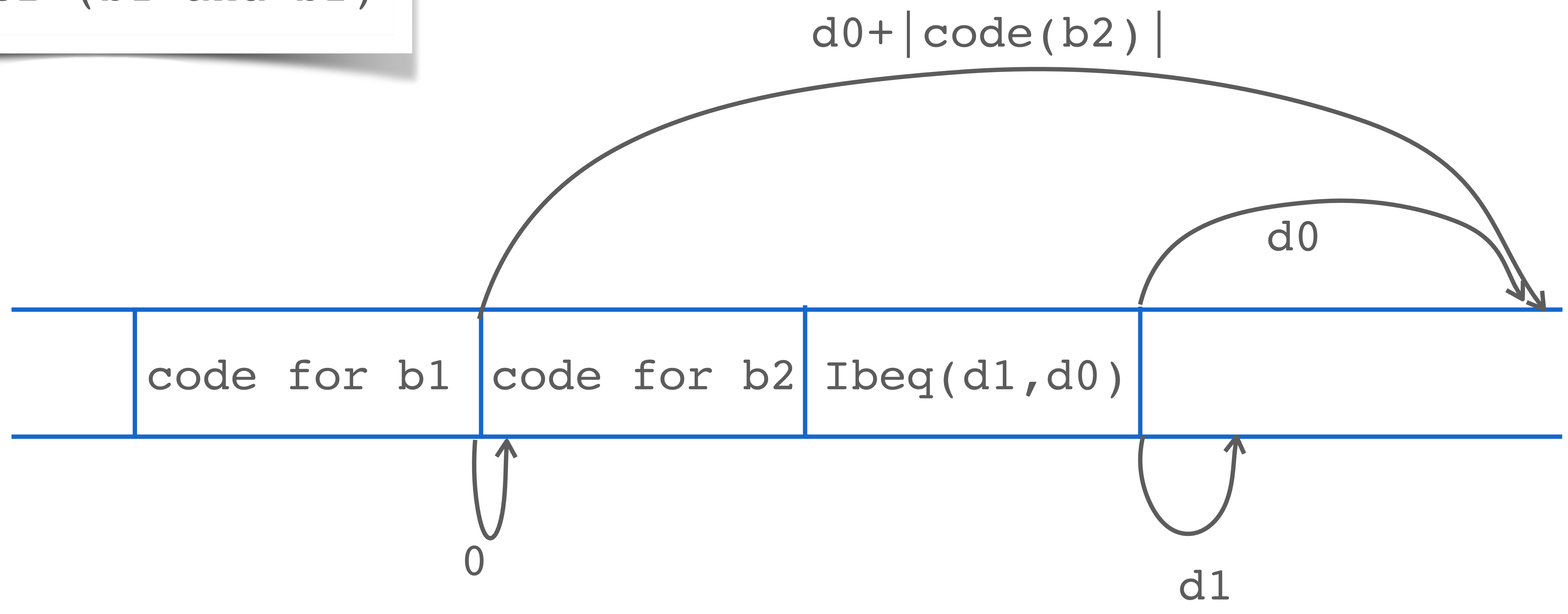
Compilation of boolean expressions

```
code for (a1 = a2)
```



Short-circuiting `and` expressions

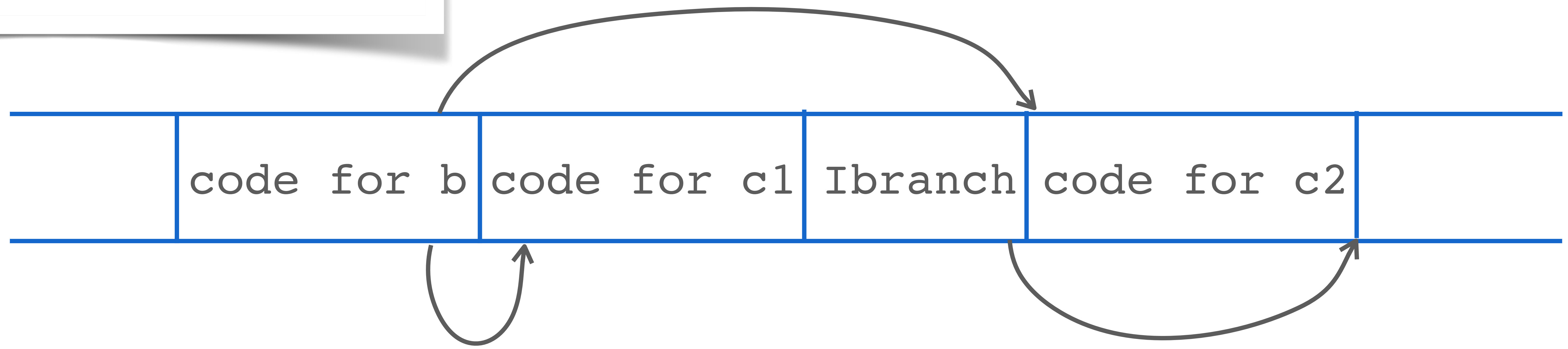
```
code for (b1 and b2)
```



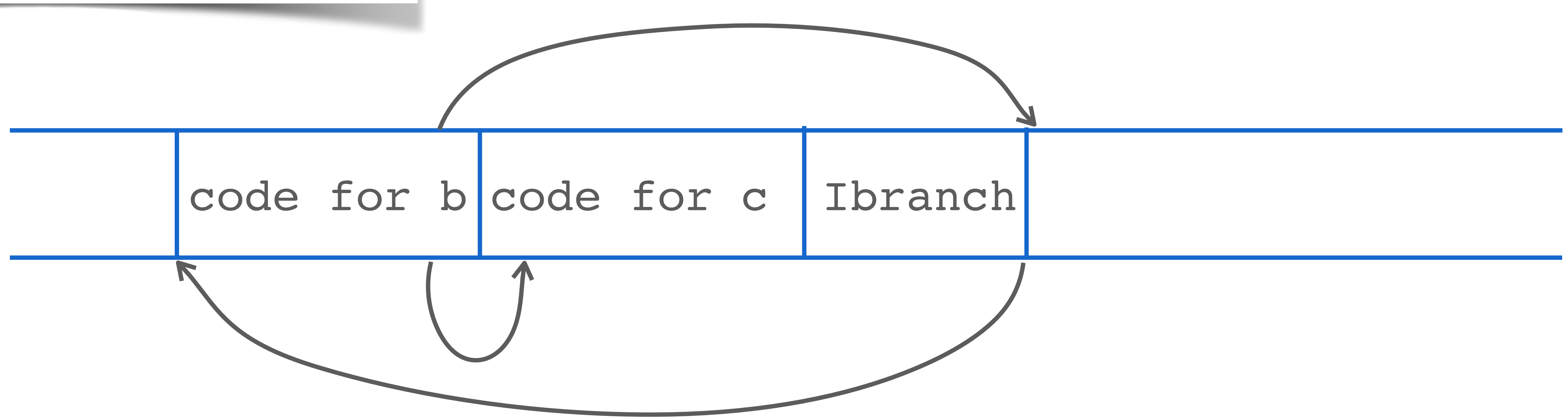
If b1 evaluates to false, so does b1 and b2: no need to evaluate b2

Compilation of commands
(`compile_com`)

```
code for (CIf b c1 c2)
```




```
code for (CWhile b c)
```



Verifying the compilation of commands

compil.v

semantics (cexec, transitions)

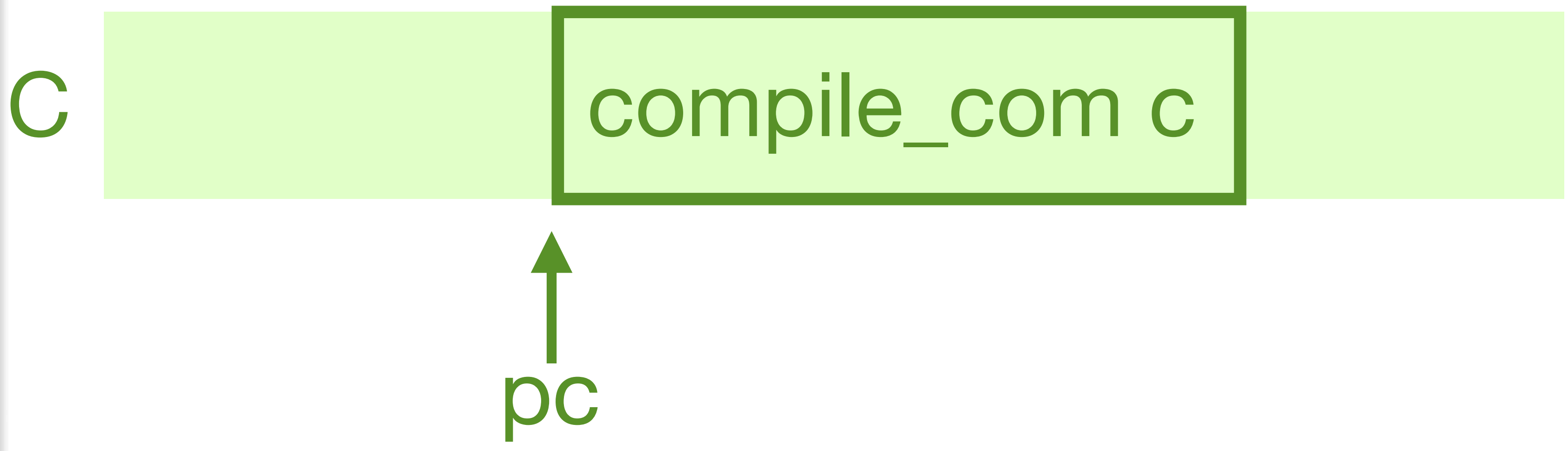
compiler (compile_com)

interactive proof

```
Lemma compile_com_correct_terminating:
  ∀ s c s',
  cexec s c s' →
  ∀ C pc stack,
  code_at C pc (compile_com c) →
  transitions C
      (pc, stack, s)
      (pc + codelen (compile_com c), stack, s').
```

remember s_compile_correct_aux!

length of the list



proof by induction on the derivation of cexec s c s’

Compiler correctness

compil.v

semantics (cexec, machine_terminates)

compiler (compile_program)

interactive proof

```
Definition compile_program (p: com) : code :=
  compile_com p ++ Ihalt :: nil.

Theorem compile_program_correct_terminating:
  ∀ s c s',
  cexec s c s' →
  machine_terminates (compile_program c) s s'.
```

```
Definition machine_terminates (C: code) (s_init s_final: store) :=
  ∃ pc, transitions C (0, nil, s_init) (pc, nil, s_final)
        ∧ instr_at C pc = Some Ihalt.
```

Part 3: summary

« The generated code must behave as prescribed by the semantics of the source program. »

IMP

Expressions: big-step semantics
Commands: **big-step** semantics

compiler

VM

Instructions: small-step semantics

compiler correctness theorem

is about

big-step semantics

observe

behaviors

termination

$c \,/\, s \Rightarrow s'$

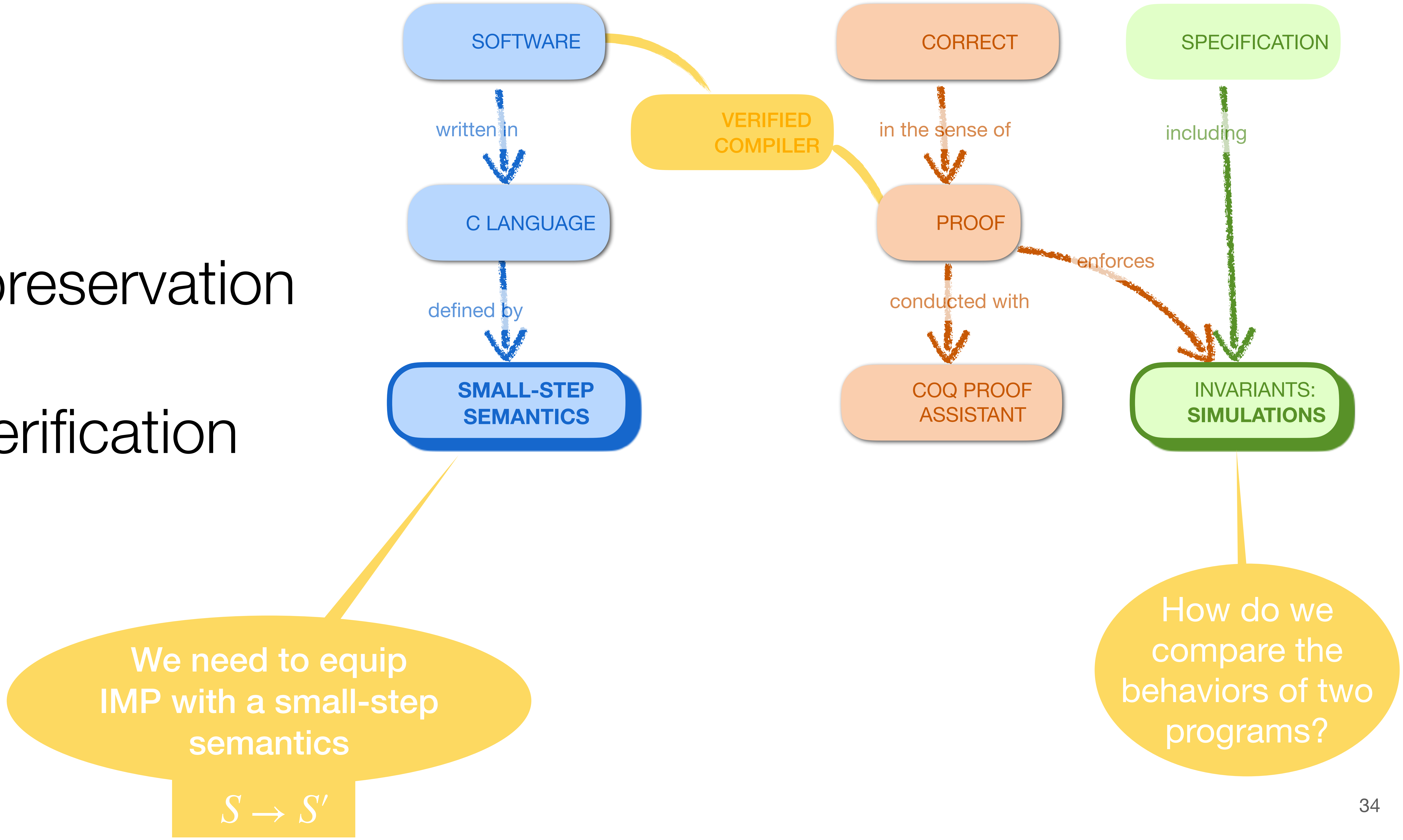
```
Theorem compile_program_correct_terminating:
  ∀ s c s',
  cexec s c s' →
  machine_terminates (compile_program c) s s'.
```

What about diverging programs?

How do we compare the behaviors of two programs?

This is not enough to conclude that the compiler is correct!

Part 4:
semantic preservation
and
compiler verification



What should be preserved?
Observable behaviors

$S \to^{*} S'$ $S \to^{+} S'$

`x := 1;` normal termination

`Ihalt`

`return 0;`

finite sequence of transitions to a final state

$S \overset{\infty}{\to}$

infinite sequence of transitions

divergence

`while true do skip end`

`Ibranch (-1)`

`for(;;) { }`

observable behaviors

impossible

abnormal termination (a.k.a. going wrong)

`Iadd`

`x = 1/0 ;`

finite sequence of transitions to a state that is stuck and not final

We need to equip IMP with a small-step semantics

$S \to S'$

IMP VM C

Notions of semantic preservation: bisimulation

The source program S and the compiled program C have exactly the same behaviors.

- Every possible behavior of S is a possible behavior of C.
- Every possible behavior of C is a possible behavior of S.

Example for the IMP to VM compiler

- (compile_com c) terminates if and only if c terminates, with the same final store
- (compile_com c) diverges if and only if c diverges
- (compile_com c) never goes wrong

Forward simulation

Forward simulation from a source program S to a compiled code C:
every possible behavior of S is a possible behavior of C

Example:

- theorem compile_program_correct_terminating
- If C diverges, (compile_com C) diverges

This looks insufficient: what if C has more behaviors than S?
For instance, if C can terminate or go wrong?

Forward simulation + determinism = bismimulation

A language is deterministic if every program has only one behavior.

Lemma
If the target language is deterministic, forward simulation implies backward simulation and therefore bisimulation.

Proof
Let C be a compiled program and S its source.
Let b be a behavior of C and b’ a behavior of S.
By forward simulation, b’ is a behavior of C.
By determinism of C, b’ = b.
Hence every behavior b of C is a behavior of S.

Reducing non-determinism during compilation

The C language is not deterministic: the evaluation order is partially unspecified.

```
int x = 0;
int f(void) { x = x + 1; return x; }
int g(void) { x = x - 1; return x; }
```

The expression `f()+g()` can evaluate either to:

- 1 if `f()` is evaluated first (returning 1), then `g()` (returning 0);
- -1 if `g()` is evaluated first (returning 1), then `f()` (returning 0).

Every C compiler chooses one evaluation order at compile-time.
The compiled code therefore has fewer behaviors than the source program (1 instead of 2). Forward simulation and bisimulation fail.

Backward simulation, a.k.a. refinement

Backward simulation from a source program S to a compiled code C:
every possible behavior of C is a possible behavior of S.
However, C may have fewer behaviors than S.

Backward simulation suffices to show the preservation of properties established by source-level verification:

If all behaviors of S satisfy a specification Spec,
then all behaviors of C satisfy Spec as well.

Should «going wrong» behaviors be preserved?

```
#include <stdio.h>
int main()
{
   int x;
   x = 1 / 0;
   return 0;
}
```

Compilers routinely optimize away going-wrong behaviors.

This program goes wrong.

However, the compiler eliminates `x=1/0;` as it is dead code.

Thus, the generated code always terminates.

Justifications

- We know that the program does not go wrong (e.g. by static analysis).
- It is the programmer's responsibility to avoid going-wrong behaviors (C standards).

Should «going wrong» behaviors be preserved?

```
#include <stdio.h>
int main()
{
   int x[2] = { 12, 34 };
   printf("x[2] = %d\n", x[2]);
   return 0;
}
```

This program goes wrong.

However, the code generated by the compiler does not check the array bounds.

The generated code may crash but in general it prints an arbitrary integer and terminates normally.

This out-of-bound access is an example of an undefined behavior (according to the ISO C standard).

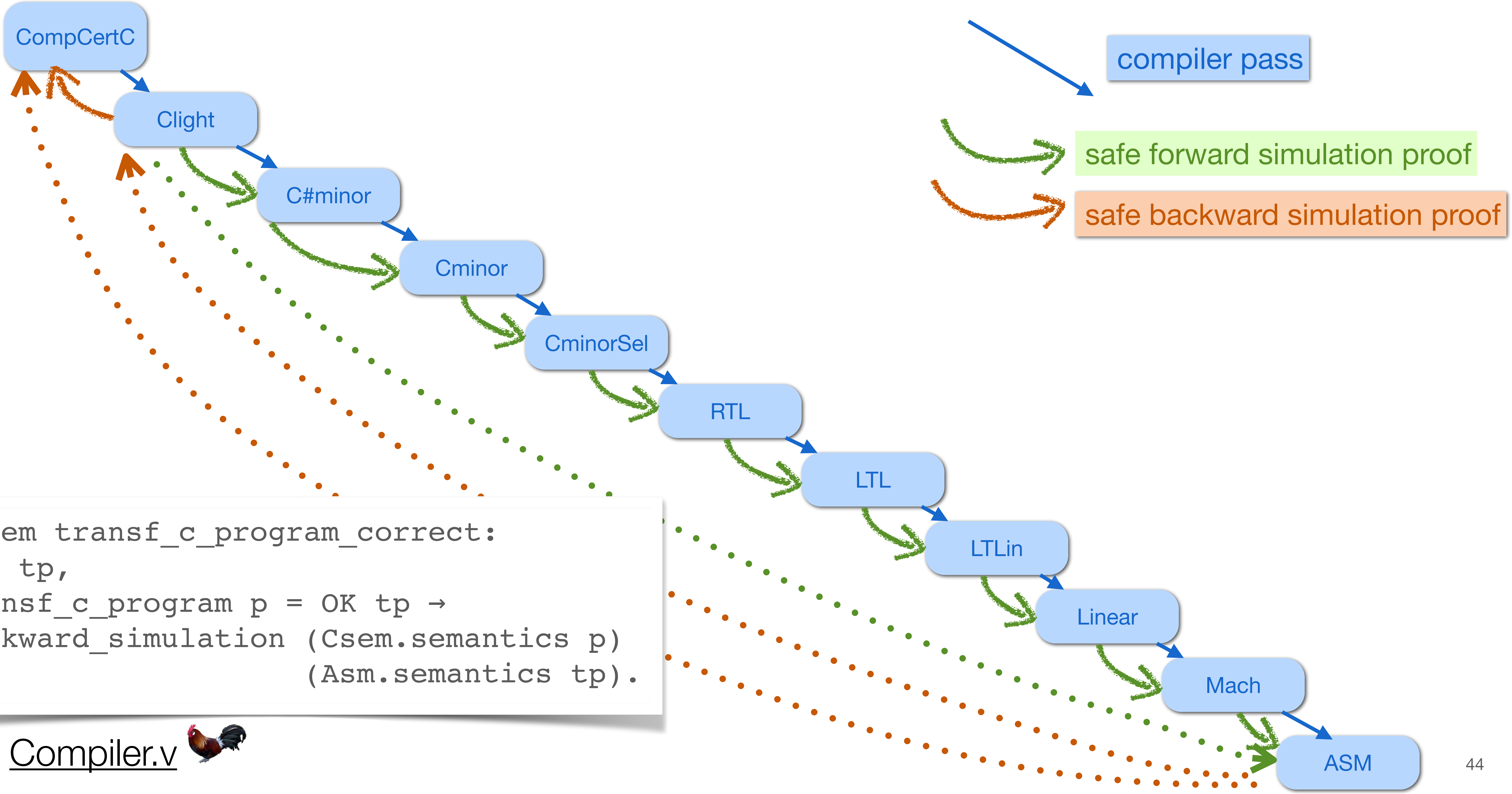
Simulations for safe programs

A program is **safe** when it either terminates or diverges.

Safe forward simulation: any behavior of the source program S other than « going wrong » is a possible behavior of the compiled code C.

Safe backward simulation: for any behavior b of the compiled code C, the source program S can either have behavior b or go wrong.

Handling multiple compilation passes




```
Theorem transf_c_program_correct:
  ∀ p tp,
  transf_c_program p = OK tp →
  backward_simulation (Csem.semantics p)
                      (Asm.semantics tp).
```

Compiler.v

Simulation diagrams

Behaviors are defined in terms of sequences of transitions.

