

Certified Abstract Interpretation with Pretty-Big-Step Semantics

Martin BODIN Thomas JENSEN Alan SCHMITT

Inria

12–14th of January

CPP'15

JSCert: A Trusted Mechanised JAVASCRIPT Specification



jscert.org

- An operational semantics for JAVASCRIPT;
- Trusted;
- *Huge* (~ 800 reduction rules).

How to derive
an abstract interpreter
from such a huge semantics?

... proven in Coq?

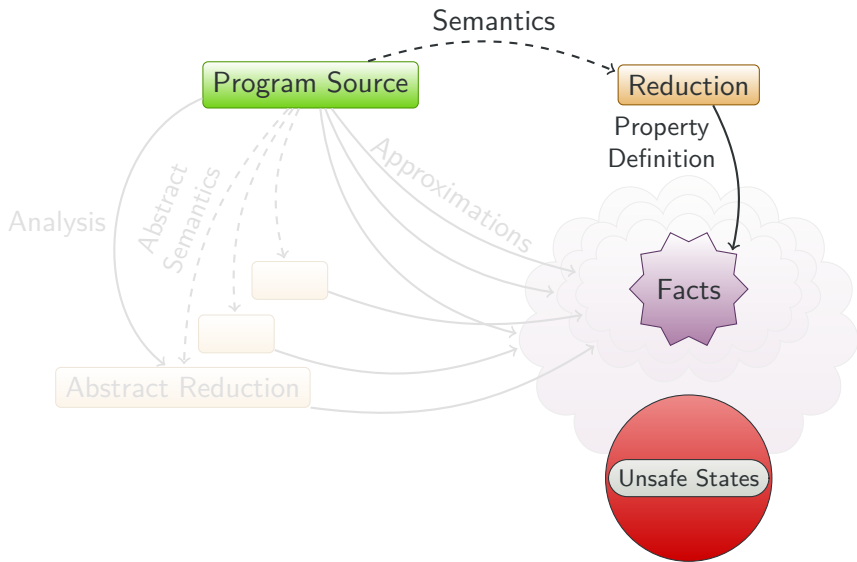
Let's make it correct *by construction!*

How to derive
an abstract interpreter
from such a huge semantics?

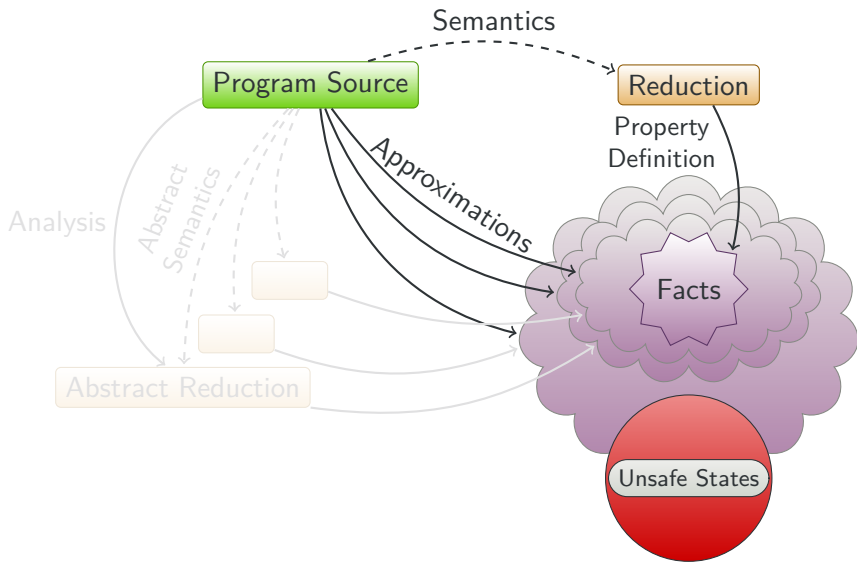
... proven in Coq?