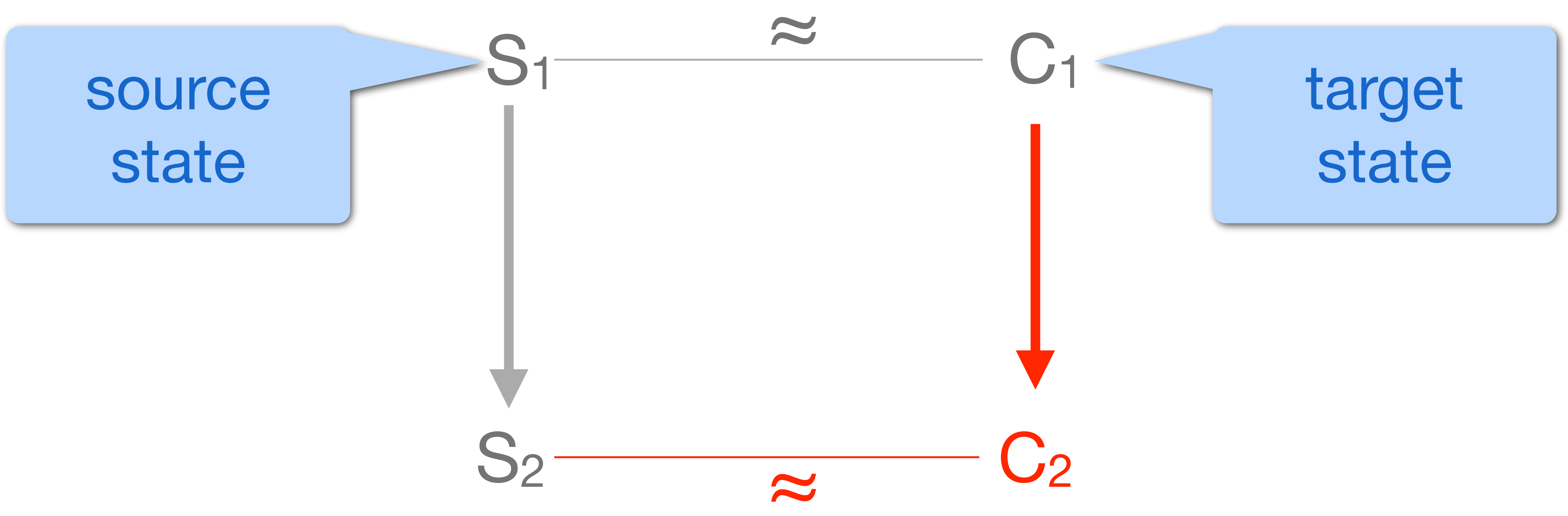
Forward simulation from a source program S to a compiled code C can be proved as follows:

- show that every transition in S is simulated by some transitions in C
- while preserving an invariant ≈ between the states of S and C

Backward simulation is similar but simulates transitions of C by transitions of S.

Lock-step simulation

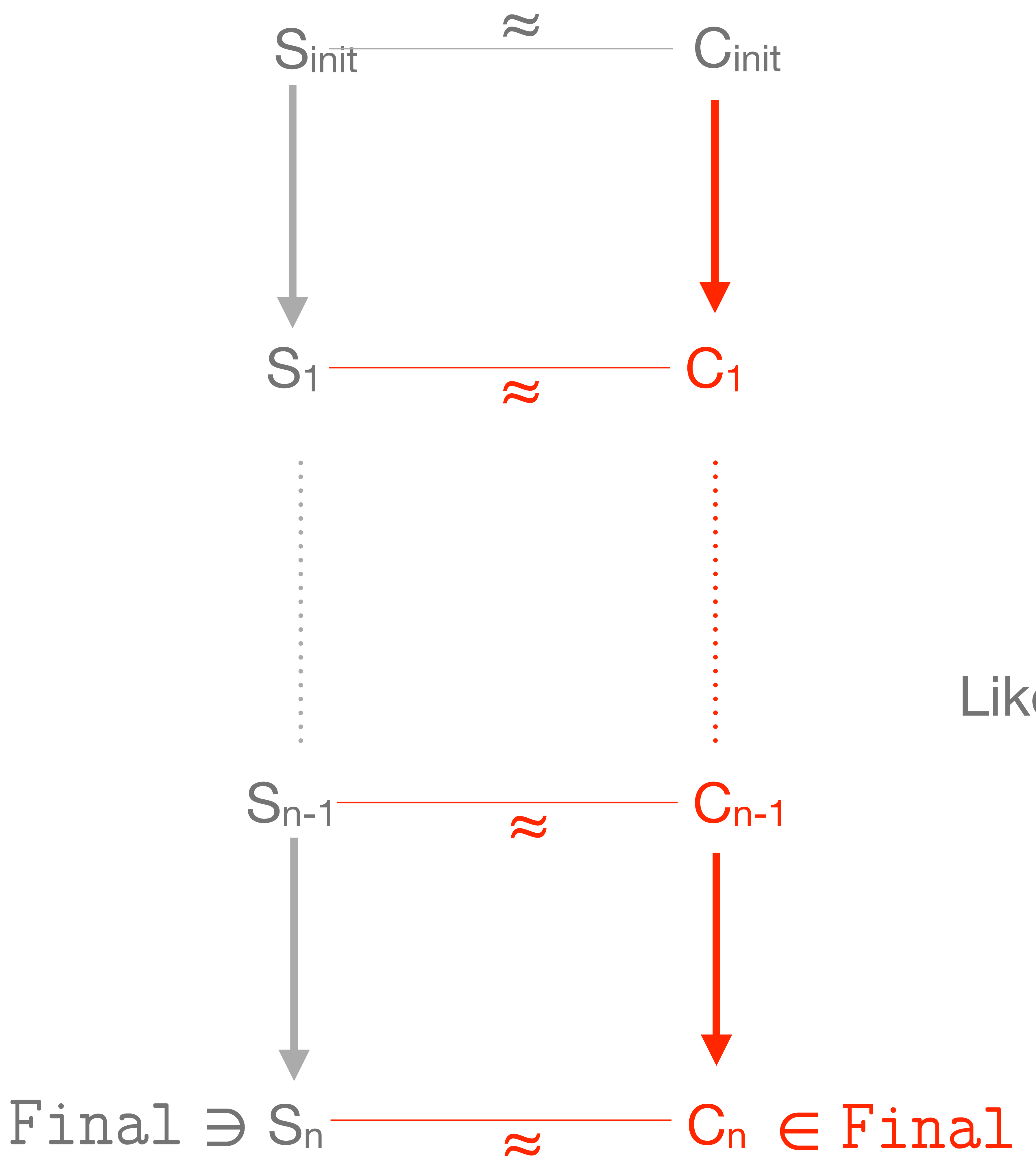
Every transition in the source S is simulated by exactly one transition in the compiled code C



Further show that initial states are related: $S_{init} \approx C_{init}$

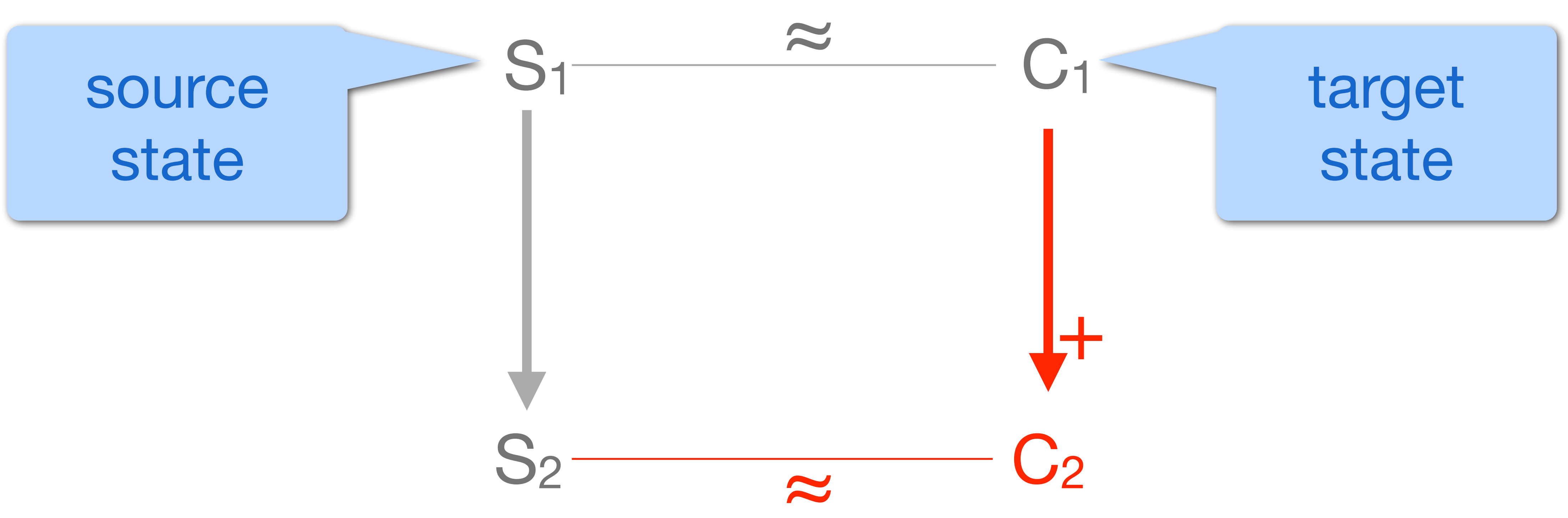
and final states are related: $S \approx C \wedge S \in \mathtt{Final} \Rightarrow C \in \mathtt{Final}$

From lock-step simulation to forward simulation



Likewise if S_{init} makes an infinity of transitions

Plus simulation



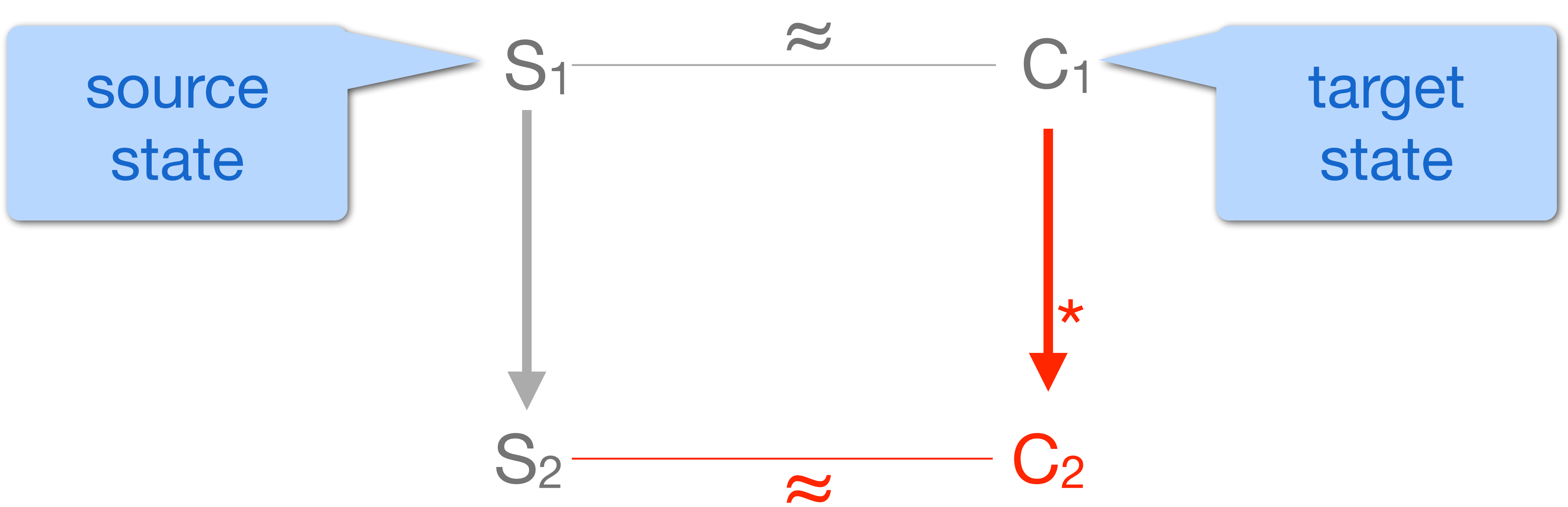
Example: compilation of `x := x + 1` into

`Ivar "x" :: Iconst 1 :: Iadd :: Isetvar "x" :: nil`

(already seen on this slide)

Forward simulation still holds

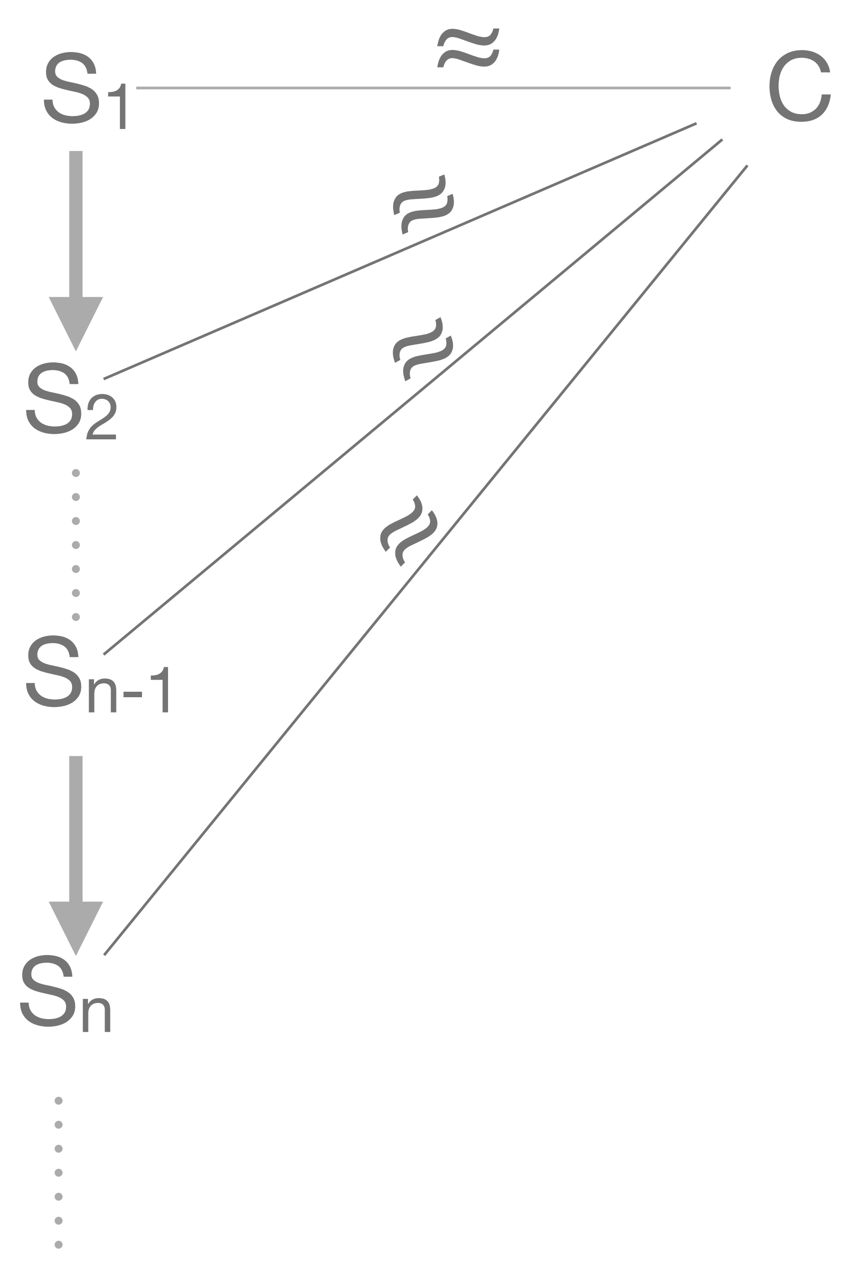
Incorrect star simulation



Forward simulation is not guaranteed:

- terminating executions are preserved,
- but diverging executions may not be preserved

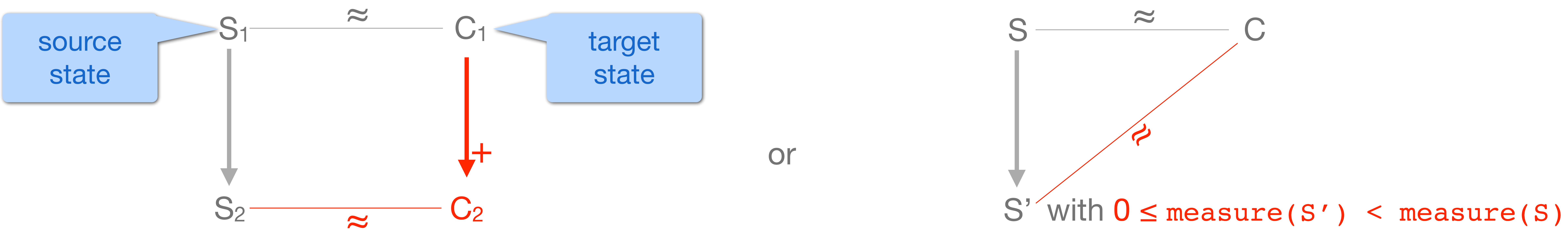
The problem of infinite stuttering



The source program diverges but the compiled code can terminate (normally or abnormally).

This denotes an incorrect optimization of a diverging program,
e.g. compiling `(while true skip)` into `skip`

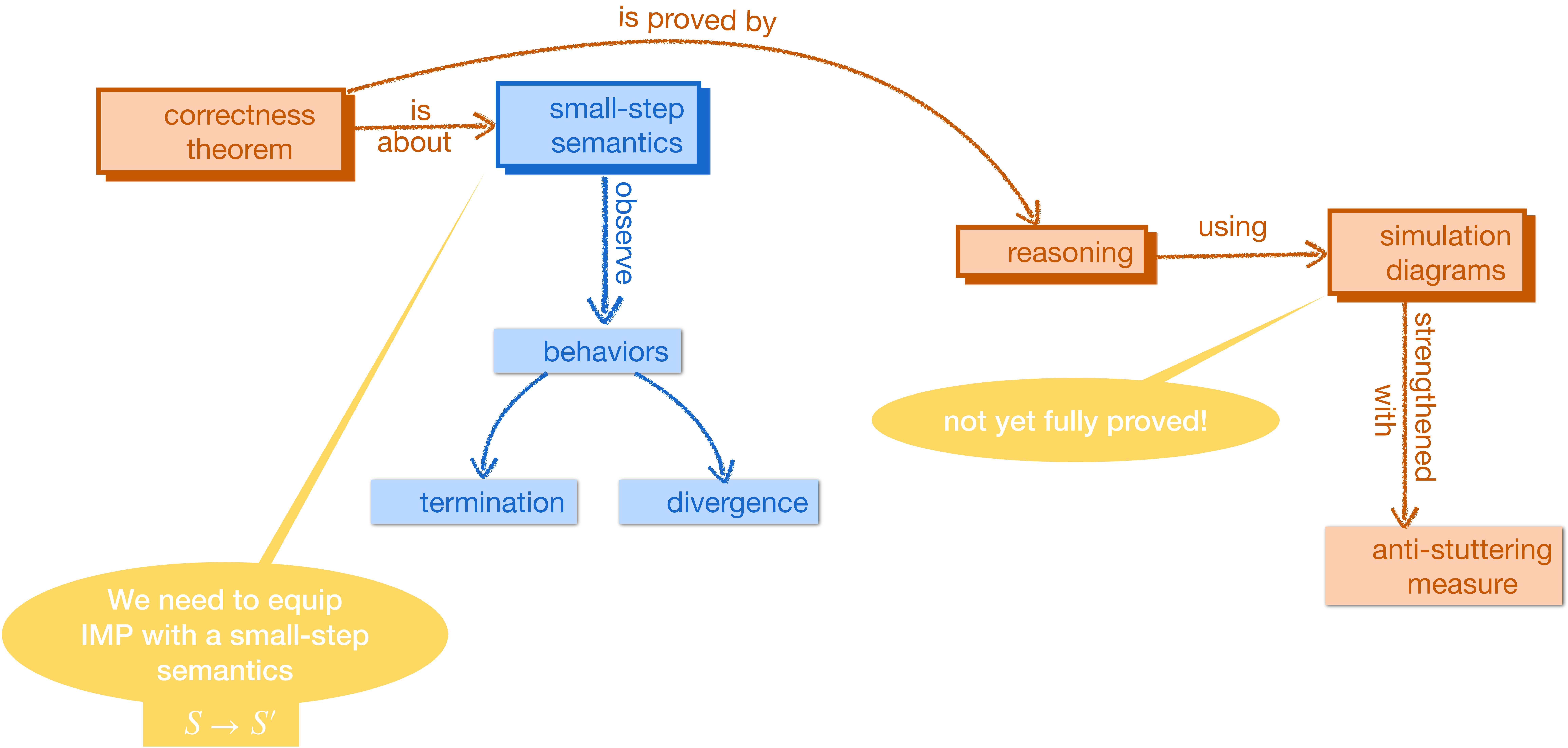
Corrected star simulation



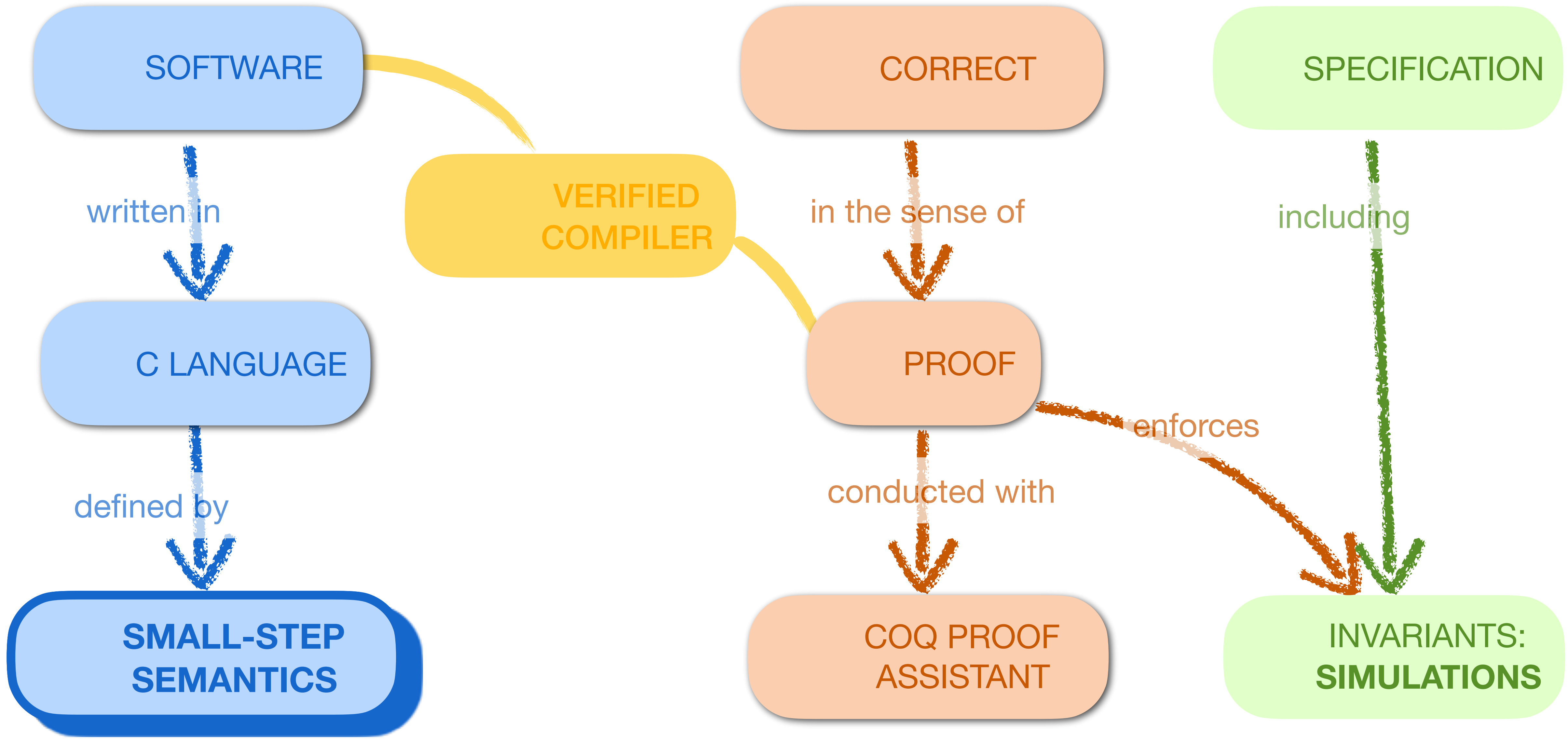
`measure(S):nat` from source states (could be to a well-founded set)

If the source program diverges, it must perform infinitely many non-stuttering steps, so the compiled code executes infinitely many transitions.