Let's make it correct *by construction!*

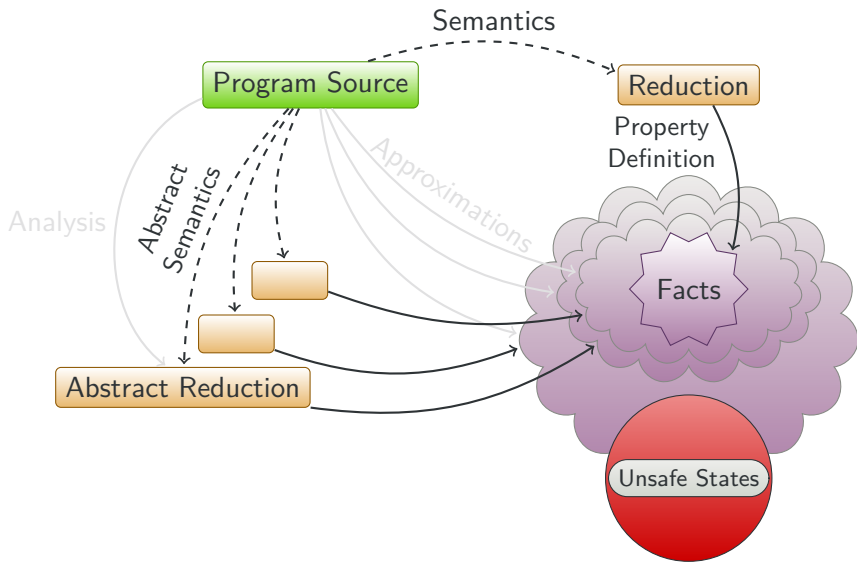
Abstract Interpretation



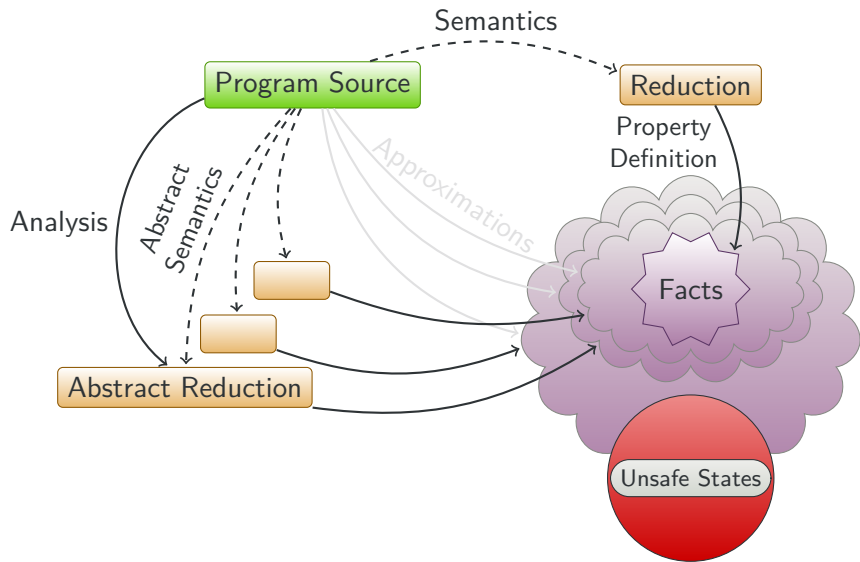
Abstract Interpretation



Abstract Interpretation



Abstract Interpretation



Concrete Domains

`int, bool`

Concrete Operations

`+, =`

Concrete Semantics

$t, \sigma \Downarrow r$

Abstract Domains

Sign

Abstract Operations

$+^\#, =^\#$

Abstract Semantics

$t, \sigma^\# \Downarrow^\# r^\#$

Abstract Interpreter

$f(t, \sigma^\#) = r^\#$

Concrete Domains

int, bool

Concrete Operations

+, =

Concrete Semantics

$t, \sigma \Downarrow r$

Abstract Domains

Sign

Abstract Operations

$+^\#, =^\#$

Abstract Semantics

$t, \sigma^\# \Downarrow^\# r^\#$

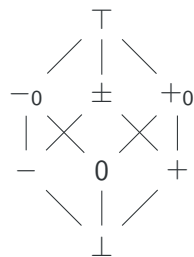
Abstract Interpreter

$f(t, \sigma^\#) = r^\#$

Defining Abstract Domains



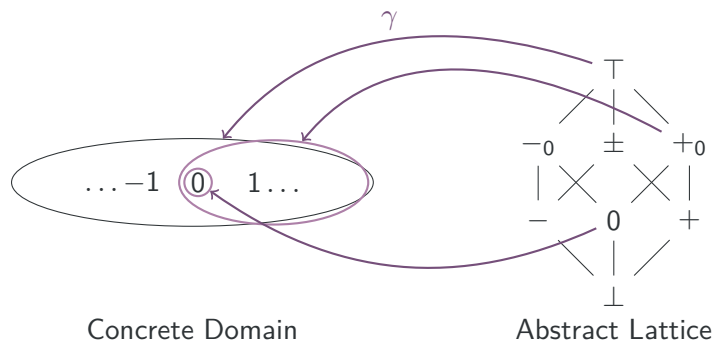
Concrete Domain



Abstract Lattice

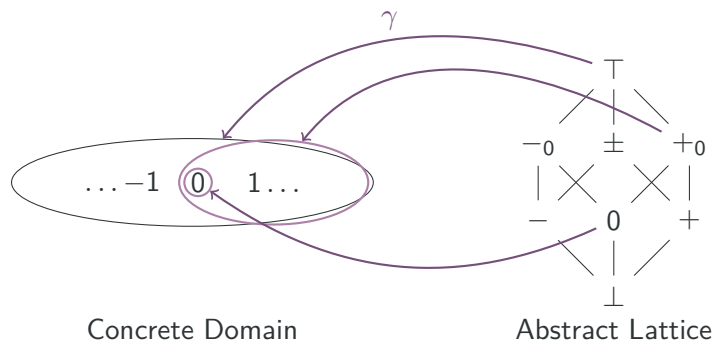
The theory has already been formalized in Coq.

Defining Abstract Domains



The theory has already been formalized in Coq.

Defining Abstract Domains



The theory has already been formalized in Coq.

Concrete Domains

int, bool

Concrete Operations

+, =

Concrete Semantics

$t, \sigma \Downarrow r$

Abstract Domains

Sign

Abstract Operations

$+^\#, =^\#$

Abstract Semantics

$t, \sigma^\# \Downarrow^\# r^\#$

Abstract Interpreter

$f(t, \sigma^\#) = r^\#$

Defining an Abstract Semantics, the Direct Approach

$$\frac{\text{IFTRUE} \quad e, E \Downarrow v \quad s_1, E \Downarrow E'}{\text{if } e \text{ } s_1 \text{ } s_2, E \Downarrow E'} \quad v \neq 0$$

$$\frac{\text{IFFALSE} \quad e, E \Downarrow v \quad s_2, E \Downarrow E'}{\text{if } e \text{ } s_1 \text{ } s_2, E \Downarrow E'} \quad v = 0$$

Let's just add \Downarrow everywhere!

Defining an Abstract Semantics, the Direct Approach

$$\frac{\text{IFTRUE} \quad e, E \Downarrow v \quad s_1, E \Downarrow E'}{\text{if } e \text{ } s_1 \text{ } s_2, E \Downarrow E'} \quad v \neq 0 \qquad \frac{\text{IFFALSE} \quad e, E \Downarrow v \quad s_2, E \Downarrow E'}{\text{if } e \text{ } s_1 \text{ } s_2, E \Downarrow E'} \quad v = 0$$

Let's just add \sharp everywhere!

$$\frac{\text{IFTRUE} \quad e, E^\sharp \Downarrow^\sharp v^\sharp \quad s_1, E^\sharp \Downarrow^\sharp E'^\sharp}{\text{if } e \text{ } s_1 \text{ } s_2, E^\sharp \Downarrow^\sharp E'^\sharp} \quad \gamma(v^\sharp) \cap \mathbb{Z}^* \neq \emptyset$$

$$\frac{\text{IFFALSE} \quad e, E^\sharp \Downarrow^\sharp v^\sharp \quad s_2, E^\sharp \Downarrow^\sharp E'^\sharp}{\text{if } e \text{ } s_1 \text{ } s_2, E^\sharp \Downarrow^\sharp E'^\sharp} \quad \gamma(v^\sharp) \cap \{0\} \neq \emptyset$$

Defining an Abstract Semantics, the Direct Approach

$$\frac{\text{IFTRUE} \quad e, E \Downarrow v \quad s_1, E \Downarrow E'}{\text{if } e \text{ } s_1 \text{ } s_2, E \Downarrow E'} \quad v \neq 0 \qquad \frac{\text{IFFALSE} \quad e, E \Downarrow v \quad s_2, E \Downarrow E'}{\text{if } e \text{ } s_1 \text{ } s_2, E \Downarrow E'} \quad v = 0$$

Let's just add \sharp everywhere!

$$\frac{\text{IFADHOC} \quad e, E^\sharp \Downarrow^\sharp v^\sharp \quad s_1, E^\sharp \Downarrow^\sharp E_1^\sharp \quad s_2, E^\sharp \Downarrow^\sharp E_2^\sharp}{\text{if } e \text{ } s_1 \text{ } s_2, E^\sharp \Downarrow^\sharp E_1^\sharp \sqcup E_2^\sharp} \quad v^\sharp = \top$$

- 1 Motivation
- 2 Pretty-Big-Step: a Generic Rule Format
- 3 Defining an Abstract Semantics Correct by Construction
- 4 Running Abstract Interpreters

Pretty-Big-Step

- Introduced by CHARGÉRAUD (ESOP 2013).
- Can be compiled from Small-Step (ESOP 2014).
- Similar to Big-Step semantics.
- But much more constrained.

Pretty-Big-Step

- Introduced by CHARGÉRAUD (ESOP 2013).
- Can be compiled from Small-Step (ESOP 2014).
- Similar to Big-Step semantics.
- But much more constrained.

AXIOM

$$\frac{}{\mathfrak{l}, \sigma \Downarrow ax(\sigma)} \quad cond(\sigma)$$

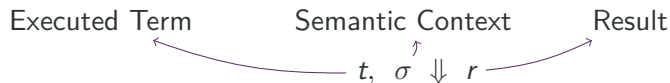
RULE1

$$\frac{u_1, up(\sigma) \Downarrow r}{\mathfrak{l}, \sigma \Downarrow r} \quad cond(\sigma)$$

RULE2

$$\frac{u_2, up(\sigma) \Downarrow r \quad n_2, next(\sigma, r) \Downarrow r'}{\mathfrak{l}, \sigma \Downarrow r'} \quad cond(\sigma)$$

Pretty-Big-Step



Each rule has

- A structural part: identifier, terms;
- A semantic part: side-conditions, transfer functions.

AXIOM

$$\frac{}{l, \sigma \Downarrow ax(\sigma)} \quad cond(\sigma)$$

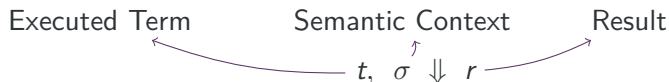
RULE1

$$\frac{u_1, up(\sigma) \Downarrow r}{l, \sigma \Downarrow r} \quad cond(\sigma)$$

RULE2

$$\frac{u_2, up(\sigma) \Downarrow r \quad n_2, next(\sigma, r) \Downarrow r'}{l, \sigma \Downarrow r'} \quad cond(\sigma)$$

Pretty-Big-Step



Each rule has

- A structural part: **identifier**, terms;
- A semantic part: side-conditions, transfer functions.