Part 4: summary



Part 5:
small-step semantics and compiler verification



We need to equip IMP with a small-step semantics

$S \to S'$

A small-step semantics for IMP

Relation $c / s \to c' / s'$

big-step semantics for expressions

$x := a / s \to \text{skip} / s\ [x \leftarrow (\text{aeval } a\ s)]$

$(\text{skip}; c) / s \to c / s$

$$\frac{c1 / s1 \to c2 / s2}{(c1\ ;\ c) / s1 \to (c2\ ;\ c) / s2}$$

$$\frac{\text{eval } s\ b = \text{true}}{(\text{if } b \text{ then } c1 \text{ else } c2) / s \to c1 / s}$$

$$\frac{\text{eval } s\ b = \text{false}}{(\text{if } b \text{ then } c1 \text{ else } c2) / s \to c2 / s}$$

$$\frac{\text{eval } s\ b = \text{false}}{(\text{while } b \text{ do } c \text{ end}) / s \to \text{skip} / s}$$

$$\frac{\text{eval } s\ b = \text{true}}{(\text{while } b \text{ do } c \text{ end}) / s \to c;\ \text{while } b \text{ do } c \text{ end} / s}$$

Equivalence with big-step semantics

IMP.v

A classic result: c / s ⇒ s' if and only if $c/s \rightarrow * \; \texttt{skip}/s'$

This proof is useful to build confidence in both semantics.

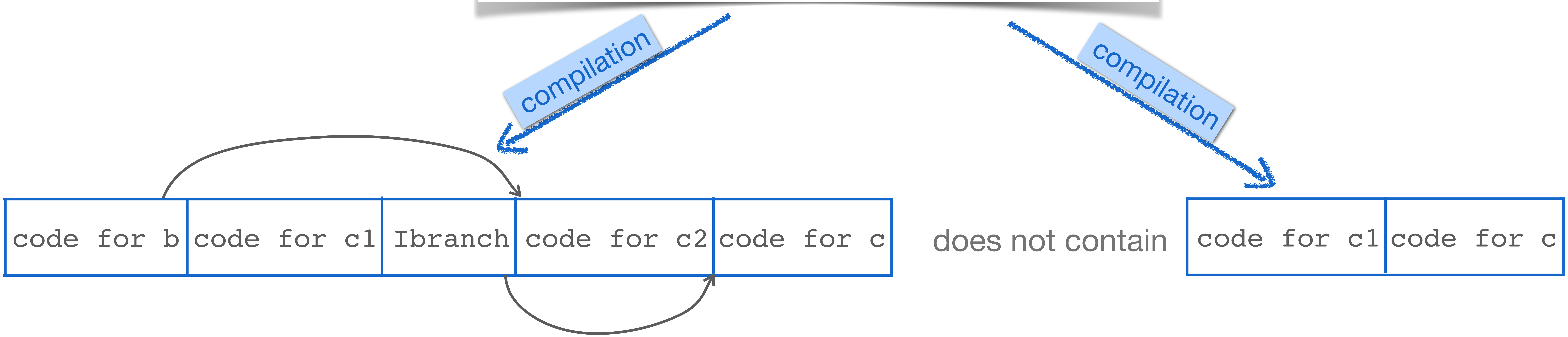
- From big-step to small-step : by induction on the ⇒ relation
- From small-step to big-step: intermediate lemma

If c1 / s1 → c2 / s2 and c2 / s2 ⇒ s' then c1 / s2 ⇒ s'

Spontaneous generation of commands

Some rules generate fresh commands that are not subterms of the source program.

(if b then c1 else c2); c / s → (c1; c) / s



Raises two issues when using simulation diagrams:

- impractical to reason on the execution relation
- difficult to define the measure

Small-step semantics with continuations

Instead of rewriting whole commands: c / s → c' / s'

rewrite pairs of (subcommand under focus, continuation): c / k / s → c' / k' / s'

Continuation

- remainder of command
- context in which it occurs (control stack)

  Kstop nothing remains to be done

  c • k execution of a sequence of two commands

  Kwhile b c k execution of a loop

Small-step semantics with continuations

c / k / s → c' / k' / s'

No generation of fresh commands: c' is always a subterm of c

(if b then c1 else c2) / k / s → c1 / k / s   when eval s b = true

New kinds of rules for dealing with continuations

(c1;c2) / k / s → c1 / c2 • k / s   Focus (on the left of a sequence)

skip / c • k / s → c / k / s   Resume (the remaining computations)

A small-step semantics for IMP

c / k / s → c' / k' / s'

x := a / k / s → skip / k / x ↦(aeval a s); s

(c1 ; c2) / k / s → c1 / c2 • k / s

$$\frac{\text{eval s b = true}}{\text{(if b then c1 else c2) / k / s} \rightarrow \text{c1 / k / s}}$$

$$\frac{\text{eval s b = false}}{\text{(if b then c1 else c2) / k / s} \rightarrow \text{c2 / k / s}}$$

$$\frac{\text{eval s b = false}}{\text{(while b do c end) / k / s} \rightarrow \text{skip / k / s}}$$

$$\frac{\text{eval s b = true}}{\text{(while b do c end) / k / s} \rightarrow \text{c / Kwhile b c k / s}}$$

skip / c • k / s → c / k / s

skip / Kwhile b c k / s → while b do c end / k / s

Program execution

Termination

Divergence

```
Definition kterminates (s: store) (c: com) (s': store) :=
  star step (c, Kstop, s) (SKIP, Kstop, s').

Definition kdiverges (s: store) (c: com) :=
  infseq step (c, Kstop, s).
```

Equivalence between small-step semantics

```
Theorem equiv_smallstep_terminates:
  ∀ s c s', terminates s c s' ↔ kterminates s c s'.

Theorem equiv_smallstep_diverges:
  ∀ s c, diverges s c ↔ kdiverges s c.
```

Full proof of compiler correctness
Simulation diagram

with $0 \leq$ `measure(c₂,k₂) < measure(c₁,k₁)`

Difficulties

- find the invariant ≈ between source and target states
- find the measure from source states to a natural number

Full proof of compiler correctness
The anti-stuttering measure

When do the source program stutter? When no VM instruction is executed.

(c1 ; c2) / k / s → c1 / c2 • k / s

skip / c • k / s → c / k / s

(if true then c1 else c2) / k / s → c1 / k / s

(while true do c end) / k / s → c / Kwhile b c k / s

`measure(c,k):` sum of the sizes of `c` and all the commands appearing in `k`

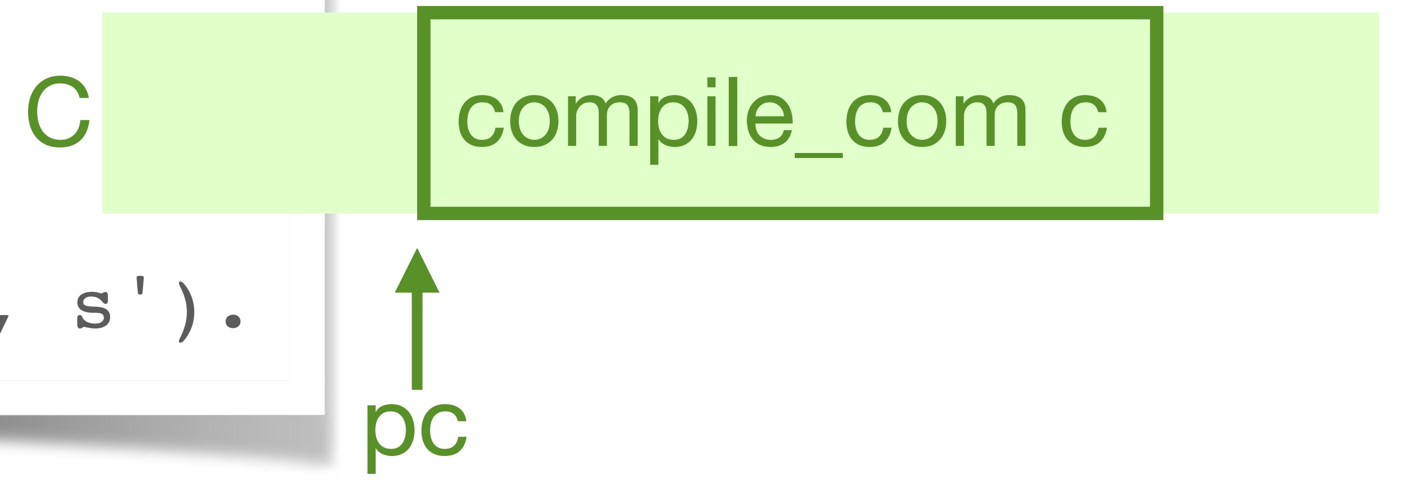
skip and := have size 1
The size of a sequence s1;s2 is the sum of the sizes of s1 and s2.

Trick: the size of Kwhile b c k is the size of k.

Full proof of compiler correctness
The simulation invariant

Remember this slide:

```
Lemma compile_com_correct_terminating:
  ∀ s c s', ceval s c s' →
  ∀ C pc stack,
  code_at C pc (compile_com c) →
  transitions C (pc, stack, s)
    (pc + codelen(compile_com c), stack, s').
```



C ⊢ **c / k / s ≈ (pc, stack, s')** is defined as:

- s = s'
- stack = [ ]
- `code_at C pc (compile_com c)` as in the previous proof
- C contains compiled code matching k at `pc + codelen(compile_com c)`

Compiler correctness: wrapping up

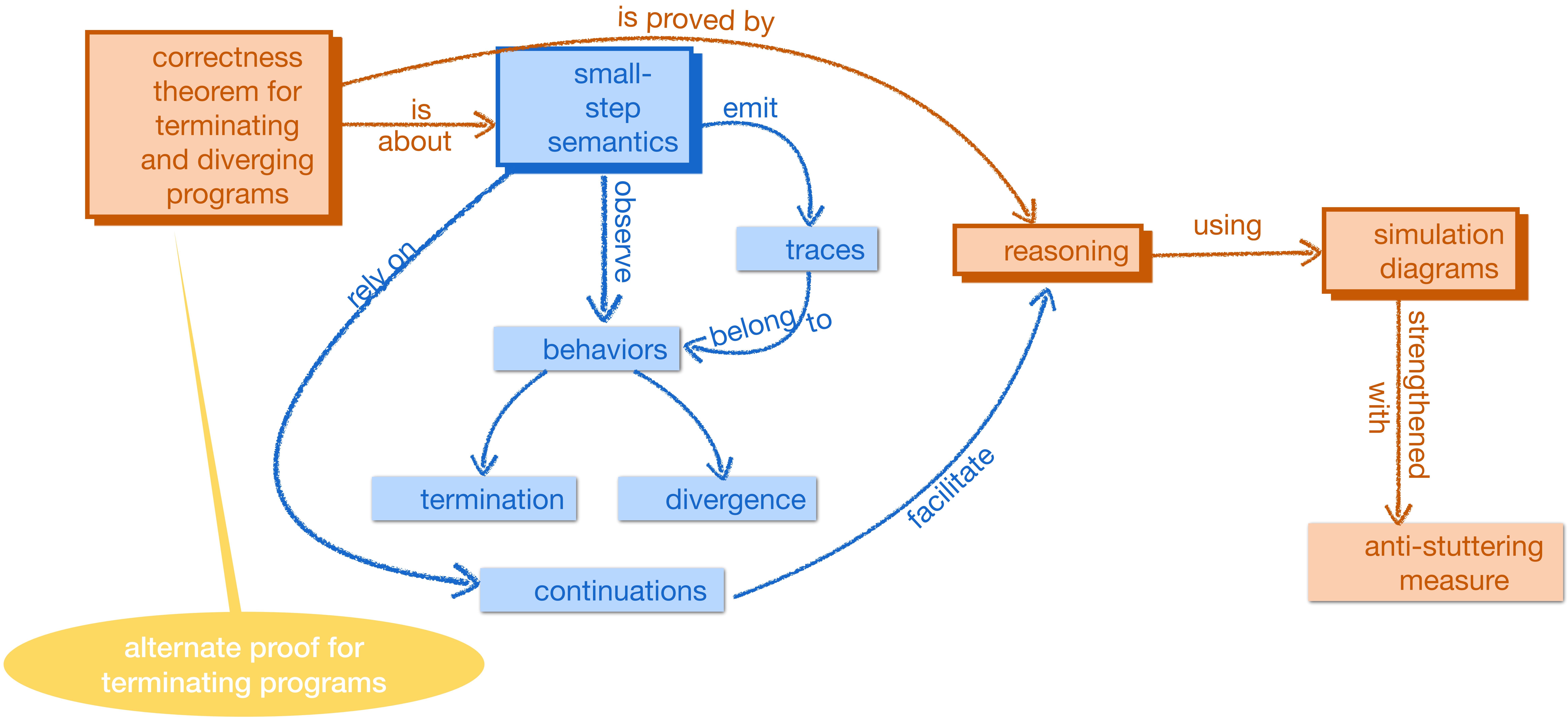
compil.v

```
Theorem compile_program_correct_terminating:
  ∀ s c s',
  cexec s c s' →
  machine_terminates (compile_program c) s s'.
```

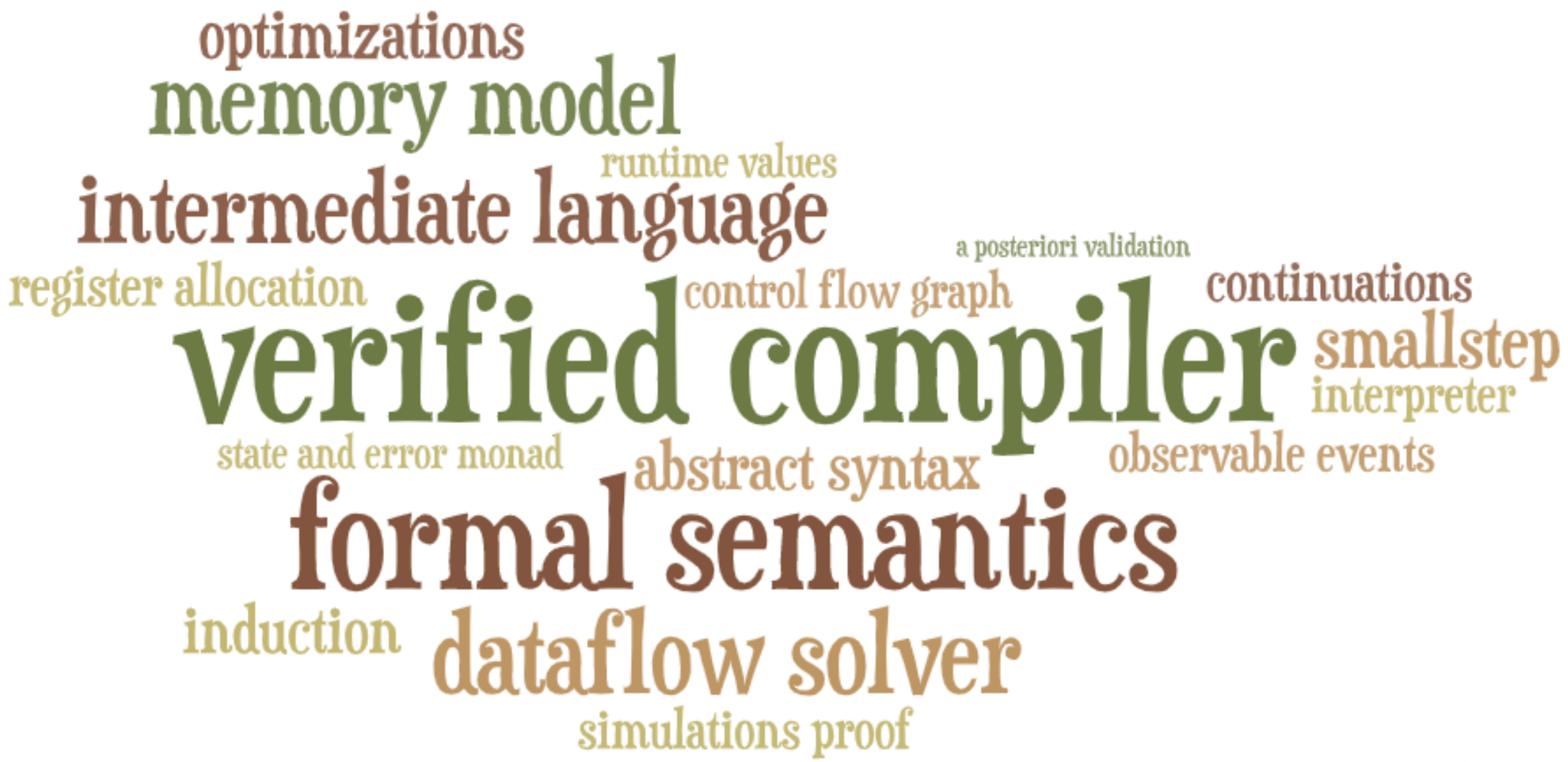
```
Theorem compile_program_correct_terminating_2:
  ∀ s c s',
  star step (c, Kstop, s) (SKIP, Kstop, s') →
  machine_terminates (compile_program c) s s'.
```

```
Theorem compile_program_correct_diverging:
  ∀ c s,
  infseq step (c, Kstop, s) →
  machine_diverges (compile_program c) s.
```

Part 5: summary



Part 6
How to turn CompCert from a prototype in a lab into a real-world compiler?



Observable behaviors

Behaviors.v and Events.v

```
program_behavior :=
   | Terminates t n
   | Diverges t
   | Reacts tinf
   | Goes_wrong t
```

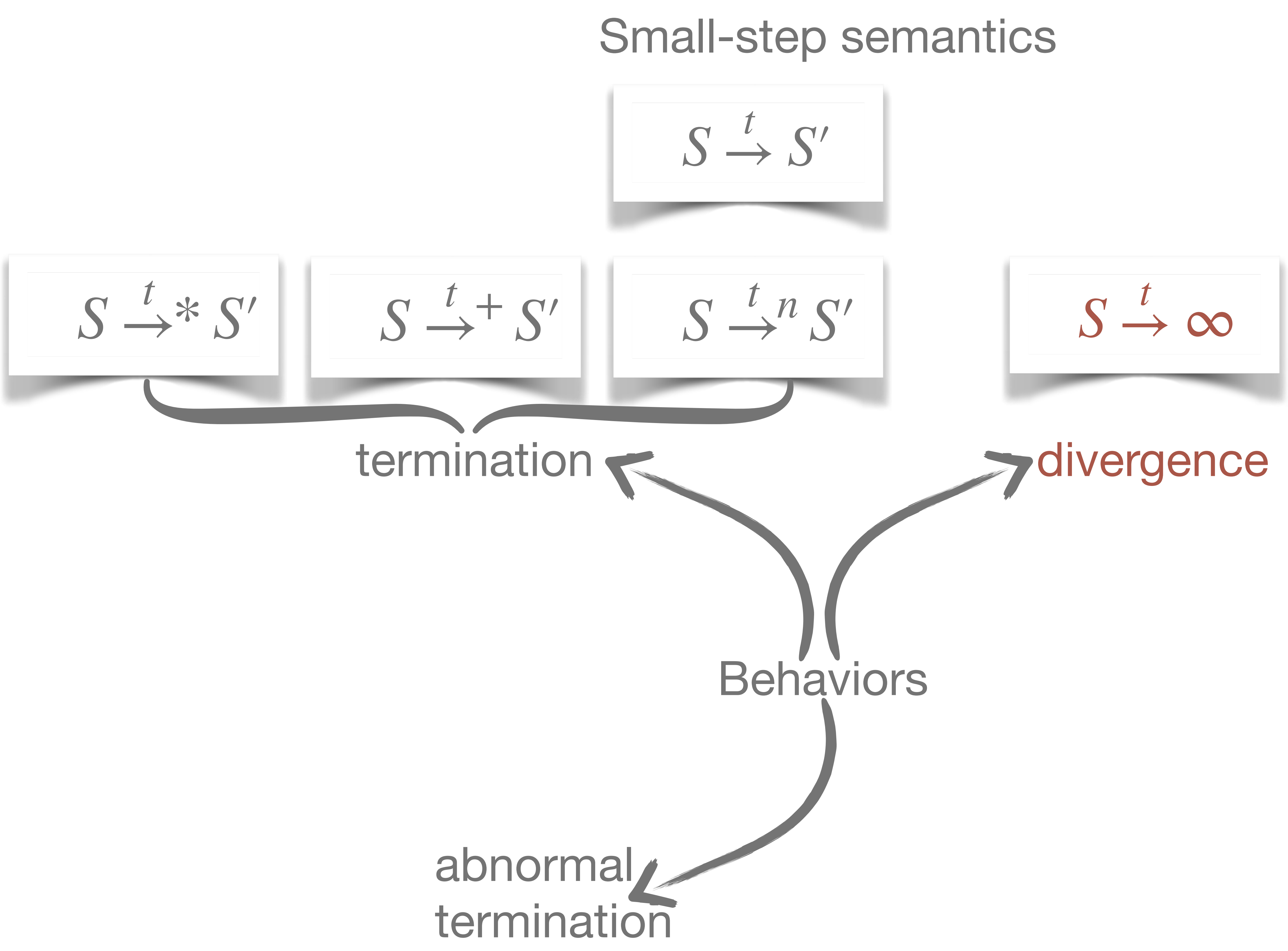
`t` = list of I/O events

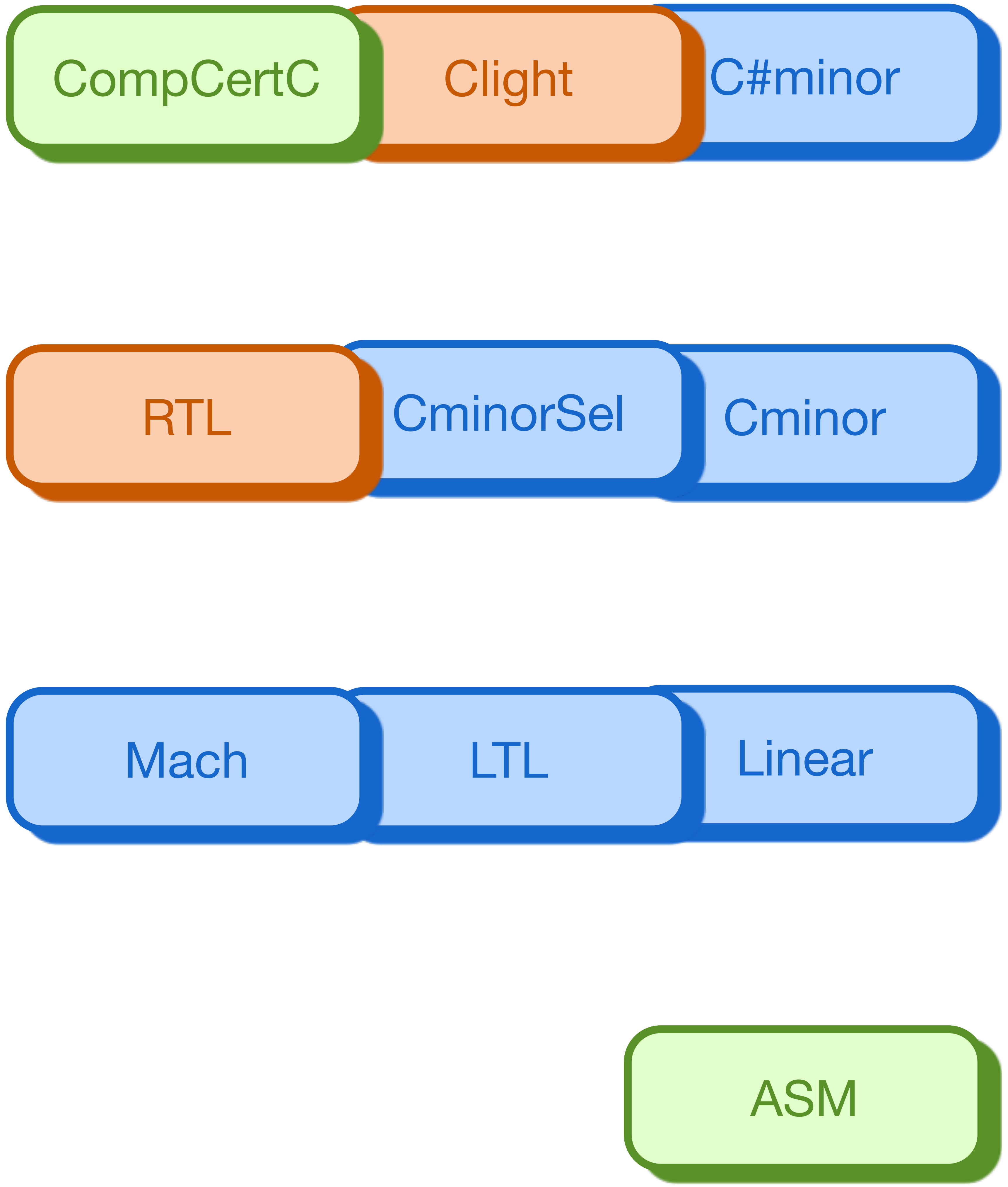
`tinf` = infinite list of I/O events

I/O event

- call to an external function (e.g. `printf`)
- memory accesses to global volatile variables (hardware devices)

CompCert compiler: 11 languages, 18 passes





General form of small-step semantics

Smallstep.v 

The semantics `L.sem` of a language `L` is defined by a step relation between semantic states, the initial and final states.

observed events

$G \vdash S \xrightarrow{t} S'$

`initial_state(`S`)`

`final_state(`S`,n)`

does not change during transitions

return value

G maps:

- each name of a function or global variable to a memory address
- each function pointer to a function definition

Semantic states S include a memory state, mapping addresses to values.

CompCert: main correctness theorem

Compiler.v

```
Theorem transf_c_program_correct:
  ∀ p tp,
  transf_c_program p = OK tp →
  backward_simulation (Csem.semantics p) (Asm.semantics tp).
```

The CompCert memory model

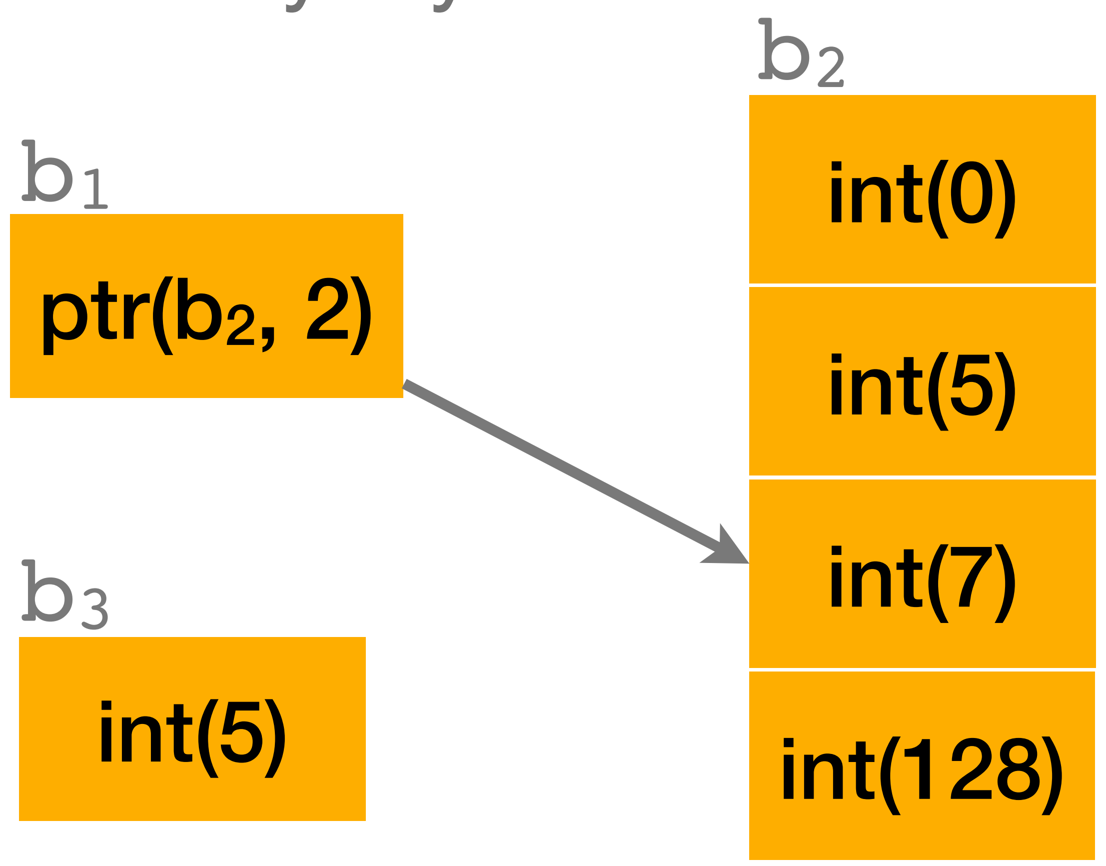
Memory.v

Shared by all the languages of the compiler

An abstract view of memory refined into a concrete memory layout

In the semantics:

```
m: mem
```



Memory: a collection of disjoint blocks
Values: machine integers, pointers, floating-point numbers

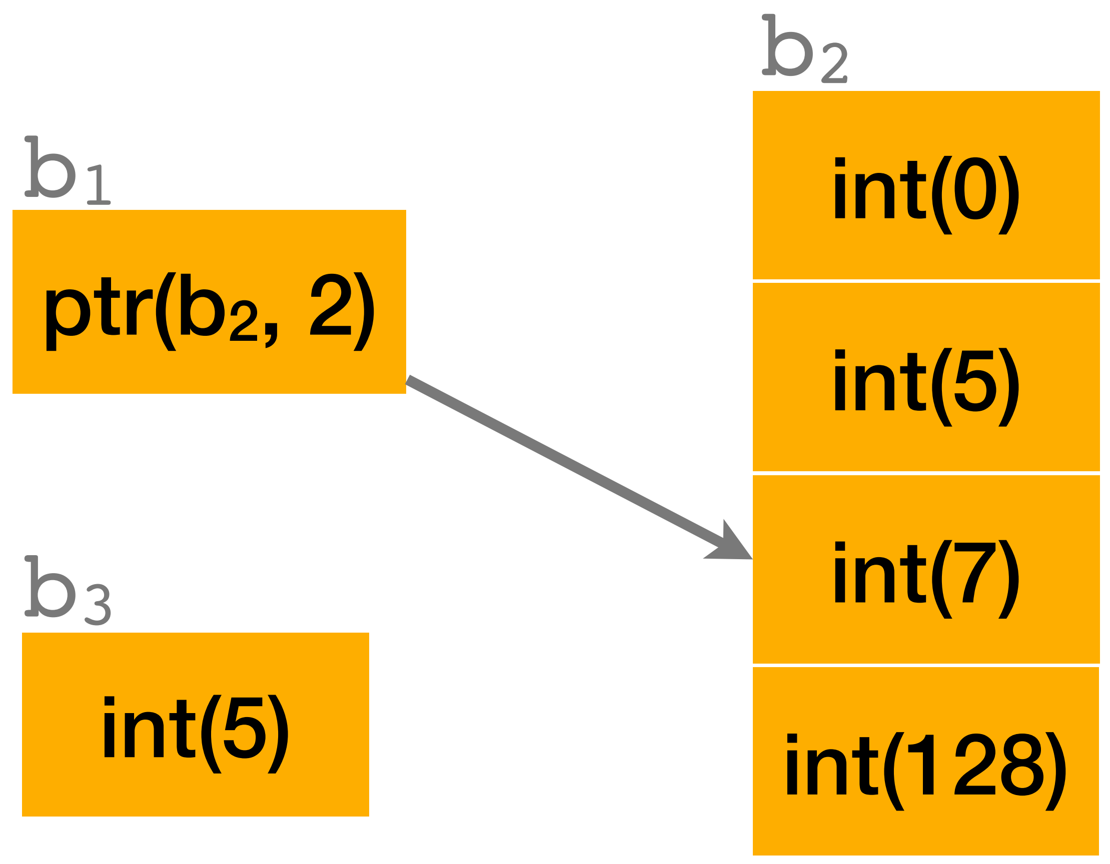
In C semantics, there are as many blocks as variables.

The number of blocks decreases during compilation.

The CompCert memory model

Memory operations (load, store, alloc, free) over values

load chunk m b ofs : option value
chunk: memory_chunk (ex. 16-bit unsigned int)



Memory safety preserved by CompCert (good variable properties)

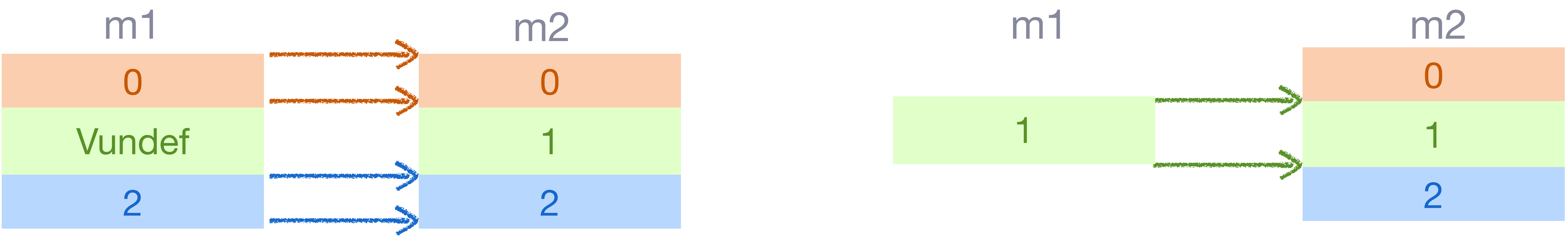
If $\texttt{alloc}(m, l, h) = \texttt{OK}(b, m') \land b' \neq b$, then $\texttt{load}(\tau, m', b', ofs) = \texttt{load}(\tau, m, b', ofs)$

If $\texttt{store}(\tau, m, b, ofs, v) = \texttt{OK}(m') \land \tau \sim \tau'$, then $\texttt{load}(\tau', m', b, ofs) = \texttt{convert}(v, \tau')$

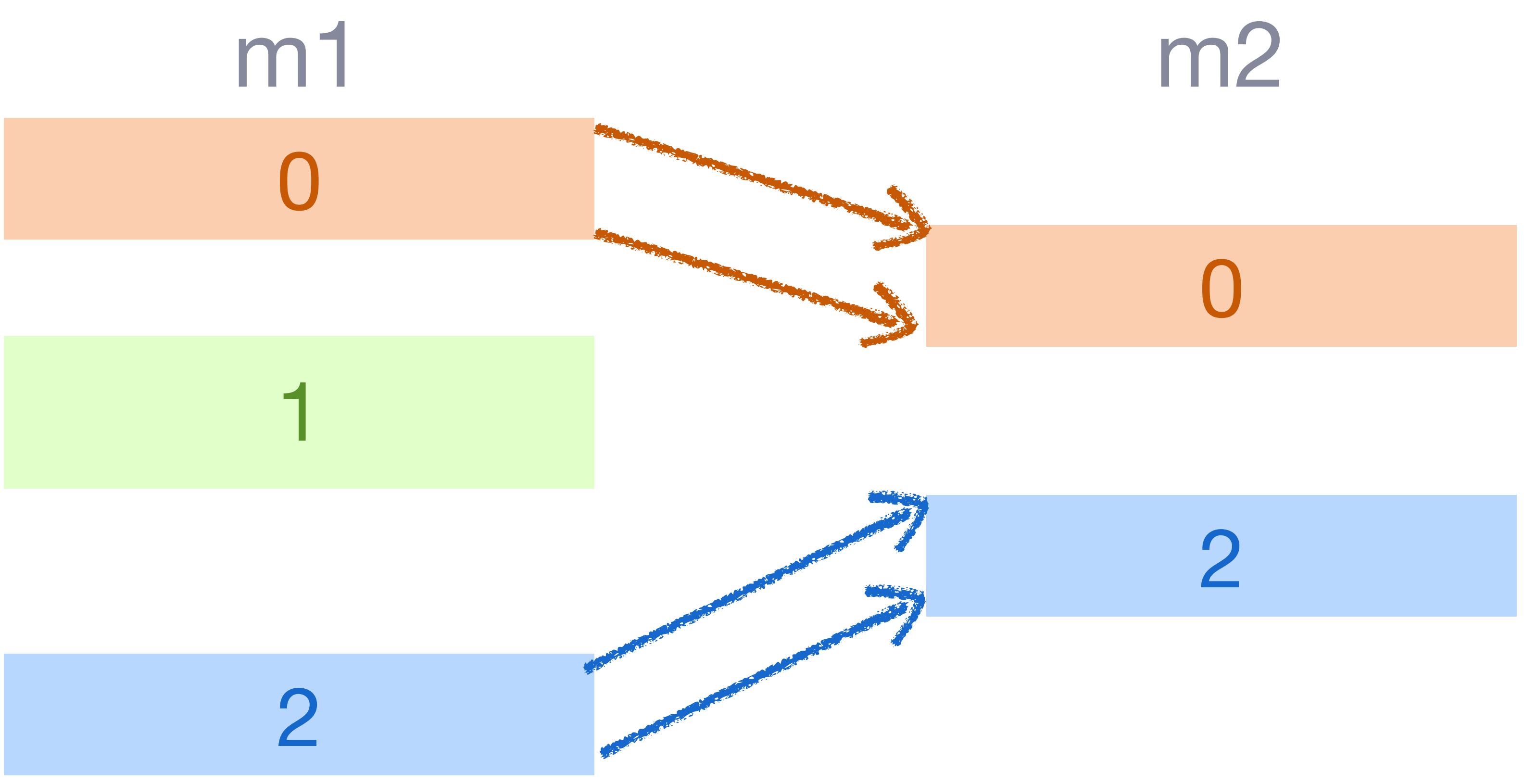
The CompCert memory model

Generic memory transformations

Memory **extensions**: m2 extends m1

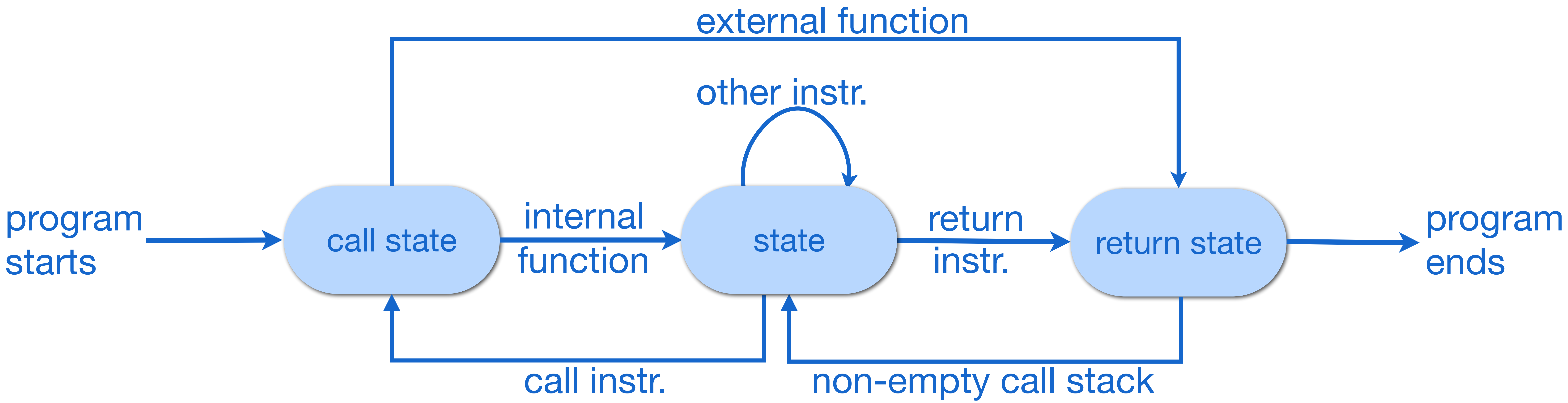


Memory **injections**: m1 is injected into m2



Memory operations are preserved by these transformations.

Semantic states



Exemple: Clight

```
state :=
    | State f stmt k env tempenv m
    | Callstate fdef args k m
    | Returnstate res k m
```

Exception: assembly languages, where a state is a pair of a memory and a mapping from processor registers to values

CompCert C source language

(see chapter 4 of the user's manual)

Expressions are annotated with their type

```
Eval(int(5), Tint(I32,Signed)): expr
```

Overloading and implicit conversions between types

Expressions have side-effects

- Assignments are expressions

Expressions implicitly classified into l-values and r-values

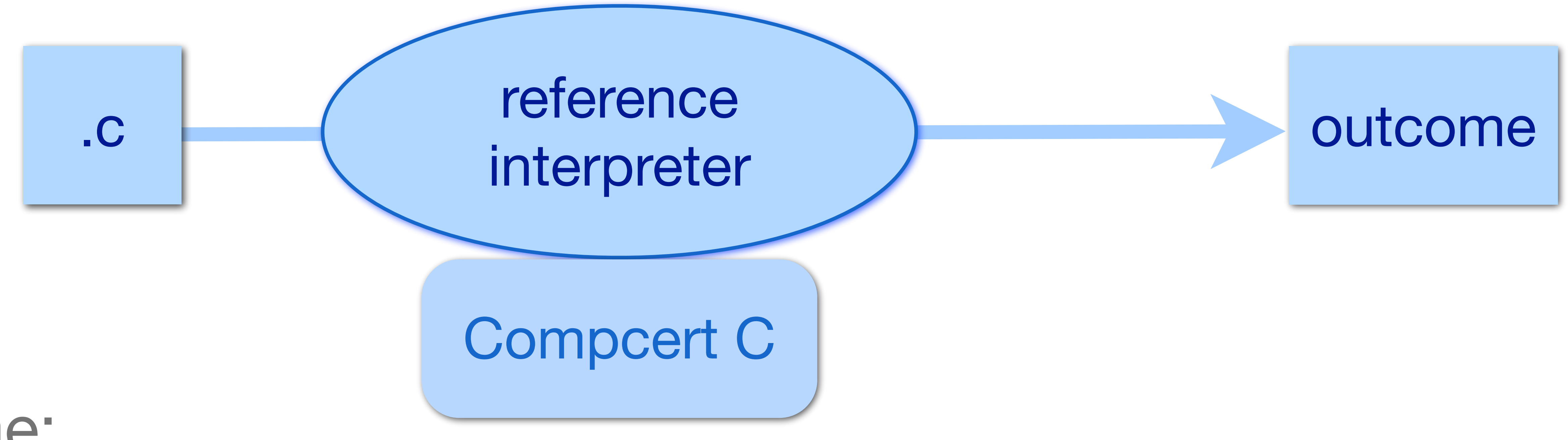
Non-deterministic evaluation of expressions (e.g., see this slide)

Commands

All C constructs: loops, `switch`, `goto`, `break`, `continue`, `return`

The CompCert C reference interpreter

Cexec.v



Outcome:

- normal termination or aborting on an undefined behavior
- observable effects (I/O events: `printf`, `volatile` memory accesses)