AXIOM

$$\frac{}{\mathfrak{l}, \sigma \Downarrow ax(\sigma)} \quad cond(\sigma)$$

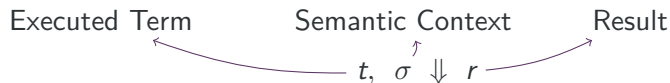
RULE1

$$\frac{u_1, up(\sigma) \Downarrow r}{\mathfrak{l}, \sigma \Downarrow r} \quad cond(\sigma)$$

RULE2

$$\frac{u_2, up(\sigma) \Downarrow r \quad n_2, next(\sigma, r) \Downarrow r'}{\mathfrak{l}, \sigma \Downarrow r'} \quad cond(\sigma)$$

Pretty-Big-Step



Each rule has

- A structural part: identifier, **terms**;
- A semantic part: side-conditions, transfer functions.

AXIOM

$$\frac{}{\mathfrak{l}, \sigma \Downarrow ax(\sigma)} \quad cond(\sigma)$$

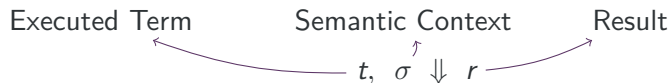
RULE1

$$\frac{\mathfrak{u}_1, up(\sigma) \Downarrow r}{\mathfrak{l}, \sigma \Downarrow r} \quad cond(\sigma)$$

RULE2

$$\frac{\mathfrak{u}_2, up(\sigma) \Downarrow r \quad \mathfrak{n}_2, next(\sigma, r) \Downarrow r'}{\mathfrak{l}, \sigma \Downarrow r'} \quad cond(\sigma)$$

Pretty-Big-Step



Each rule has

- A structural part: identifier, terms;
- A semantic part: **side-conditions**, transfer functions.

AXIOM

$$\frac{}{l, \sigma \Downarrow ax(\sigma)} \quad \text{cond}(\sigma)$$

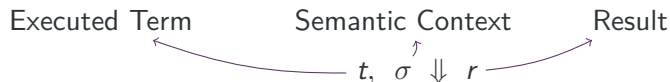
RULE1

$$\frac{u_1, up(\sigma) \Downarrow r}{l, \sigma \Downarrow r} \quad \text{cond}(\sigma)$$

RULE2

$$\frac{u_2, up(\sigma) \Downarrow r \quad n_2, next(\sigma, r) \Downarrow r'}{l, \sigma \Downarrow r'} \quad \text{cond}(\sigma)$$

Pretty-Big-Step



Each rule has

- A structural part: identifier, terms;
- A semantic part: side-conditions, **transfer functions**.

AXIOM

$$\frac{}{l, \sigma \Downarrow \mathbf{ax}(\sigma)} \quad \mathit{cond}(\sigma)$$

RULE1

$$\frac{u_1, \mathbf{up}(\sigma) \Downarrow r}{l, \sigma \Downarrow r} \quad \mathit{cond}(\sigma)$$

RULE2

$$\frac{u_2, \mathbf{up}(\sigma) \Downarrow r \quad n_2, \mathbf{next}(\sigma, r) \Downarrow r'}{l, \sigma \Downarrow r'} \quad \mathit{cond}(\sigma)$$

Each rule has

- A structural part: identifier, terms;
- A semantic part: side-conditions, transfer functions.

Abstract Rules

Shared between the concrete and abstract semantics

Each rule has

- A structural part: identifier, terms;
- A semantic part: side-conditions, transfer functions.

To be specified in the abstract semantics.
To be *locally* proved correct.

- The abstract semantics will follow the exact same structure as the concrete semantics.

But we don't define \Downarrow and $\Downarrow^\#$ the same way from the rules!

Concrete Semantics \Downarrow

At each step,
apply *one* rule that applies

Abstract Semantics $\Downarrow^\#$

At each step,
apply *all* the rules that apply

But we don't define \Downarrow and $\Downarrow^\#$ the same way from the rules!

Concrete Semantics \Downarrow

At each step,
apply *one* rule that applies

Inductive interpretation
of the rules

$$\Downarrow = \text{lfp}(\mathcal{F})$$

Abstract Semantics $\Downarrow^\#$

At each step,
apply *all* the rules that apply

Co-inductive interpretation
of the rules

$$\Downarrow^\# = \text{gfp}(\mathcal{F}^\#)$$

But we don't define \Downarrow and \Downarrow^\sharp the same way from the rules!