Faithful to the formal semantics of CompCert C; the interpreter displays all the behaviors according to the semantics

```
step: genv → state → trace → state → Prop
```

predicate

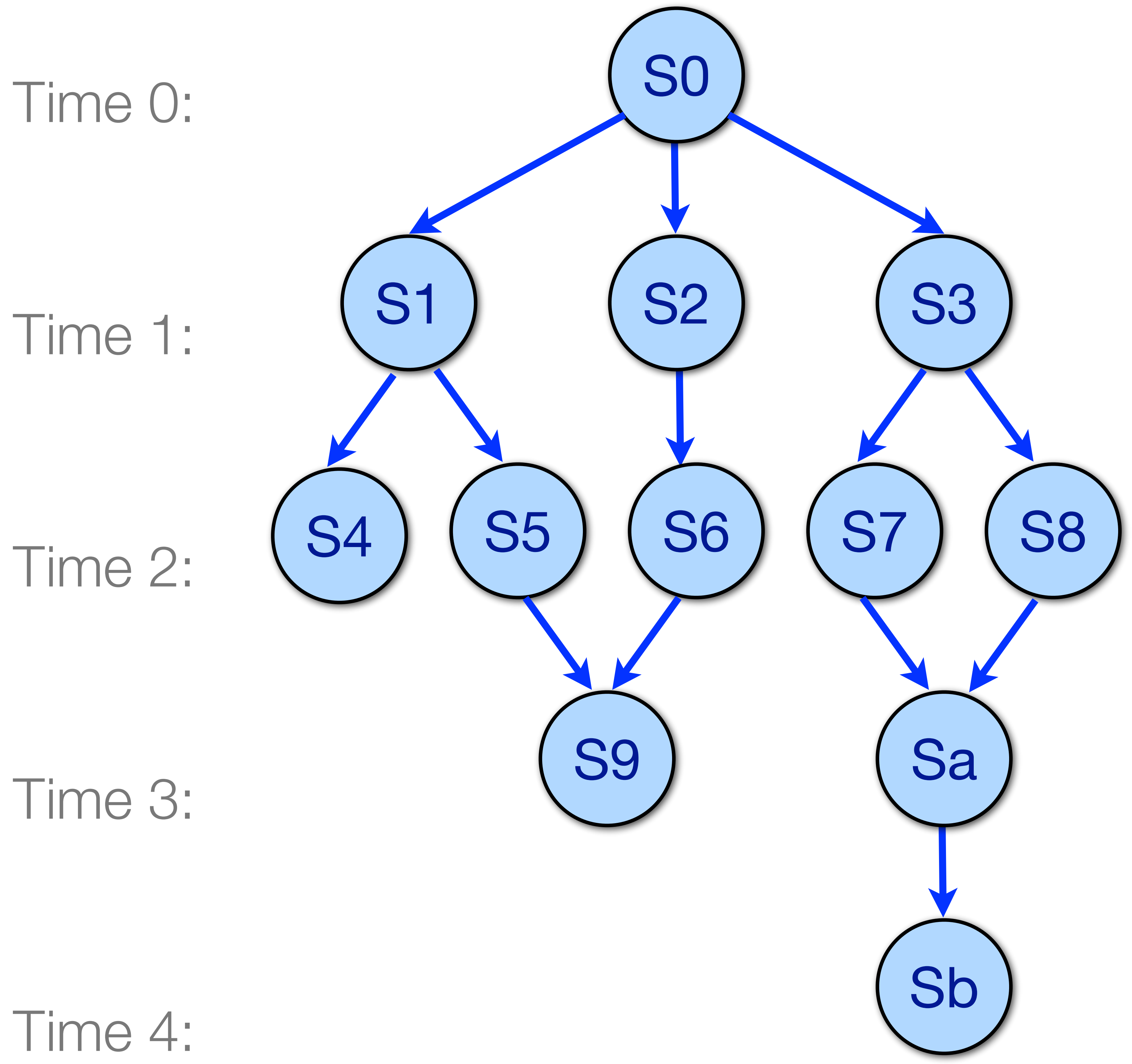
$$G \vdash S \xrightarrow{t} S'$$

```
do_step: world → genv → state → list (trace * state)
```

external world: uniquely determines the results of external calls

function

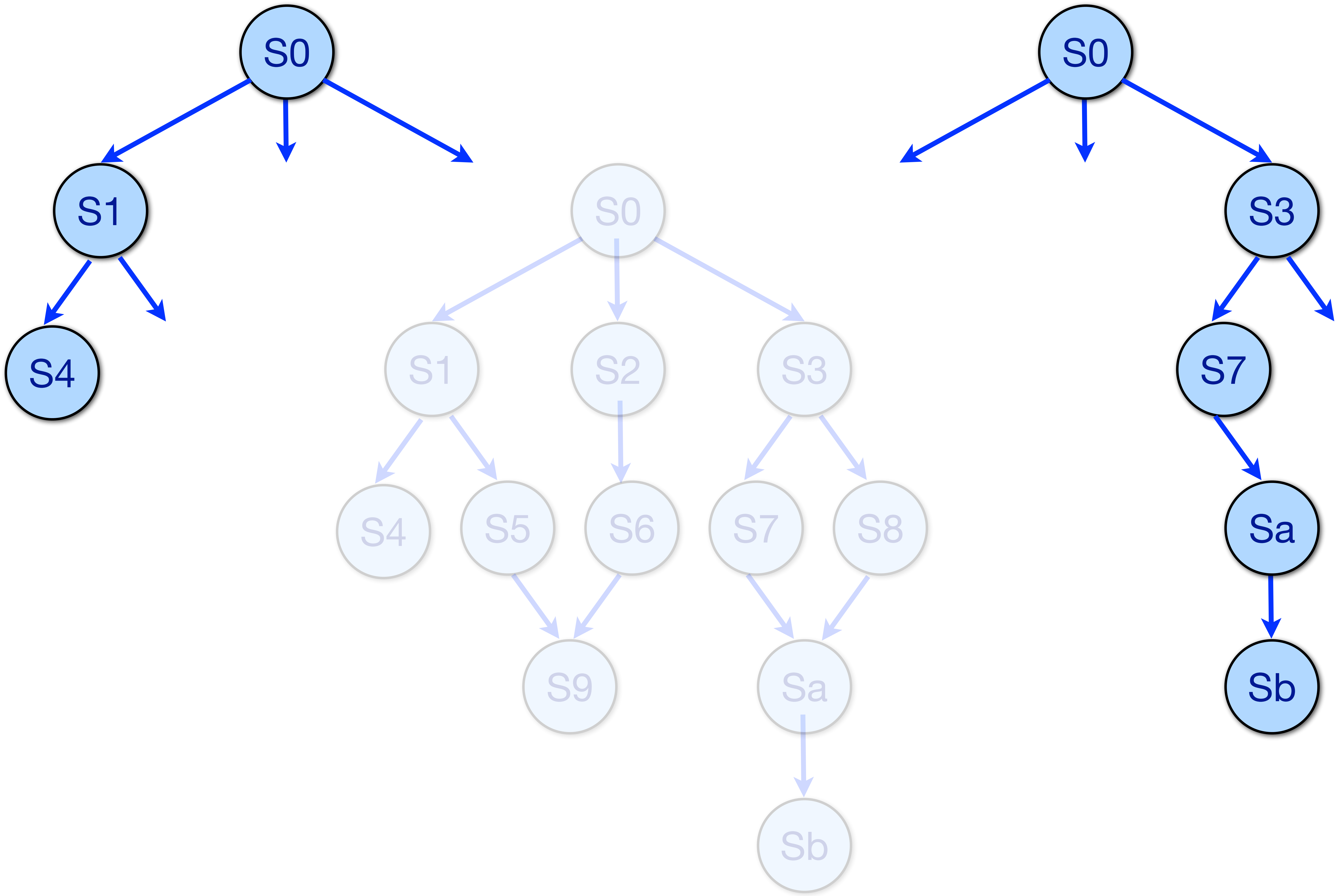
Using the reference interpreter: exhaustive exploration



Using the reference interpreter: randomized exploration

First choice

Randomized



Using the reference interpreter
A first example

```c
int f(int n) {
 int x = 1;
 for (int i = 1; i < n; i++)
    if (x < 9) x = x + 2;
    else if (x > 50) x = x + 1;
    else x = 2 * x;
    return x;  }

int main(void) {
    int res = f(12);
    printf("Result is %d \n",res);
    return 0; }
```

reference interpreter

```
Result is 76
Time 387: observable event: extcall printf(& __stringlit_1, 76)
Time 392: program terminated (exit code = 0)
```

number of execution steps

Using the reference interpreter
A first example with a detailed trace of execution

```
int f(int n) {
 int x = 1;
 for (int i = 1; i < n; i++)
    if (x < 9) x = x + 2;
    else if (x > 50) x = x + 1;
    else x = 2 * x;
    return x;  }

int main(void) {
    int res = f(12);
    printf("Result is %d \n",res);
    return 0; }
```

```
Time 0: calling main()
—[step_internal_function]—>
Time 1: in function main, statement
  res = f(12); printf(__stringlit_1, res); return 0;
—[step_seq]—>
…
Time 8: calling f(12)
—[step_internal_function]—>
Time 9: in function f, statement x = 1; for (…)…
—[step_seq]—>
Time 10: in function f, statement x = 1;
—[step_do_1]—>
Time 11: in function f, expression x = 1
—[red_var_local]—>
Time 12: in function f, expression <loc x> = 1
—[red_assign]—>
…
```

name of the semantic rule

Using the reference interpreter
A second example

```c
int main(void)
{  int x[2] = { 12, 34 };
  printf("x[2] = %d\n", x[2]);
  return 0;  }
```

reference interpreter

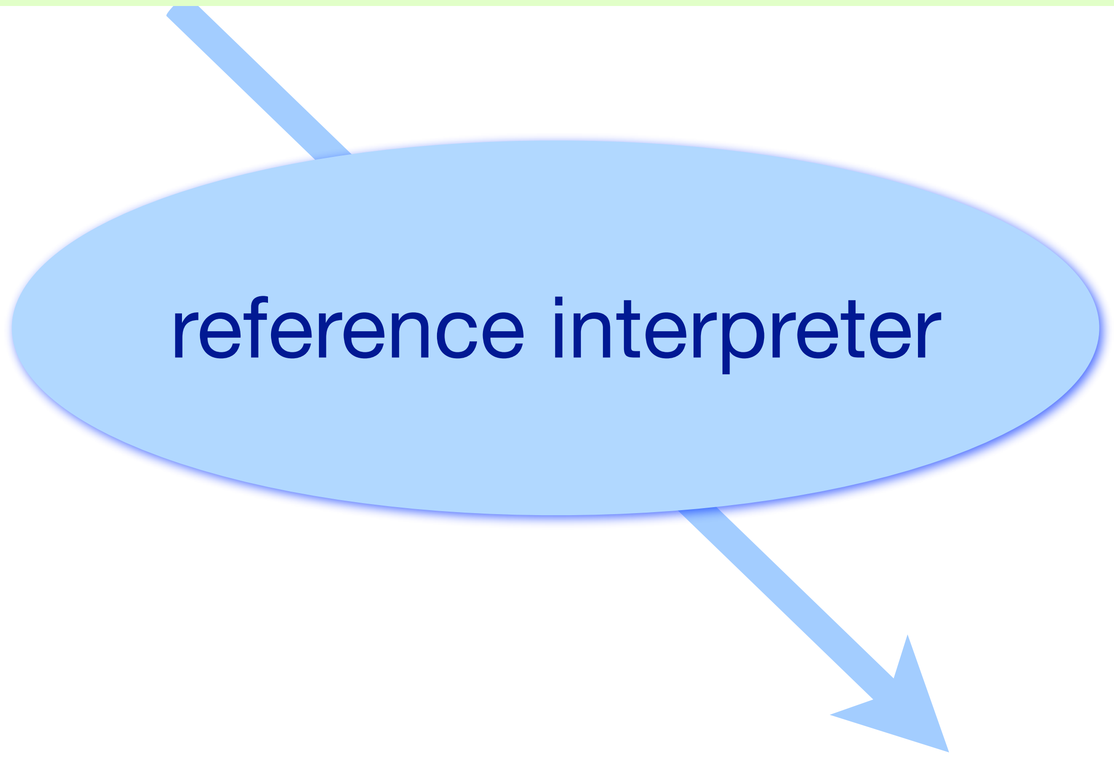
The interpreter stops on this undefined behavior.
This is not the case for the compiled code.

```
Stuck state: in function main, expression
  <printf>(<ptr __stringlit_1>, <loc x+8>)
Stuck subexpression: <loc x+8>
ERROR: Undefined behavior
```

Using the reference interpreter
A third example: randomized exploration

```c
int a() { printf("a "); return 1; }
int b() { printf("b "); return 2; }
int c() { printf("c "); return 3; }

int main () { printf("%d\n", a() + (b() + c())); return 0;  }
```




```
a Time 14: observable event: extcall printf(& __stringlit_1)
b Time 28: observable event: extcall printf(& __stringlit_2)
c Time 42: observable event: extcall printf(& __stringlit_3)
6
Time 50: observable event: extcall printf(& __stringlit_4, 6)
Time 55: program terminated (exit code = 0)
```

Clight language

Clight.v

Expressions are annotated with their type

```
Econst_int(int(5), Tint(I32,Signed)): expr
```

No overloading and explicit conversions between types and arithmetic operators

Expressions are pure

Temporary variables do not reside in memory

~~`&v`~~

Commands

Assignments are commands

Single syntax for loops, continue command

```
Sloop s1 s2
```

- C loops are derived forms

RTL language

RTL.v (a.k.a. 3-address code)

Each function is represented by its CFG

Elementary instructions only

Unlimited supply of pseudo-registers

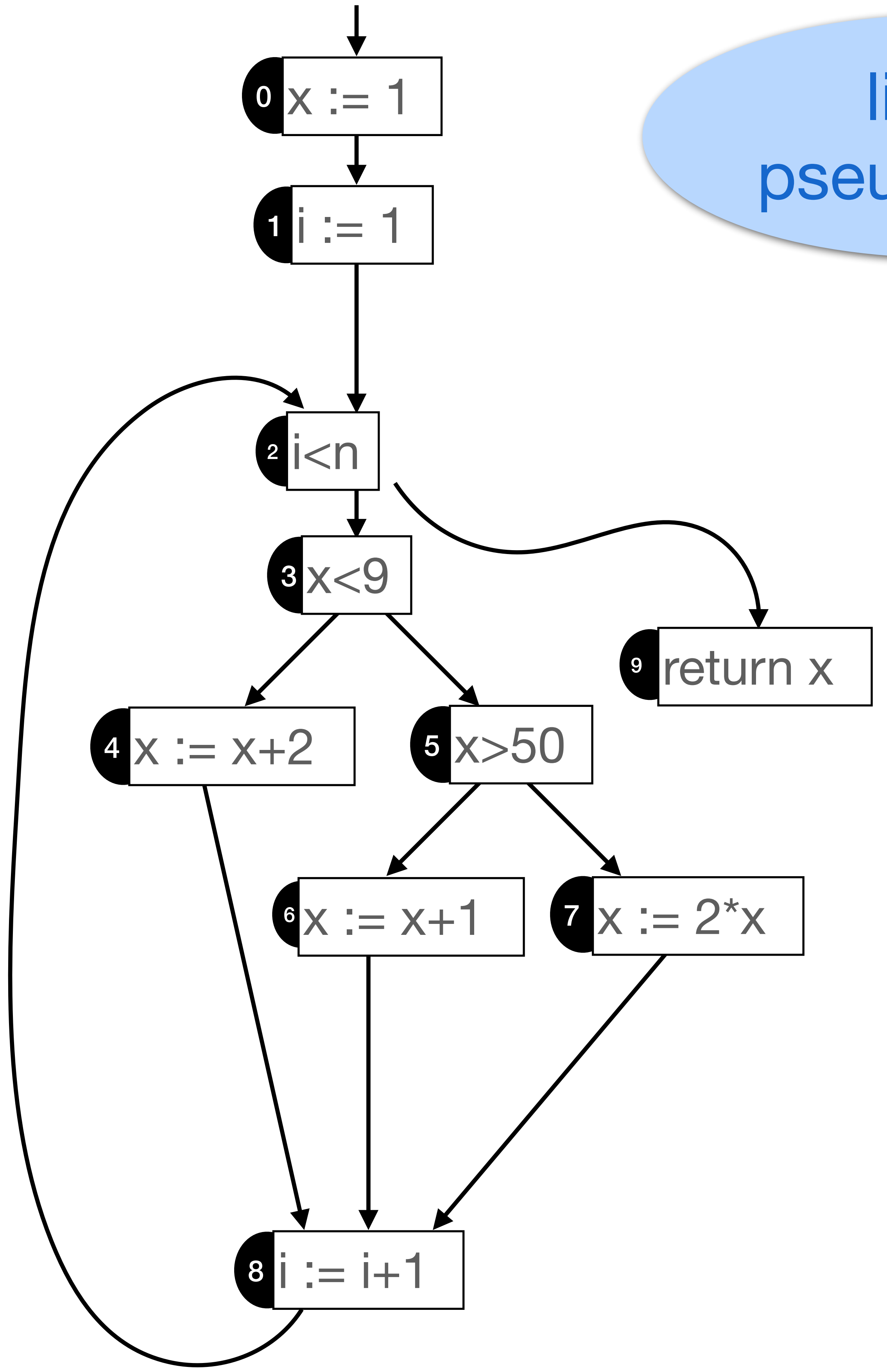
```
Iop(int(5), args, dest, succ): instruction
```

list of pseudo-regs

register to store the result

successor node

```
int f(int n) {
 int x = 1;
 for (int i = 1; i < n; i++)
   if (x < 9) x = x + 2;
   else if (x > 50) x = x + 1;
   else x = 2 * x;
 return x;  }
```



Assembly languages

A lot of lost information, including expressions, control flow, types, variable identifiers

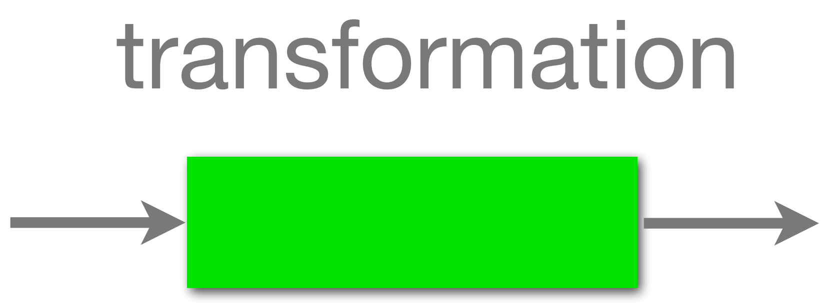
```
exec_instr (f:function) (i: instruction) (rs:regset) (m:mem): outcome
```

Next(rs’,m’) or Stuck

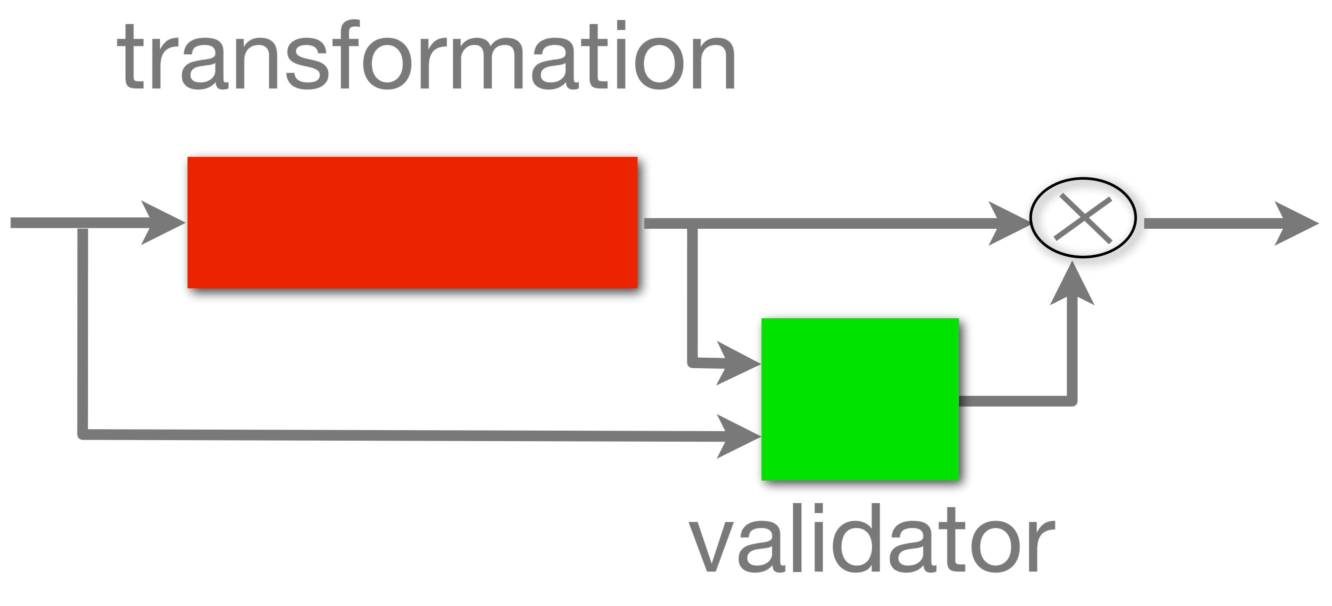
CompCert compiler: 10 languages, 18 passes

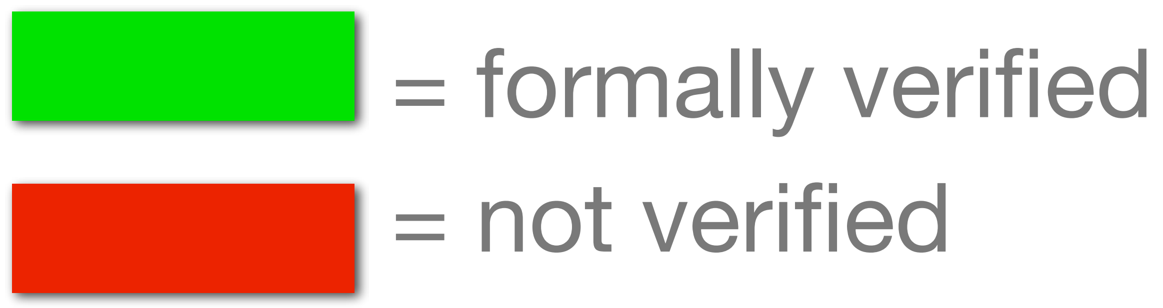
Verification patterns

Verified transformation



Verified translation validation

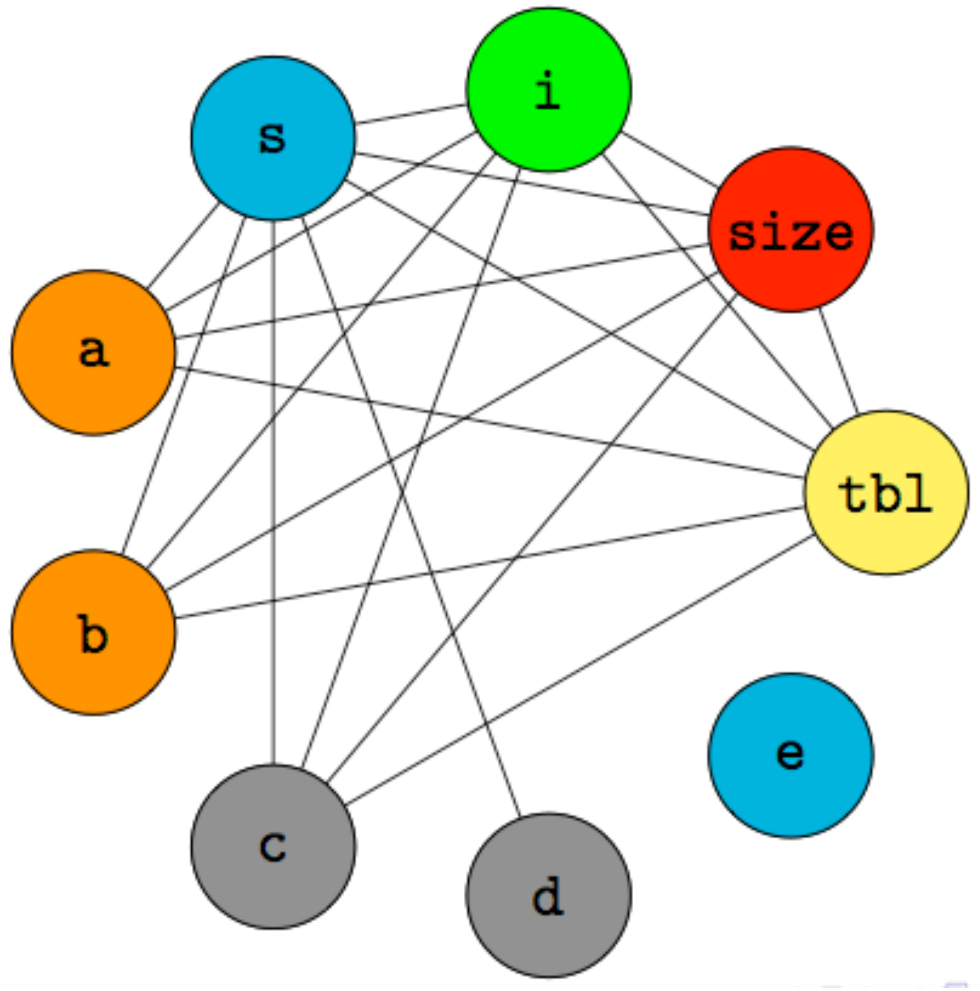




Verified validator

- Less to prove (if validator simpler than transformation)
- Validator reusable for several variants of an optimization
- Can be efficient (cheap enough to be invoked on every compiler run)

Example: register allocation with advanced spilling and splitting



Part 6: summary

Proving a compiler pass mainly amounts to proving a simulation diagram

Many reusable libraries:

- simulations, memory model, C semantics, Clight and RTL languages
- machine integers, dataflow solver

Some useful compilation options

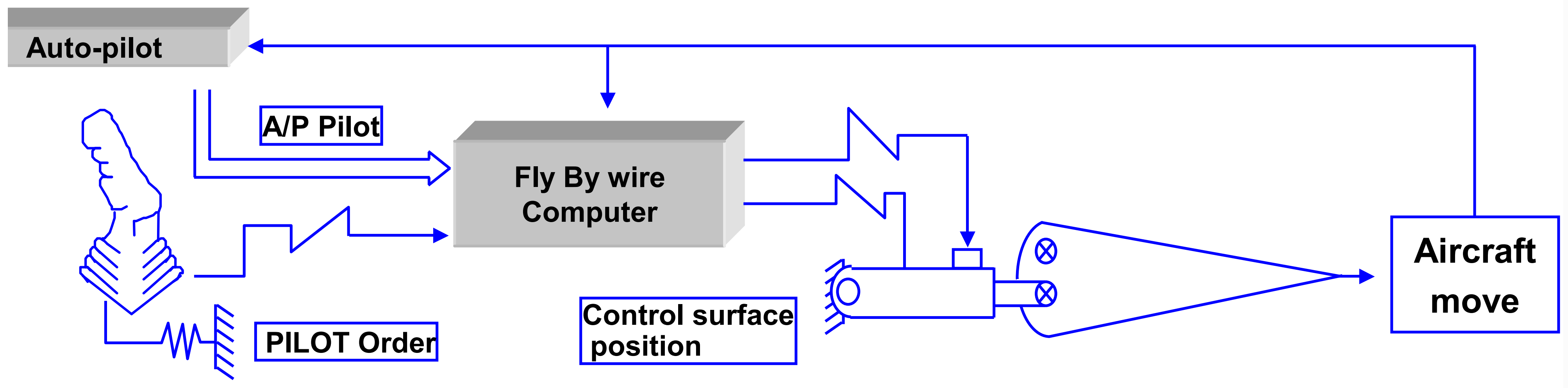
- using the CompCert C interpreter: `-interp` (`-trace`, `-all`, `-random`)
- tracing options: `-dc`, `-dclight`, `-drtl, …`
- show the time spent in compiler passes: `-timing`

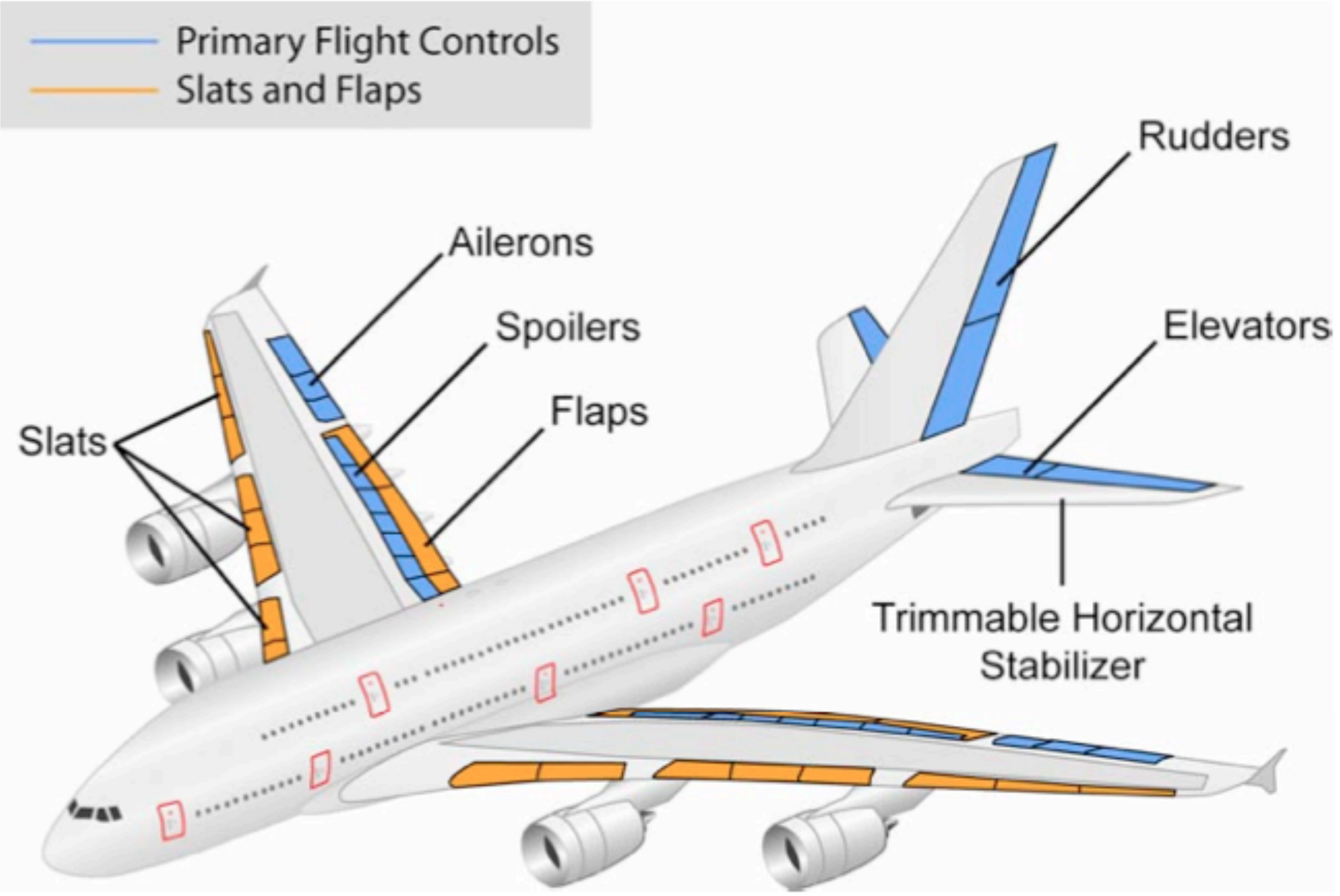
Part 7: Compiling critical embedded software with CompCert

Fly-by-wire software

Execute pilot's commands

Flight assistance: keep aircraft within safe flight envelope





Mostly control-command code (Scade) +
a minimalistic OS (C)

100k - 1M LOC code, but mostly generated from block diagrams (Simulink, Scade)

Fly-by-wire software

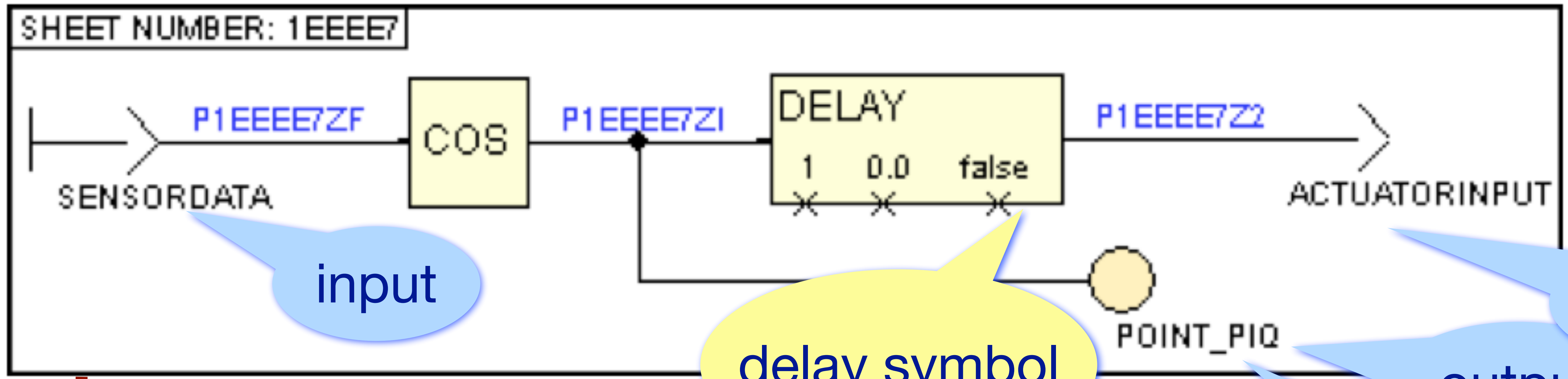
The qualification process (DO-178)

Rigorous validation: review (qualitative), analysis (quantitative), testing (huge amounts)

Conducted at multiple levels, from design to final product

Meticulous development process; extensive documentation

From block diagrams to assembly



input

delay symbol

output

output

observation point

code generator

compiler

delay macro

variable stored in RAM

```
/*Sheet Number: 1EEEE7*/

#include "other_includes.h"
#include "delay.mac"
#include "cos.mac"
#include "piq_x.mac"

m_START_NUM_INPUT_ZONE
m_INPUT(SENSORDATA,NUM)
m_END_NUM_INPUT_ZONE

m_START_NUM_OUTPUT_ZONE
m_OUTPUT(ACTUATORINPUT,NUM)
m_END_NUM_OUTPUT_ZONE

m_START_NUM_OBS_ZONE
m_OBS_NUM(POINT_PIQ)
m_END_NUM_OBS_ZONE

m_START_CST_INPUT_ZONE
m_CST(c_DELAY_P3_2_N1EEEE7,ENT, 1)
m_END_CST_INPUT_ZONE

m_START_STATIC_ARRAY_ZONE
m_STATIC_ARRAY(t_DELAY_2_N1EEEE7, 1,NUM)
m_END_STATIC_ARRAY_ZONE

m_START_SHEET(1E,EE,E7)

m_CONNECTION(P1EEEE7Z2_N1EEEE7,NUM)
m_CONNECTION(loc_c_DELAY_P3_2_N1EEEE7,ENT)
m_CONNECTION(P1EEEE7ZF_N1EEEE7,NUM)
m_CONNECTION(P1EEEE7ZI_N1EEEE7,NUM)

m_VtS(SENSORDATA,P1EEEE7ZF_N1EEEE7)
m_CtS(c_DELAY_P3_2_N1EEEE7,loc_c_DELAY_P3_2_N1EEEE7)

COS(0_N1EEEE7,P1EEEE7ZF_N1EEEE7,P1EEEE7ZI_N1EEEE7)
PIQ_X(1_N1EEEE7,P1EEEE7ZI_N1EEEE7,POINT_PIQ)
DELAY(2_N1EEEE7,P1EEEE7ZI_N1EEEE7,\
loc_c_DELAY_P3_2_N1EEEE7,P1EEEE7Z2_N1EEEE7,t_DELAY_2_N1EEEE7)

m_StV(P1EEEE7Z2_N1EEEE7,ACTUATORINPUT)
m_END_SHEET
```

delay macro

delay symbol

```
    Nm_1EEEE7:
    ; annotation: Symbol DELAY number 2_N1EEEE7 ,\
 inputs: f3, r31 and one static
100 addis      r12, 0, (_DELAY_2_N1EEEE7_R2)@ha
104 lwz        r4, (_DELAY_2_N1EEEE7_R2)@l(r12)
    ; annotation: Variable to search: loc_c_DELAY_P3_2_N1EEEE7
    ; annotation: DELAY; is entered with r4 = from 0 to
108 mr         r7, r4
10c addis      r12, 0, (t_DELAY_2_N1EEEE7)@ha
110 addi       r8, r12, (t_DELAY_2_N1EEEE7)@l
114 rlwinm     r10, r7, 3, 0, 28 ; 0xfffffff8
118 add        r10, r8, r10
11c lfd        f2, 0(r10)
    ; annotation: Variable to search: loc_c_DELAY_P3_2_N1EEEE7
    ; annotation: DELAY; is entered with r4 = from 0 to
120 mr         r8, r4
124 addis      r12, 0, (t_DELAY_2_N1EEEE7)@ha
128 addi       r6, r12, (t_DELAY_2_N1EEEE7)@l
12c rlwinm     r5, r8, 3, 0, 28 ; 0xfffffff8
130 add        r9, r6, r5
134 stfd       f3, 0(r9)
138 addi       r4, r4, 1
13c cmpw       cr0, r4, r31
140 bt         0, .L101
144 addi       r4, 0, 0
    .L101:
148 addis      r12, 0, (_DELAY_2_N1EEEE7_R2)@ha
14c stw r4, (_DELAY_2_N1EEEE7_R2)@l(r12)
    ; annotation: End of DELAY number 2_N1EEEE7 , output: f2
```

delay symbol

Program annotations

A mechanism to attach annotations to program points

- Mark specific program points
- Provide information about the location of C variables
- Ensure that some variables are preserved (e.g. x must be kept in a register)

Annotations are preserved during compilation.

- Each annotation generates an observable event
- The correctness theorem ensures preservation of the sequencing of 1) symbols, and 2) of accesses to hardware devices (volatile variables)

Conformance to the qualification process: CompCert gives traceability guarantees

```
_annot("Begin of a loop");
...
x = 1;
_annot("Here x is at %1",x);
...
_annot("End of a loop");
```

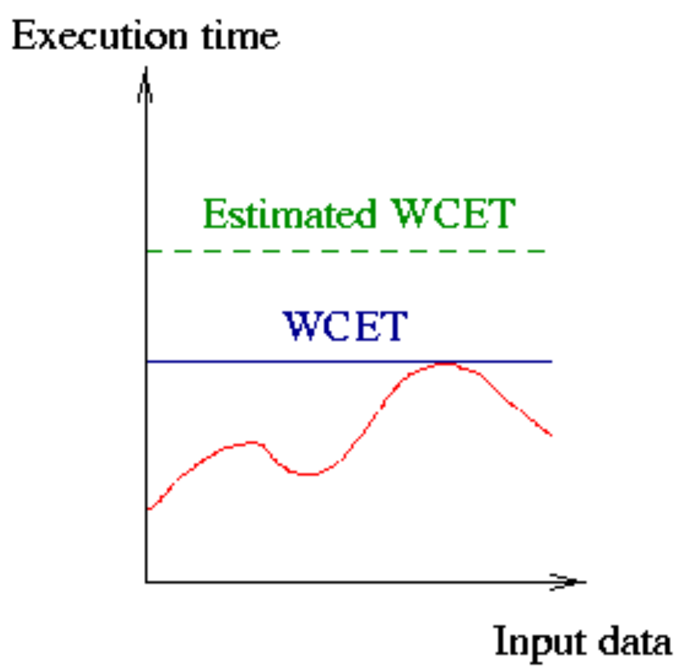
compiler

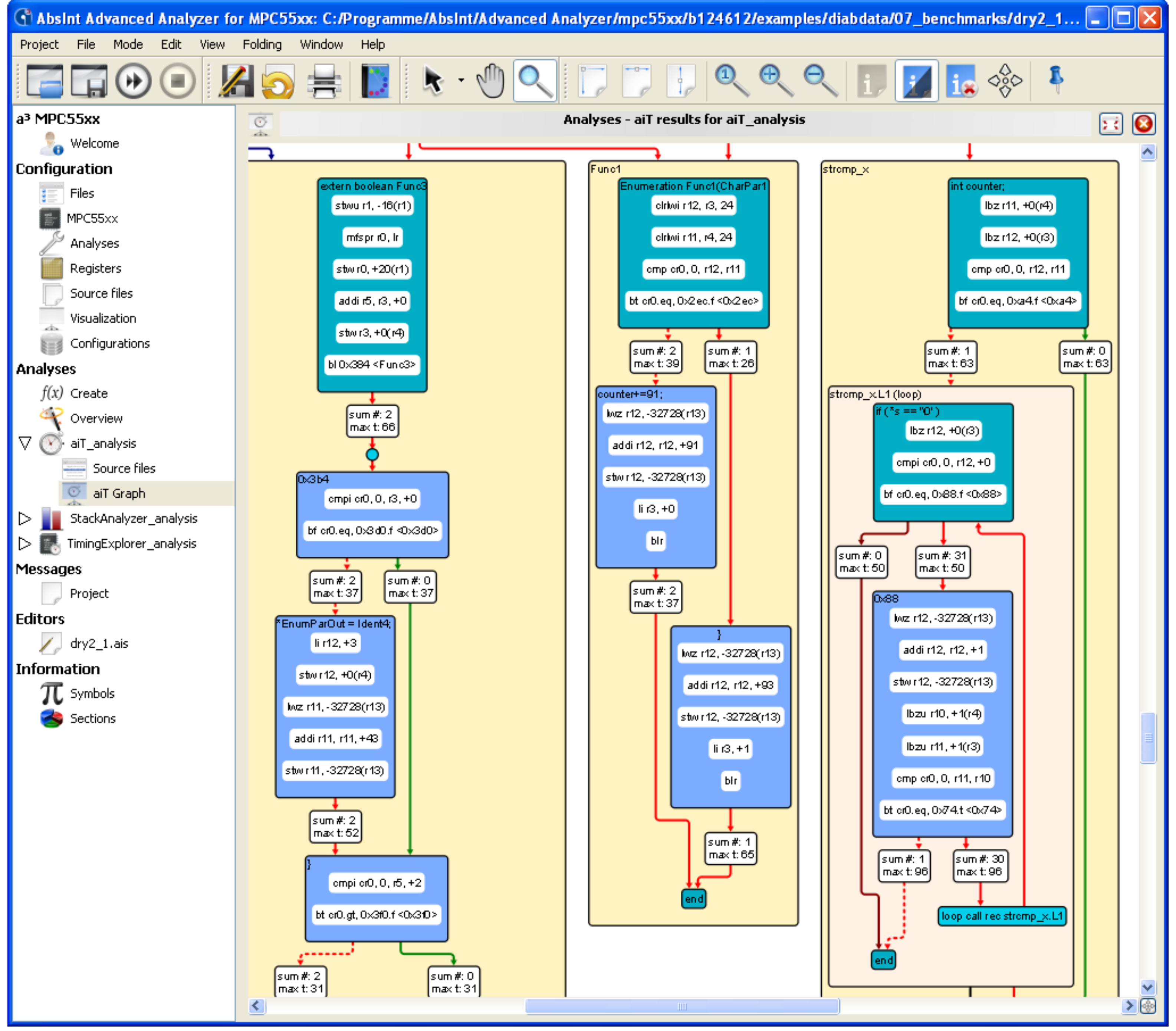
```
; annotation: Begin of a loop
...
addi r3, 0, 1
; annotation: Here x is at r3
...
; annotation: End of a loop
```

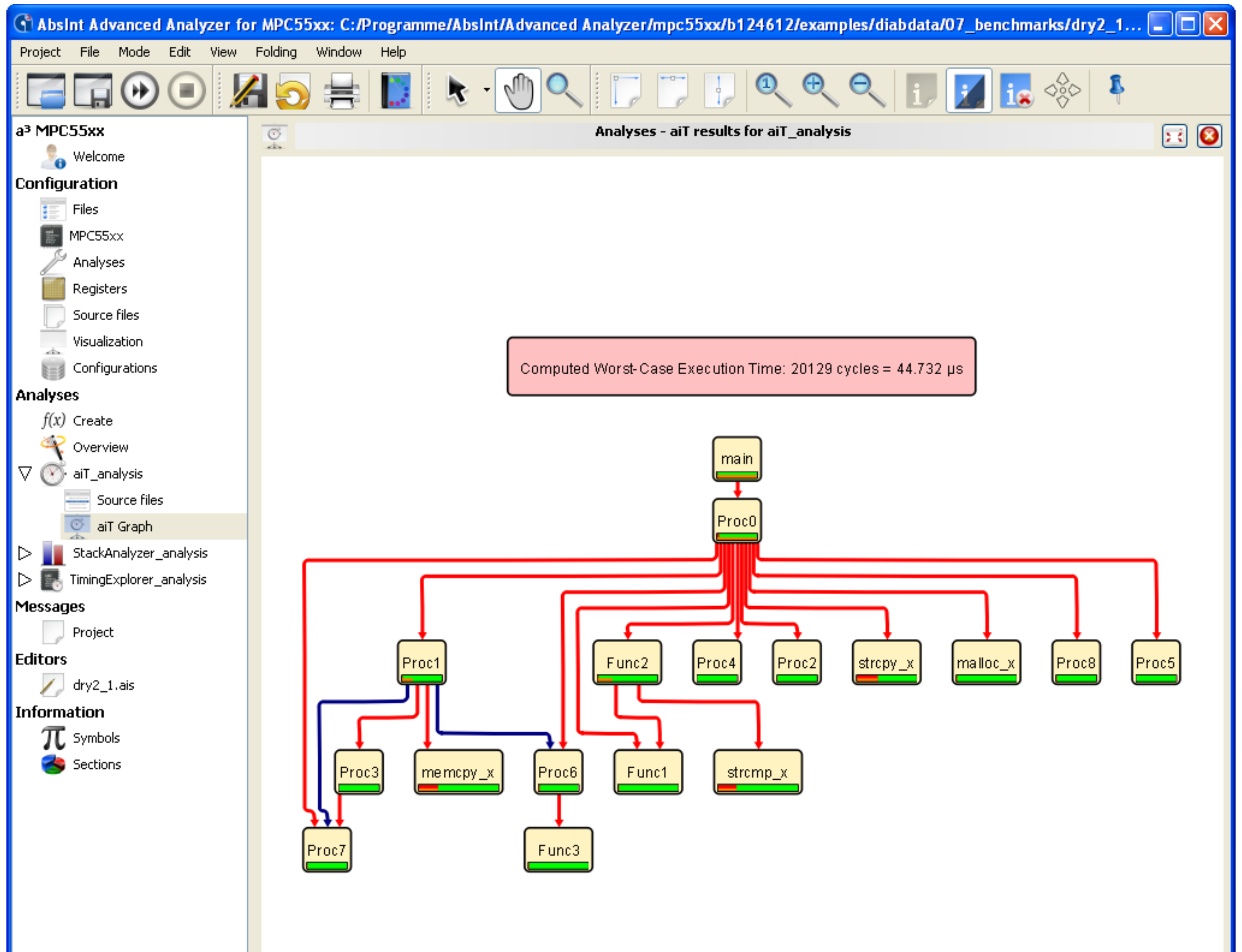
How good is the compiled code ?