Concrete Semantics \Downarrow

At each step,
apply *one* rule that applies

Inductive interpretation
of the rules

$$\Downarrow = \text{lfp}(\mathcal{F})$$

Abstract Semantics \Downarrow^\sharp

At each step,
apply *all* the rules that apply

Co-inductive interpretation
of the rules

$$\Downarrow^\sharp = \text{gfp}(\mathcal{F}^\sharp)$$

Allow approximations

Example of Concrete Rules

$$\frac{\text{WHILE}(e, s) \quad \text{while}_1 e s, \text{ret } E \Downarrow o}{\text{while } e s, E \Downarrow o}$$

$$\frac{\text{WHILE1}(e, s) \quad e, E \Downarrow o \quad \text{while}_2 e s, (E, o) \Downarrow o'}{\text{while}_1 e s, \text{ret } E \Downarrow o'}$$

$$\frac{\text{WHILE2TRUE}(e, s) \quad s, E \Downarrow o \quad \text{while}_1 e s, o \Downarrow o'}{\text{while}_2 e s, (E, \text{val } v) \Downarrow o'} \quad v \neq 0$$

$$\frac{}{\text{while}_2 e s, (E, \text{val } v) \Downarrow \text{ret } E} \quad v = 0$$

Example of Abstract Rules

$$\frac{\text{WHILE}(e, s) \quad \text{while}_1 e s, E^\# \Downarrow^\# o^\#}{\text{while } e s, E^\# \Downarrow^\# o^\#}$$

$$\frac{\text{WHILE1}(e, s) \quad e, E^\# \Downarrow^\# v^\# \quad \text{while}_2 e s, (E^\#, v^\#) \Downarrow^\# o^\#}{\text{while}_1 e s, E^\# \Downarrow^\# o^\#}$$

$$\frac{\text{WHILE2TRUE}(e, s) \quad s, E^\# \Downarrow^\# o \quad \text{while}_1 e s, o^\# \Downarrow^\# o'^\#}{\text{while}_2 e s, (E^\#, v^\#) \Downarrow^\# o'^\#} \quad \gamma(v^\#) \cap \mathbb{Z}^* \neq \emptyset$$

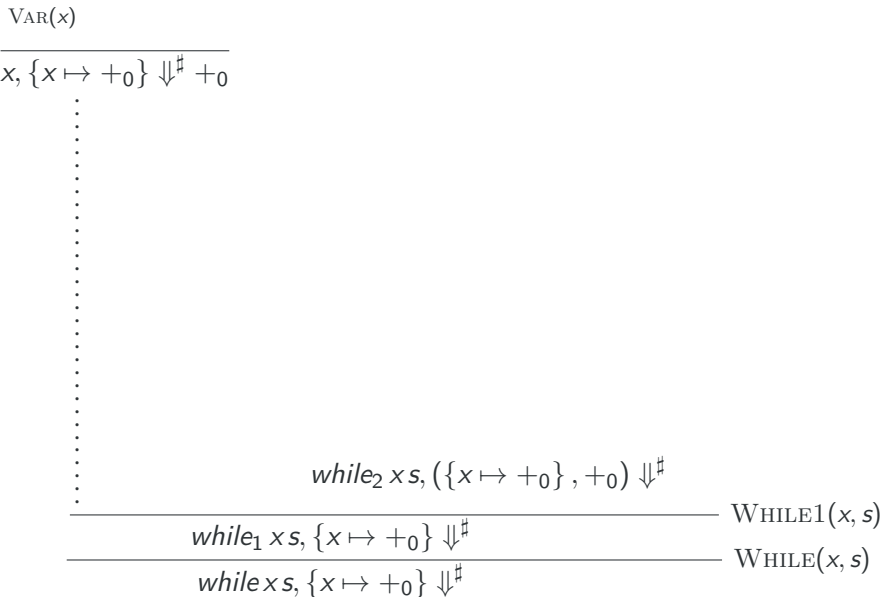
$$\frac{\text{WHILE2FALSE}(e, s)}{\text{while}_2 e s, (E^\#, v^\#) \Downarrow^\# E^\#} \quad 0 \in \gamma(v^\#)$$

Example of Abstract Derivation Tree

$$\frac{\text{WHILE1}(e, s) \quad e, E \Downarrow o \quad \text{while}_2 e s, (E, o) \Downarrow o'}{\text{while}_1 e s, \text{ret } E \Downarrow o'}$$

$$\frac{\text{WHILE1}(x, s) \quad \text{while}_1 x s, \{x \mapsto +0\} \Downarrow \#}{\text{WHILE}(x, s) \quad \text{while } x s, \{x \mapsto +0\} \Downarrow \#}$$

Example of Abstract Derivation Tree



Example of Abstract Derivation Tree

VAR(x)

$x, \{x \mapsto +0\} \Downarrow^\# +0$

⋮

$$\frac{\text{WHILE2TRUE}(e, s) \quad s, E^\# \Downarrow^\# o \quad \text{while}_1 e s, o^\# \Downarrow^\# o'^\#}{\text{while}_2 e s, (E^\#, v^\#) \Downarrow^\# o'^\#} \quad \gamma(v^\#) \cap \mathbb{Z}^* \neq \emptyset$$

WHILE2FALSE(e, s)

$$\frac{}{\text{while}_2 e s, (E^\#, v^\#) \Downarrow^\# E^\#} \quad 0 \in \gamma(v^\#)$$

$\text{while}_2 x s, (\{x \mapsto +0\}, +0) \Downarrow^\#$

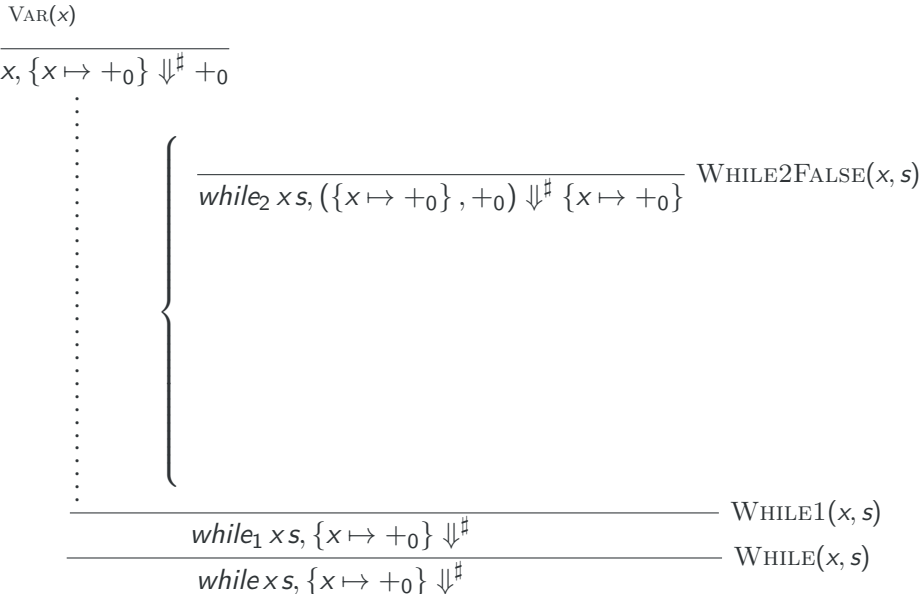
$\text{while}_1 x s, \{x \mapsto +0\} \Downarrow^\#$

WHILE1(x, s)

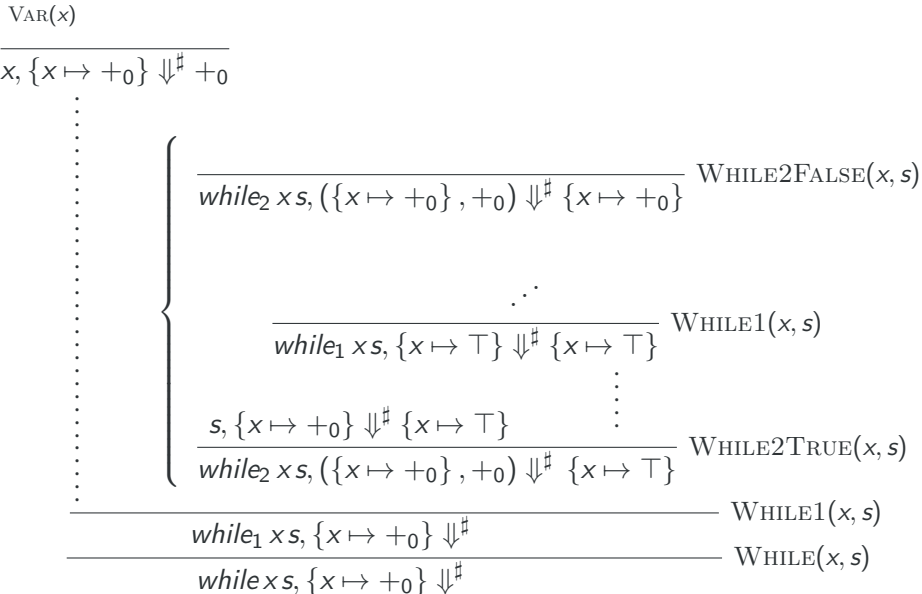
$\text{while } x s, \{x \mapsto +0\} \Downarrow^\#$

WHILE(x, s)