Trade-off between
• traceability guarantees
• and efficiency of the generated code

Low-level verifications
• reviews of the assembly
• computation of a WCET estimation



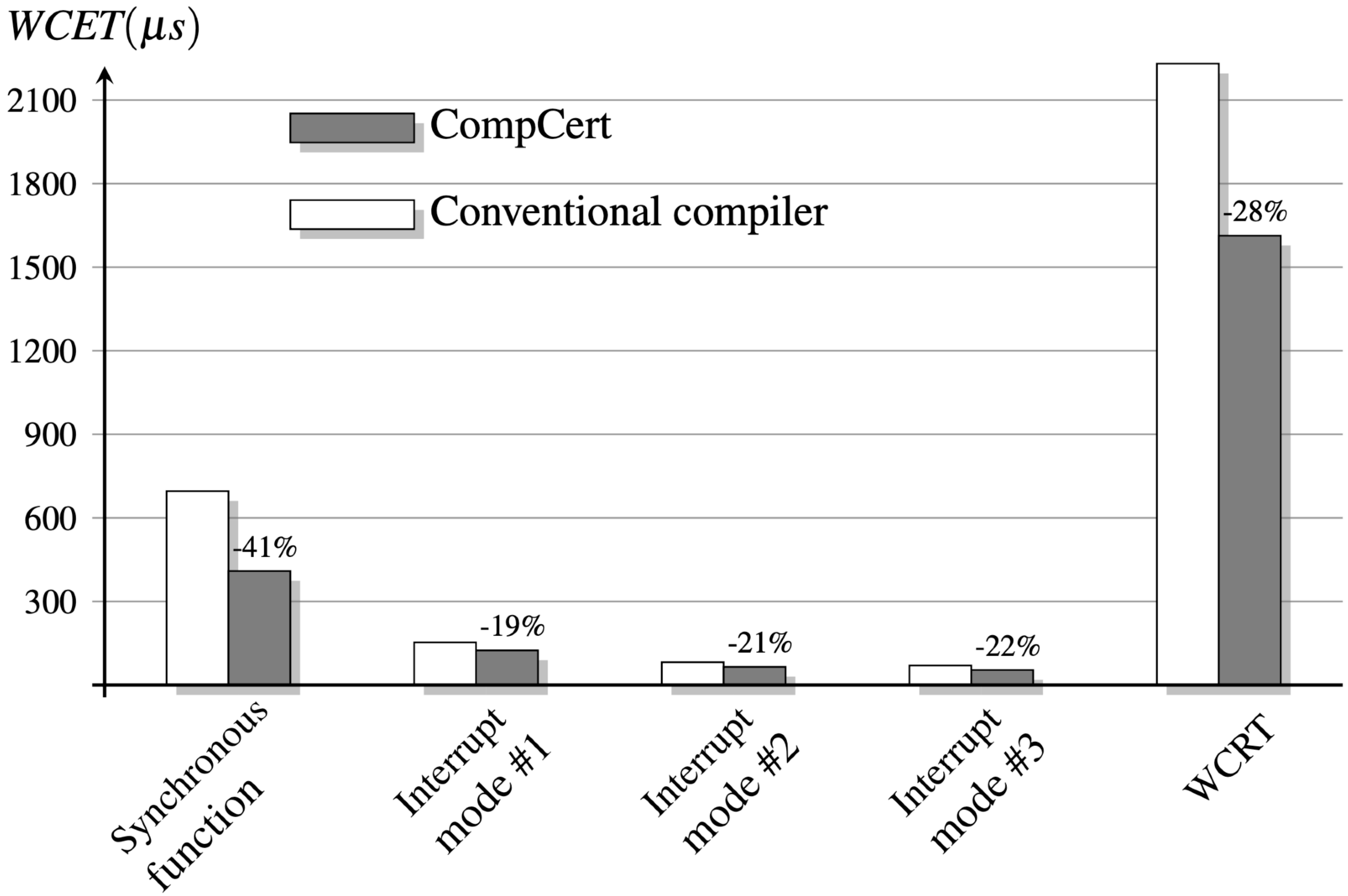


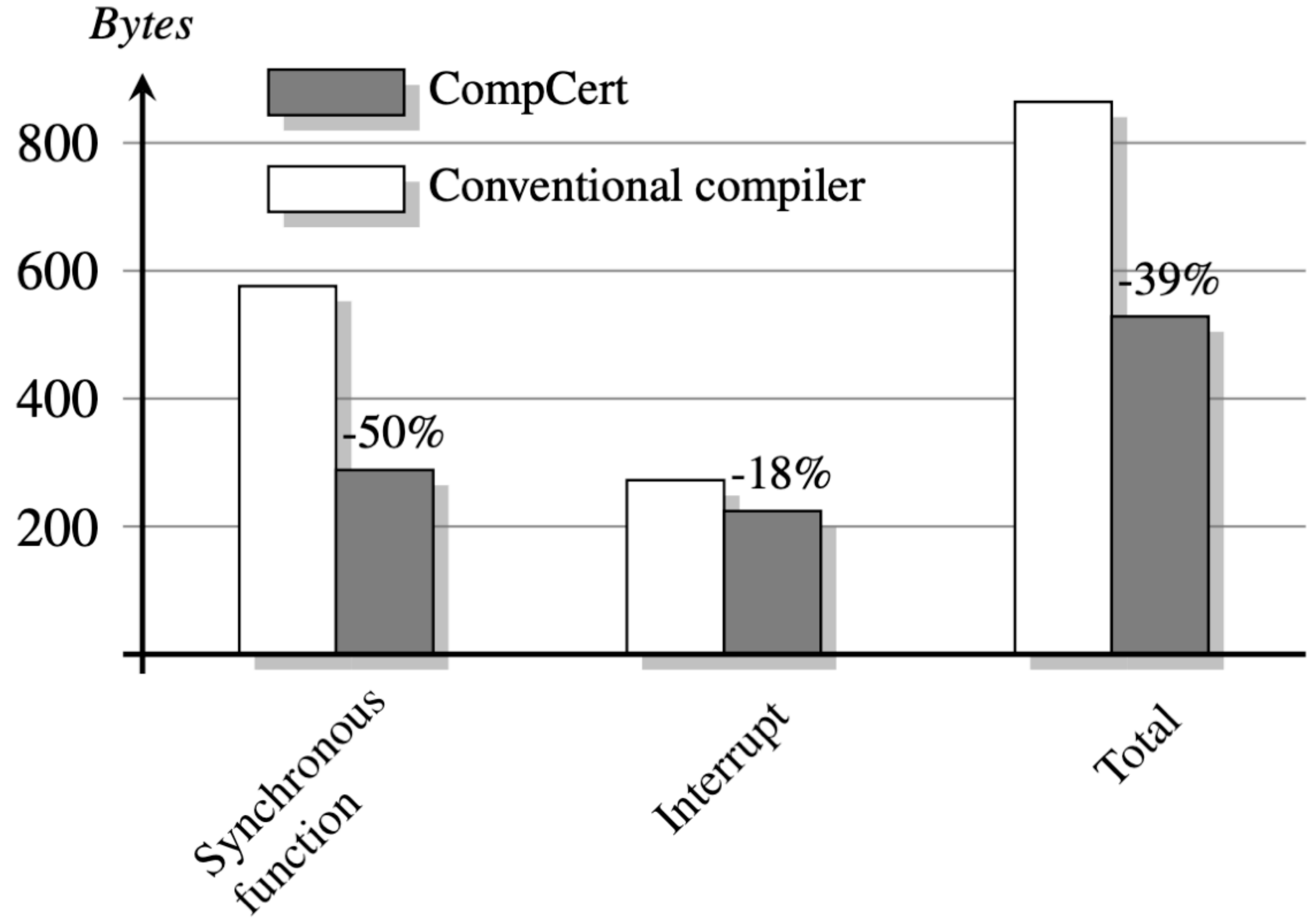


Compiling critical embedded software

[Kästner et al., CompCert: Practical experience on integrating and qualifying a formally verified optimizing compiler, ERTS'18]

WCET and stack use improvements on a real-time application
while providing proven traceability guarantees thanks to annotations





Overall assessment

The improvement mainly comes from the register allocation pass.

- From: no register allocation
- To: sharing of local variables among available registers

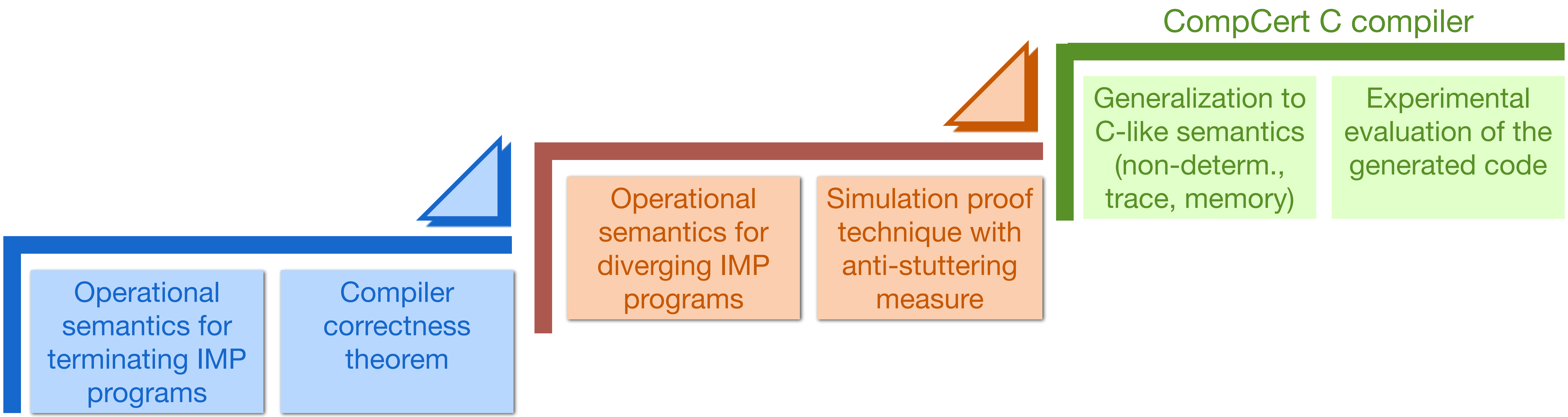
Traceability guarantees

- From: tracking of all program variables
- To: tracking of meaningful variables (used in block diagrams)

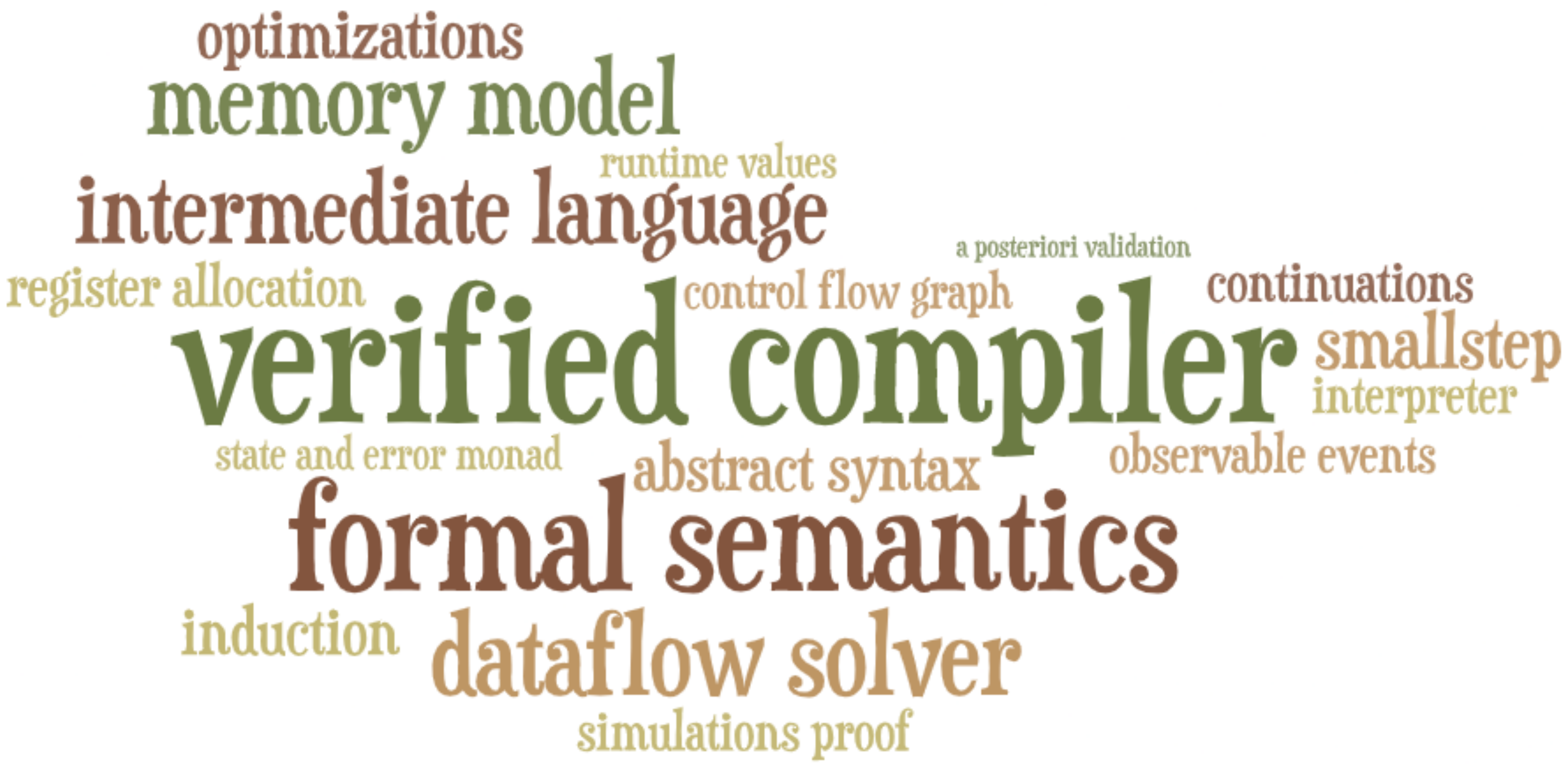
From early intuitions to fundamental formalisms …

verification tools that automate these ideas …

actual use in the critical software industry



Part 8:
CompCert, a shared infrastructure for ongoing research



Turning CompCert into a secure compiler

CT-CompCert

[Barthe, Blazy, Grégoire, Hutin, Laporte, Pichardie, Trieu, POPL'20]

Cryptographic constant-time (CCT) programming discipline

```
unsigned nok-function (unsigned x, unsigned y, bool secret)
{ if (secret) return y; else return x; }
```

```
unsigned ok-function (unsigned x, unsigned y, bool secret)
{ return x ^ ((y ^ x) & (-(unsigned)secret)); }
```

How to turn CompCert into a formally-verified secure compiler?

```
Theorem compiler-correct:
  ∀ S C b,
  compiler S = OK C →
  execCompCertC S b →
  execASM C b.
```

```
Theorem compiler-preserves-CCT:
  ∀ S C,
  compiler S = OK C →
  isCCT S →
  isCCT C.
```

Which proof technique for the isCCT policy?

Observational non-interference: observing program leakage (boolean guards and memory accesses) during execution does not reveal any information about secrets

```
Theorem compiler-preserves-CCT:
  ∀ S C,
  compiler S = OK C →
  isCCT S →
  isCCT C.
```

isCCT S

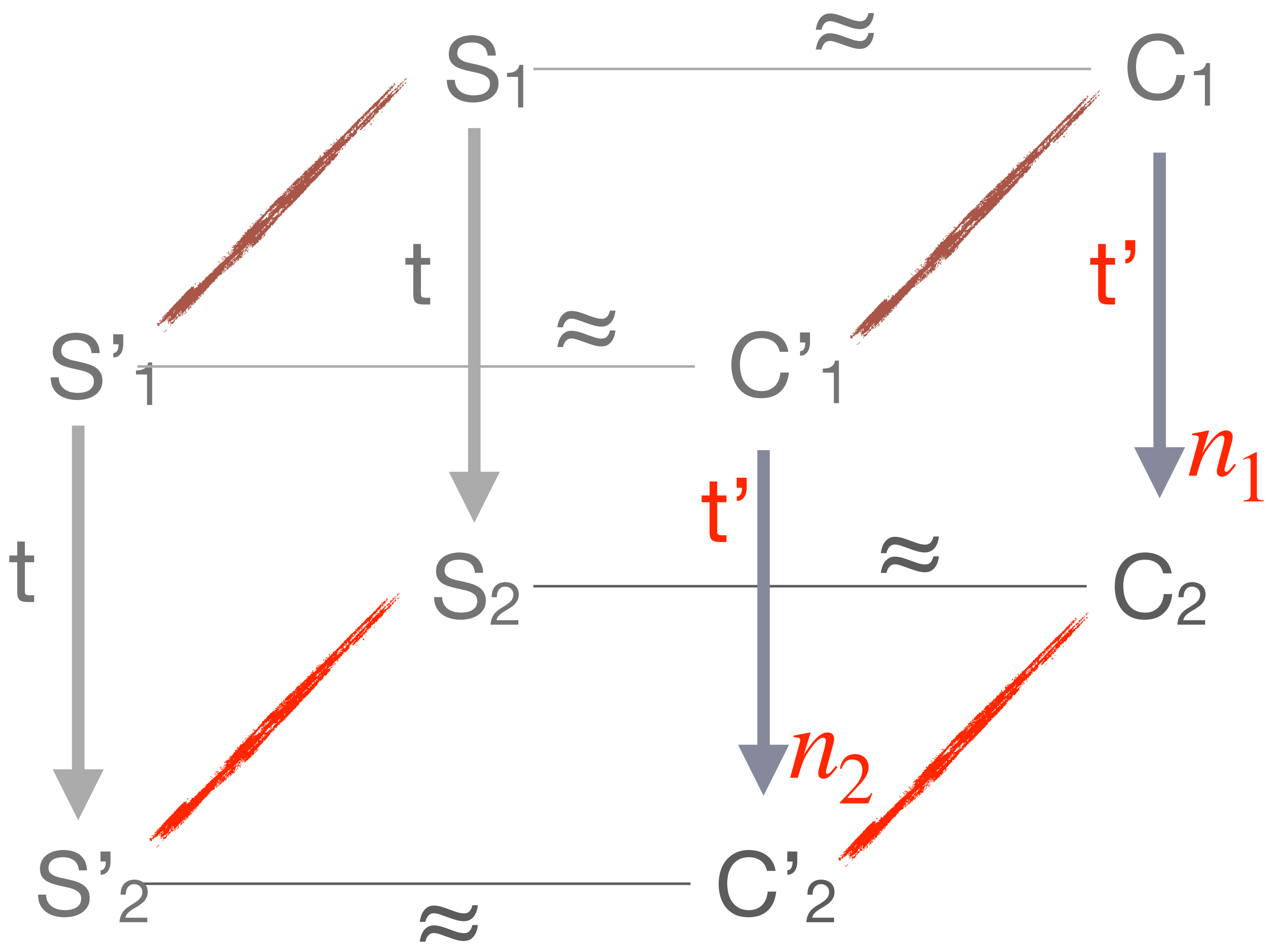
$S_1 \xrightarrow{\ell} S_2$

$S'_1 \xrightarrow{\ell'} S'_2$

with $\varphi(S_1, S'_1)$

implies $\ell = \ell'$

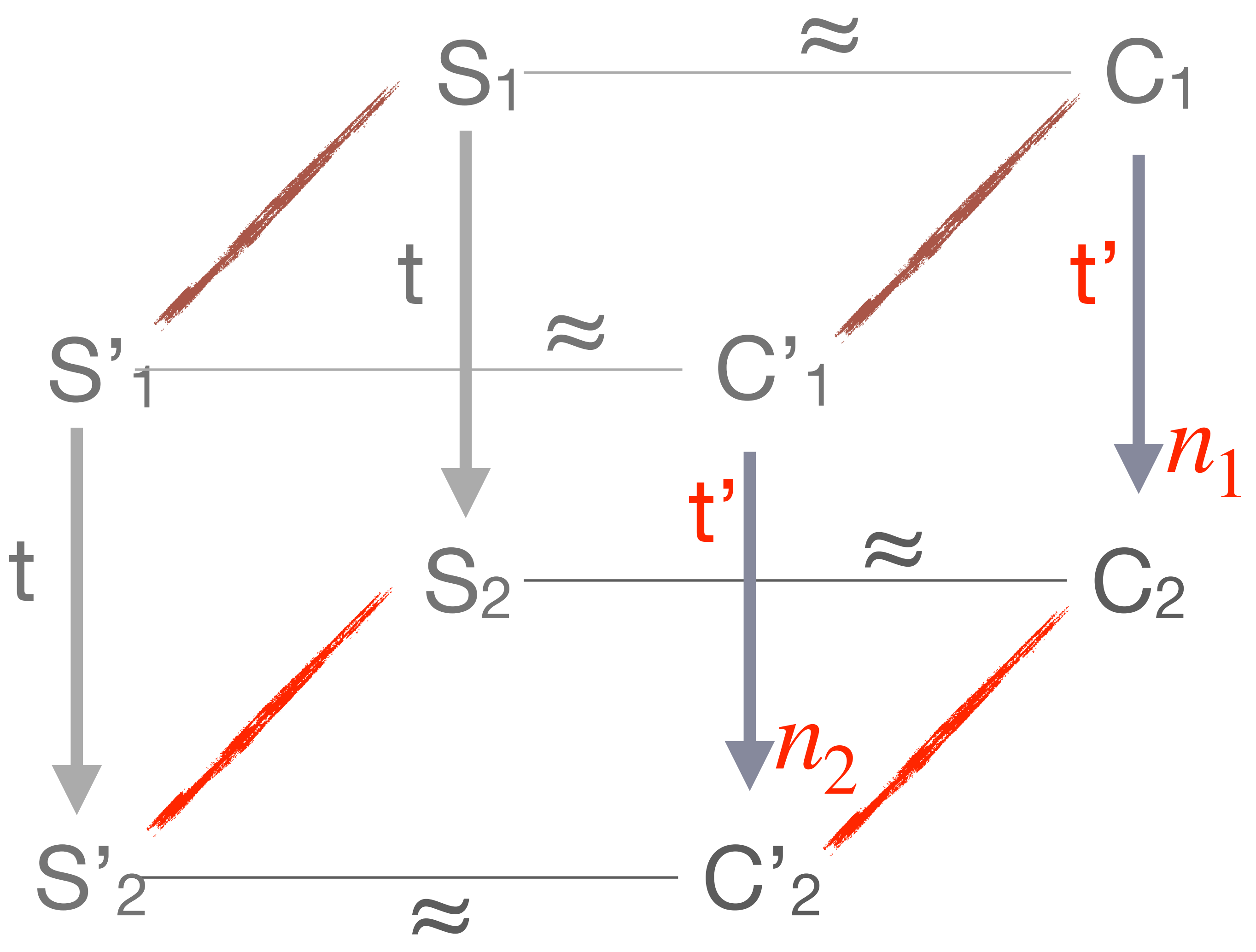
Indistinguishability property $\varphi(S_i, S'_i)$: share public values, but may differ on secret values

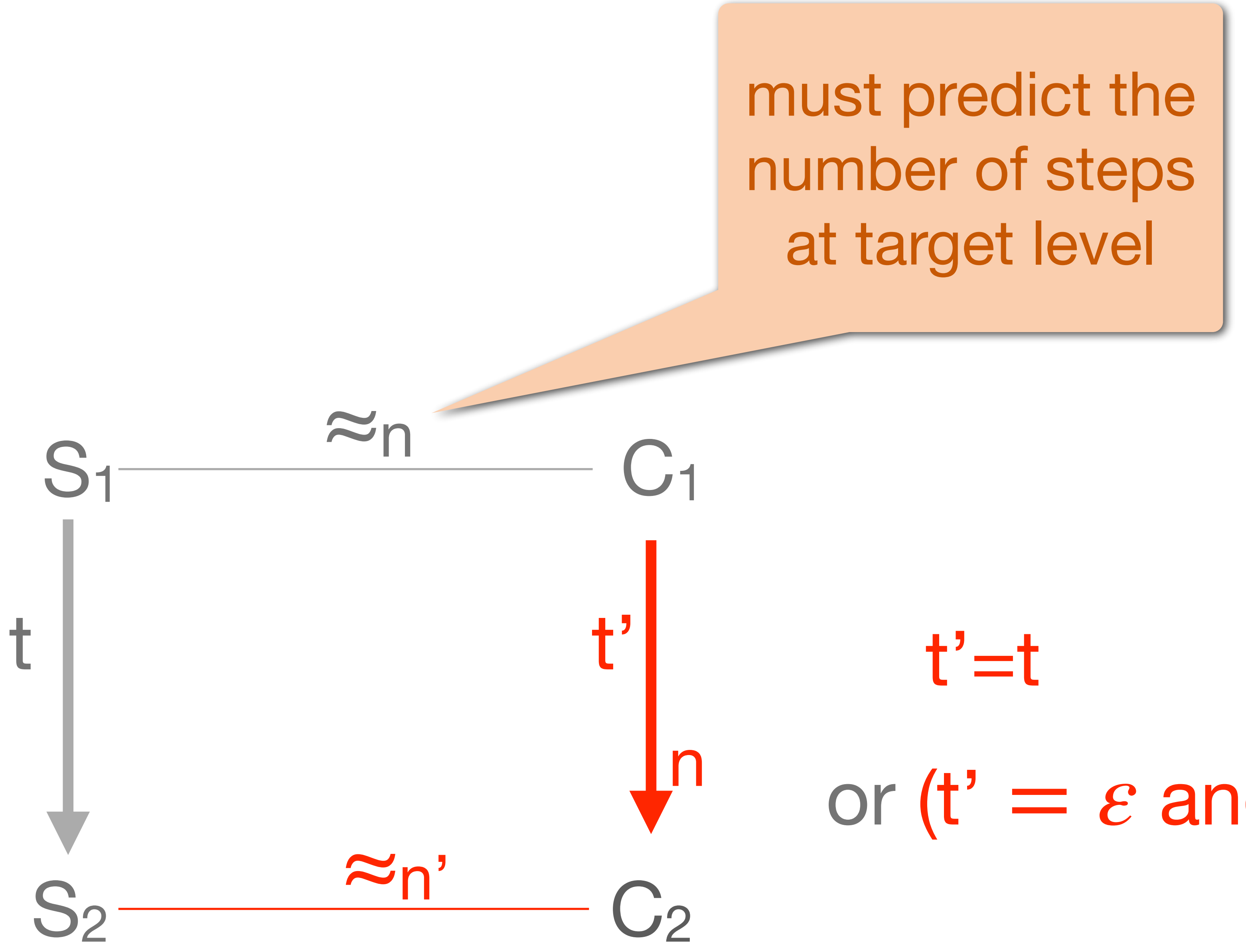


Difficulty: tricky proofs!