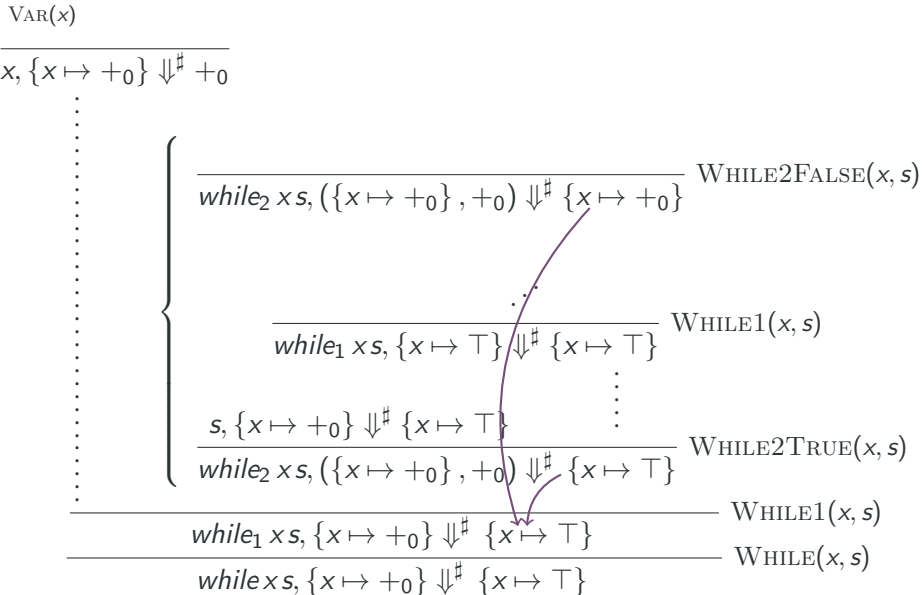
Example of Abstract Derivation Tree



Example of Abstract Derivation Tree



Example of Abstract Derivation Tree



An Abstract Semantics Correct by Construction

Hypotheses:

- Correctness of the side-conditions,
- Correctness of the transfer functions.

Theorem (Correctness)

Let t a term, σ and $\sigma^\#$ a concrete and an abstract semantic contexts, and r and $r^\#$ a concrete and an abstract results.

$$\text{If } \begin{cases} \sigma \in \gamma(\sigma^\#) \\ t, \sigma \Downarrow r \\ t, \sigma^\# \Downarrow^\# r^\# \end{cases} \text{ then } r \in \gamma(r^\#).$$



An Abstract Semantics Correct by Construction

Hypotheses:

- Correctness of the side-conditions,
- Correctness of the transfer functions.

Theorem (Correctness)

Let t a term, σ and $\sigma^\#$ a concrete and an abstract semantic contexts, and r and $r^\#$ a concrete and an abstract results.

$$\text{If } \begin{cases} \sigma \in \gamma(\sigma^\#) \\ t, \sigma \Downarrow r \\ t, \sigma^\# \Downarrow^\# r^\# \end{cases} \text{ then } r \in \gamma(r^\#).$$



Proven independently of
the rules!

Concrete Domains

`int, bool`

Concrete Operations

`+, =`

Concrete Semantics

$t, \sigma \Downarrow r$

Abstract Domains

Sign

Abstract Operations

$+^\#, =^\#$

Abstract Semantics

$t, \sigma^\# \Downarrow^\# r^\#$

Abstract Interpreter

$f(t, \sigma^\#) = r^\#$

Defining Abstract Interpreters: a Verifier

- An abstract interpreter is a function building an abstract derivation.
- But this abstract semantic tree can be infinite!

A Verifier

- It takes an oracle, i.e., a set O of triples $t, \sigma^\#, r^\#$.



It tries to prove $O \subseteq \mathcal{F}^{\#*}(O)$.

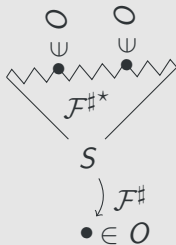
By PARK's principle, this implies $O \subseteq \Downarrow^\#$.

Defining Abstract Interpreters: a Verifier

- An abstract interpreter is a function building an abstract derivation.
- But this abstract semantic tree can be infinite!

A Verifier

- It takes an oracle, i.e., a set O of triples $t, \sigma^\#, r^\#$.



It tries to prove $O \subseteq \mathcal{F}^{\#*}(O)$.

By PARK's principle, this implies $O \subseteq \Downarrow^\#$.

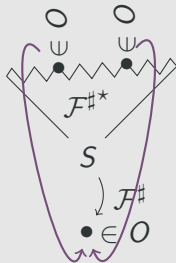


Defining Abstract Interpreters: a Verifier

- An abstract interpreter is a function building an abstract derivation.
- But this abstract semantic tree can be infinite!

A Verifier

- It takes an oracle, i.e., a set O of triples $t, \sigma^\#, r^\#$.



It tries to prove $O \subseteq \mathcal{F}^{\#*}(O)$.

By PARK's principle, this implies $O \subseteq \Downarrow^\#$.



Generic Abstract Interpreters

- We have built some *generic* abstract interpreters.
- We can extract them to OCaml and run them.

```
a := 6; b := 7; r := 0; n := a; while n (r := r + b; n := n - 1)
```

$$(\{r \mapsto +, b \mapsto +, a \mapsto +, n \mapsto \top\}, \perp)$$

Generic Abstract Interpreters

- We have built some *generic* abstract interpreters.
- We can extract them to OCaml and run them.

```
a := 6; b := 7; prod(n) := {if n (prod(n - 1); r := r + b) (r := 0)}; prod(a)
```

$(\{r \mapsto +, b \mapsto +, a \mapsto +\}, \perp)$

Generic Abstract Interpreters

- We have built some *generic* abstract interpreters.
- We can extract them to OCaml and run them.

```
a := 6; b := 7; prod(n) := {if n (prod(n - 1); r := r + b) (r := 0)}; prod(a)
```

$(\{r \mapsto +, b \mapsto +, a \mapsto +\}, \perp)$

Conclusion and Future Works

We have investigated how to define, in COQ , certified abstract interpreters for pretty-big-step semantics.

Recipe

- 1 define the concrete semantics;
- 2 define the abstract domains and operations on the abstract domain,
 - this automatically defines an abstract semantics;
- 3 prove the abstract operations are correct,
 - this implies the abstract semantics is correct;
- 4 define an analysis.