Proving CCT preservation: back to simulation diagrams

Proof-engineering: leverage the **existing proof scripts** as much as possible



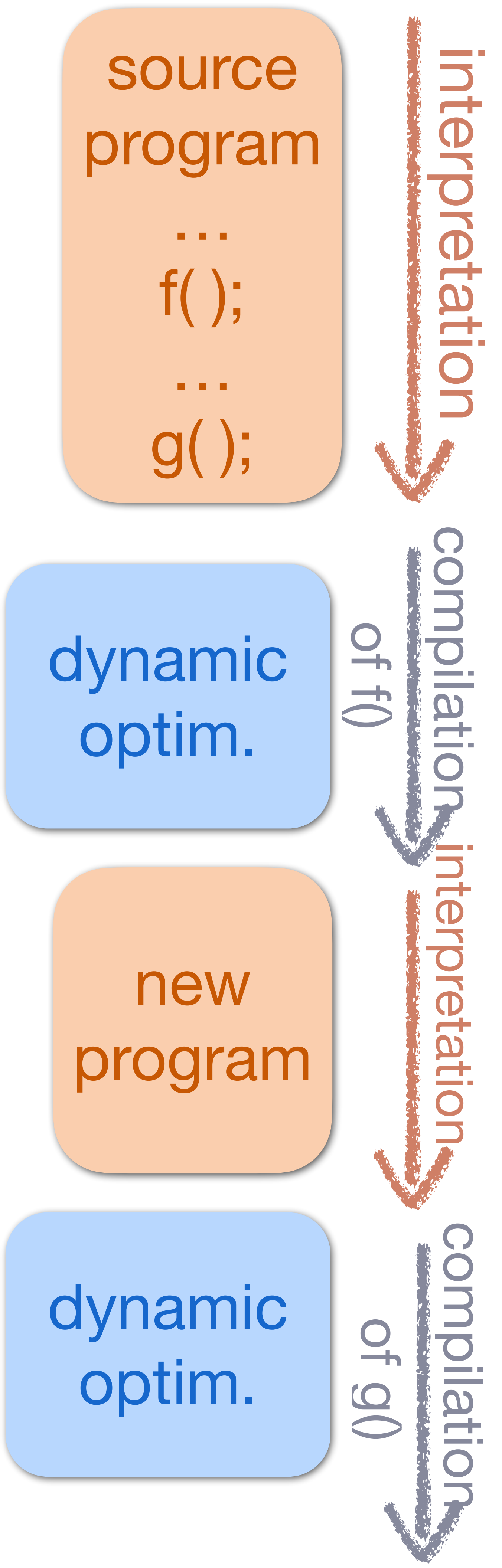


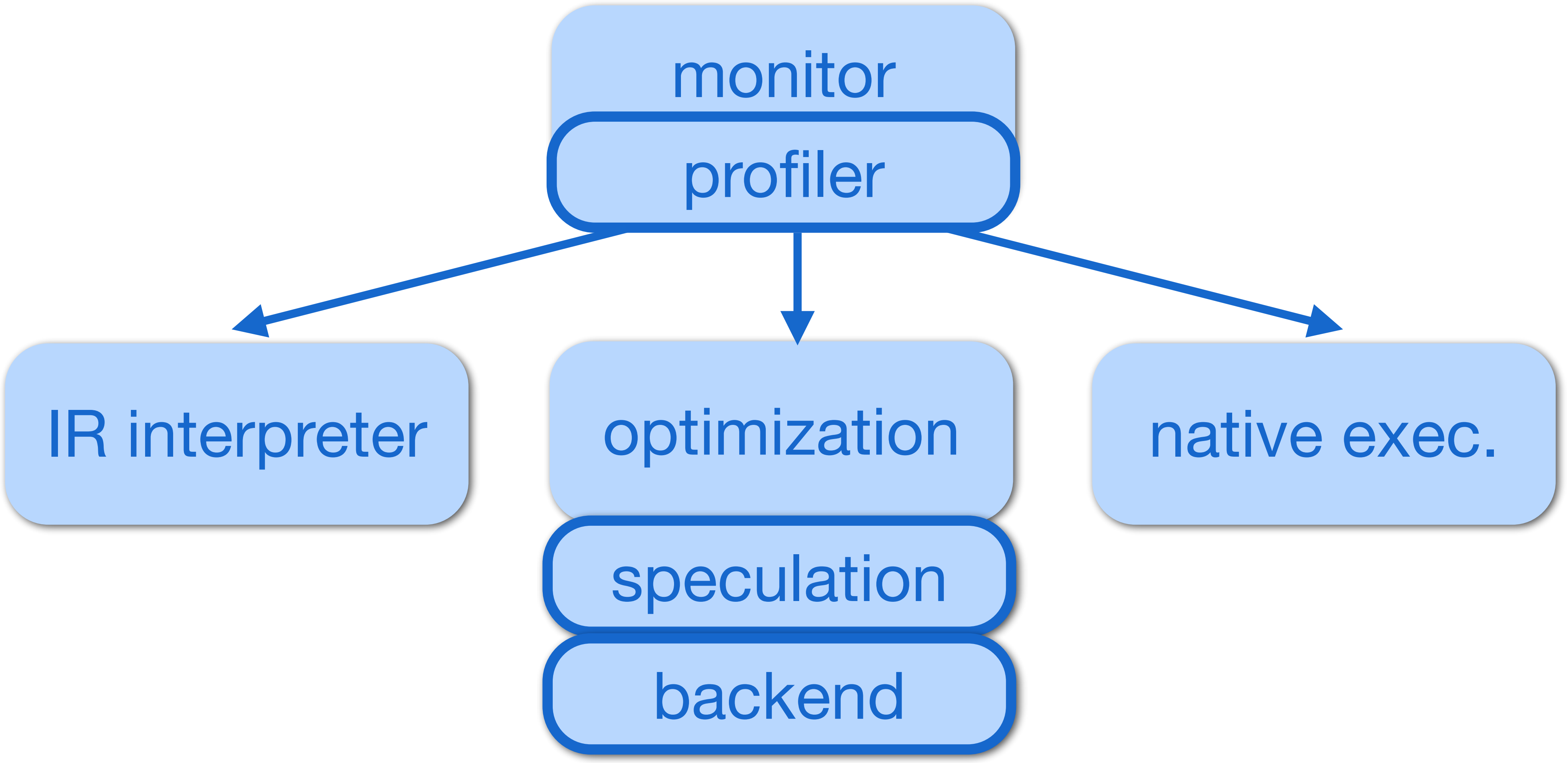
t'=t

or (t' $=\varepsilon$ and t is leak only)

Verifying just-in-time (JIT) compilation [Barrière's PhD 12/2022]

[Barrière, Blazy, Flückiger, Pichardie, Vitek, POPL'21] and [Barrière, Blazy, Pichardie, POPL'23]



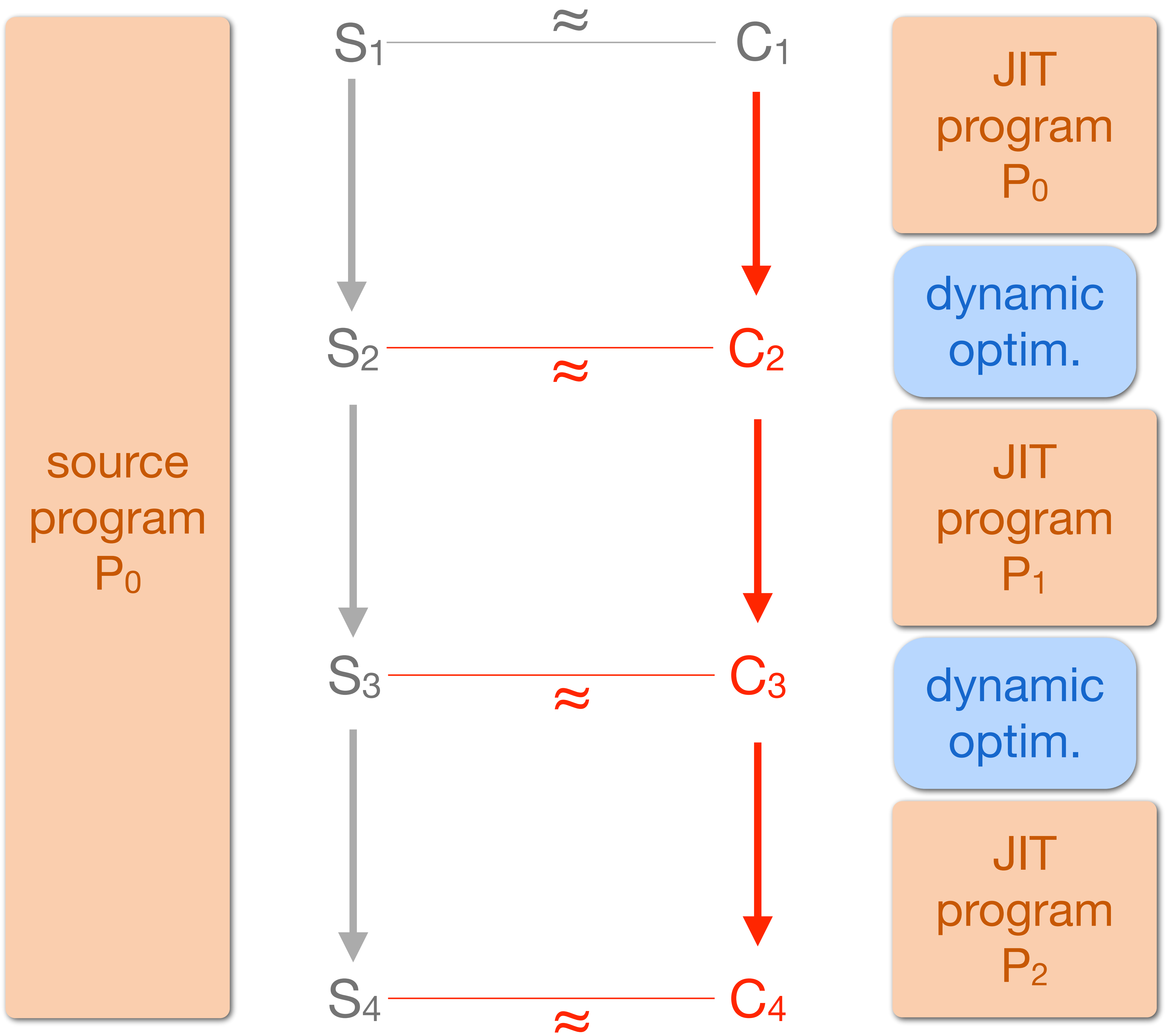


A JIT compiler **interleaves** the execution of a program with its optimizations

Dynamic speculation: specializes functions, requires deoptimization

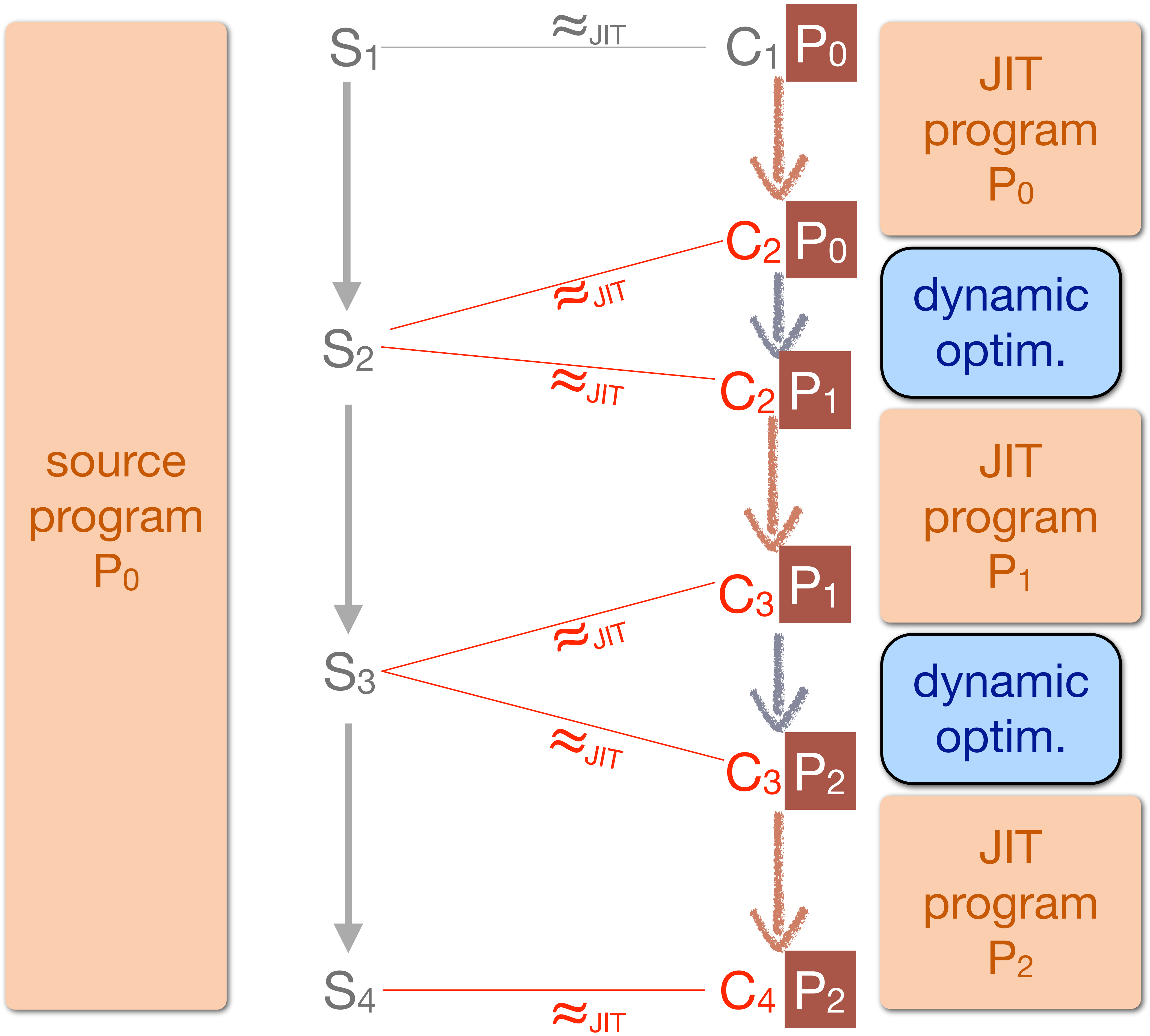
Non-deterministic semantics: either deoptimize to the source program or continue to the next instruction in the optimized program

Proving semantics preservation: the simulation approach



Both the program and the execution state are evolving

Nested simulations for JIT verification



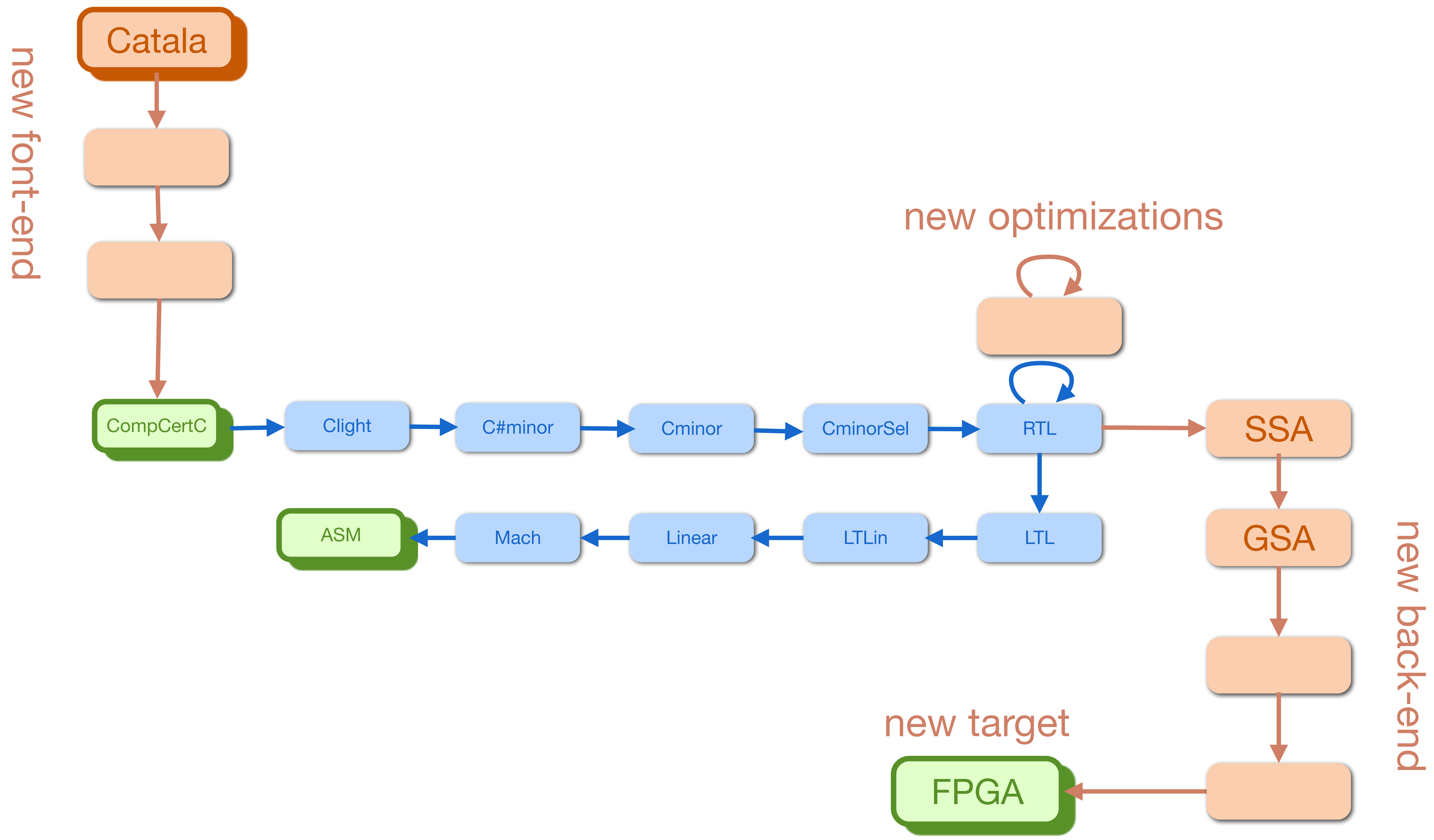
Both the program and the execution state are evolving

Invariant $\approx_{JIT}$: at any point during JIT execution

- the current state C_i corresponds to a source state S_i
- the curent JIT program P_i is equivalent to the source program P_0

Nested simulation: this equivalence is expressed with another simulation

Work in progress 🏗



Conclusion and perspectives

CompCert is a shared infrastructure for ongoing research

- **compilation** : ProbCompCert (Boston College, USA), L2C (Tsinghua, China), Velus (DIENS, Fr), CompCertO (Yale, USA), VeriCert (Imperial College, GB), CompCert-KVX (Verimag, Fr)
- **program logics**: VST (Princeton, USA), Gillian (Imperial College, GB), VeriFast (KUL, Be)
- **static analysis** : Verasco (Inria, Fr)

Opens the way to the trust of development tools

Bibliography

- Boyer, R. S., Moore, J S. MJRTY - A Fast Majority Vote Algorithm. *Essays in Honor of Woody Bledsoe*. 1991.
- Yang, Chen, Eide, Regehr. Finding and understanding bugs in C compilers. PLDI’11.
- Leroy. Formal verification of a realistic compiler. Communications of the ACM 52(7), 2009.
- Leroy. A formally verified compiler back-end. JAR 43(4), 2009.
- Appel, Blazy. Separation logic for small-step Cminor. TPHOLs’07.
- Blazy, Leroy. Mechanized semantics for the Clight subset of the C language. JAR 43(3), 2009.
- Leroy, Appel, Blazy, Stewart. The CompCert memory model. 2014. Program Logics for Certified Compilers.
- Leroy, Blazy, Kästner, Schommer, Pister, Ferdinand. CompCert - A formally verified optimizing compiler. ERTS2’16.
- Kumar, Myreen, Norrish, Owens. CakeML: a verified implementation of ML. POPL’14.
- Jourdan, Laporte, Blazy, Leroy, Pichardie. A formally-verified static analyzer. POPL’15.
- Blazy, Laporte, Pichardie. An Abstract Memory Functor for Verified C Static Analyzers. ICFP’16.
- Barthe, Blazy, Grégoire, Hutin, Laporte, Pichardie, Trieu. Formal verification of a constant-time preserving C compiler. POPL’20.
- Barrière, Blazy, Flückiger, Pichardie, Vitek. Formally verified speculation and deoptimization in a JIT compiler. POPL’21.
- Barrière, Blazy, Pichardie. Formally verified native code generation in an effectful JIT - or: Turning the CompCert backend into a formally verified JIT compiler. POPL’23.
- Barthe, Demange, Pichardie. Formal Verification of an SSA-based middle-end for CompCert. TOPLAS’14.
- Herklotz, Demange, Blazy. Mechanised semantics for gated static single assignment. CPP’23.
- Merigoux, Chataing, Protzenko. Catala: a programming language for the law. Proc. ICFP’21.