Future Works

- Apply it to JSCert.
- Allow non-local reasoning.
- Taking into account non-terminating behaviours.

Thanks You for Listening!

Concrete Domains

`int, bool`

Concrete Operations

`+, =`

Concrete Semantics

$t, \sigma \Downarrow r$

Abstract Domains

Sign

Abstract Operations

$+^\#, =^\#$

Abstract Semantics

$t, \sigma^\# \Downarrow^\# r^\#$

Abstract Interpreter

$f(t, \sigma^\#) = r^\#$

Bonus Slides

$$\begin{array}{l}
 \text{apply}_i(\Downarrow_0) := \\
 \left| \begin{array}{l}
 \text{match rule}(i) \text{ with} \\
 | Ax(ax) \quad \Rightarrow \{(l_i, \sigma, r) \mid ax(\sigma) = \text{Some}(r)\} \\
 \\
 | R_1(up) \quad \Rightarrow \left\{ (l_i, \sigma, r) \mid \begin{array}{l} up(\sigma) = \text{Some}(\sigma') \\ \wedge u_{1,i}, \sigma' \Downarrow_0 r \end{array} \right\} \\
 \\
 | R_2(up, next) \Rightarrow \left\{ (l_i, \sigma, r) \mid \begin{array}{l} up(\sigma) = \text{Some}(\sigma') \\ \wedge u_{2,i}, \sigma' \Downarrow_0 r_1 \\ \wedge next(\sigma, r_1) = \text{Some}(\sigma'') \\ \wedge n_{2,i}, \sigma'' \Downarrow_0 \text{Some}(r) \end{array} \right\}
 \end{array} \right.
 \end{array}$$

$$\Downarrow = \text{lfp}(\mathcal{F})$$

$$\mathcal{F}(\Downarrow_0) = \{(t, \sigma, r) \mid \exists i, \text{cond}_i(\sigma) \wedge (t, \sigma, r) \in \text{apply}_i(\Downarrow_0)\}$$

$$\mathit{apply}_i^\#(\Downarrow_0^\#) = \left\{ (t, \sigma, r) \mid \begin{array}{l} \exists \sigma_0, \exists r_0, \\ \sigma \sqsubseteq^\# \sigma_0 \wedge r_0 \sqsubseteq^\# r \wedge \\ (t, \sigma_0, r_0) \in \mathit{apply}_i(\Downarrow_0^\#) \end{array} \right\}$$

$$\Downarrow^\# = \mathit{gfp}(\mathcal{F}^\#)$$

$$\mathcal{F}^\#(\Downarrow_0^\#) = \left\{ (t, \sigma, r) \mid \begin{array}{l} \forall i. t = l_i \Rightarrow \mathit{cond}_i(\sigma) \Rightarrow \\ (t, \sigma, r) \in \mathit{apply}_i^\#(\Downarrow_0^\#) \end{array} \right\}$$

$if\ x (r := 0) (r := x)$

Analysing in $\{x \mapsto +\}$

- Only the rule IF_{TRUE} applies.
- We get $r \mapsto 0$.

Analysing in $\{x \mapsto \top\}$

- Both rules IF_{TRUE} and IF_{FALSE} apply.
- We get $r \mapsto 0$ from IF_{TRUE} .
- We get $r \mapsto \top$ from IF_{FALSE} .
- We get $r \mapsto \top$ at the end.

```

CoInductive aeval : term -> ast -> ares -> Prop :=
  | aeval_cons : forall t sigma r,
    (forall n,
      t = left n ->
      acond n sigma ->
      aapply n sigma r) ->
    aeval t sigma r
with aapply : name -> ast -> ares -> Prop :=
  | aapply_cons : forall n sigma sigma' r r',
    sigma  $\sqsubseteq$  sigma' ->
    r'  $\sqsubseteq$  r ->
    aapply_step n sigma' r' ->
    aapply n sigma r

```

```

with aapply_step : name -> ast -> ares -> Prop :=
| aapply_step_Ax : forall n ax sigma r,
  rule_struct n = Rule_struct_Ax _ ->
  arule n = Rule_Ax ax ->
  ax sigma = Some r ->
  aapply_step n sigma r
| aapply_step_R1 : forall n t up sigma sigma' r,
  rule_struct n = Rule_struct_R1 t ->
  arule n = Rule_R1 _ up ->
  up sigma = Some sigma' ->
  aeval t sigma' r ->
  aapply_step n sigma r
| aapply_step_R2 : forall n t1 t2 up next
  sigma sigma1 sigma2 r r',
  rule_struct n = Rule_struct_R2 t1 t2 ->
  arule n = Rule_R2 up next ->
  up sigma = Some sigma1 ->
  aeval t1 sigma1 r ->
  next sigma r = Some sigma2 ->
  aeval t2 sigma2 r' ->
  aapply_step n sigma r'.

```

- 1 Motivation
- 2 Pretty-Big-Step: a Generic Rule Format
- 3 Defining an Abstract Semantics Correct by Construction
- 4 Running Abstract Interpreters