

Compilation and Program Analysis (#2a): Semantics

Ludovic Henrio

Master 1, ENS de Lyon et Dpt Info, Lyon1

2025-2026



Intro

Contact me:

web: lhenrio.github.io

email: ludovic.henrio@ens-lyon.fr

Credits: JC Filliâtre / JC Fernandez / Nielson-Nielson-Hankin /
Laure Gonnord

Note on organisation:

1: Course

2: **exercises and proofs during the course** ;

3: **exercises and proofs done at the end the course if we
have the time**

- 1 Generalities on semantics
- 2 Operational semantics for mini-while
- 3 Comparing the different semantics

Semantics

We will first define an abstract syntax for our language.

Example : arithmetic expressions, $x \in V$ a set of variables

$$e ::= x \mid n \mid e + e \mid e * e \mid \dots$$

This is just another view of the AST obtained after parsing.

On the abstract syntax we will define one or several semantics.

Different kinds of semantics:

- axiomatic
- denotational
- by translation
- **operational semantics (natural, structural)**

Axiomatic Semantics (Hoare logic)

(*An axiomatic basis for computer programming*, 1969)

Characterisation by properties on variables, using triples of the form:

$$\{P\} i \{Q\}$$

“if P is true before the instruction i , then Q is true afterwards”

Example :

$$\{x \geq 0\} x := x + 1 \{x > 0\}$$

Example of generating rule:

$$\{P[x \leftarrow E]\} x := E \{P(x)\}$$

► proving properties of programs.

Denotational Semantics

Associates to an expression e its mathematical meaning $\llbracket e \rrbracket$ that represents its computation in a mathematical domain \mathcal{D} .

Example : arithmetic expressions, $x \in V$ a set of variables

$$e ::= x \mid n \mid e + e \mid e * e \mid \dots$$

You must choose a domain for the mathematical meaning with adequate operations.

Trivial example for expressions with $\mathcal{D} = \text{env} \rightarrow \mathbb{N}$.

$$\llbracket x \rrbracket \rho = \rho(x)$$

$$\llbracket n \rrbracket \rho = \mathcal{N}(n)$$

$$\llbracket e_1 + e_2 \rrbracket \rho = \llbracket e_1 \rrbracket \rho + \llbracket e_2 \rrbracket \rho$$

$$\llbracket e_1 * e_2 \rrbracket \rho = \llbracket e_1 \rrbracket \rho \times \llbracket e_2 \rrbracket \rho$$

Semantics by translation

(Definitional interpreters for higher-order programming languages, Reynolds, 1972)

We can define the semantics of a language by translation into a language whose semantics is already known.

$$\begin{aligned} \llbracket x = v + v' \rrbracket &= \begin{array}{l} y = \text{get } v; \\ z = \text{get } v'; \\ x = y + z \end{array} \end{aligned}$$

- ▶ Inherit for free the meta-theory from the host language.
- ▶ Not always very illuminating in terms of behaviour
- ▶ ... but sometimes a good specification in terms of implementation or compilation.

Operational Semantics

Describes the computation as an evaluation from the program to its computed value. Operates directly on the abstract syntax. 2 kinds:

- “natural” or “*big-steps semantics*”, evaluates the program in one step

$$e \Downarrow v$$

- “by reduction” or “*small-steps semantics*”, repeat the evaluation until a result is obtained:

$$e \rightarrow e_1 \rightarrow e_2 \rightarrow \dots \rightarrow v$$

A relation describes an atomic reduction, and the semantics consider the transitive (reflexive) closure of this relation.

Results do not need to be a value.

Note: different notations (arrows) exist: \Downarrow / \Rightarrow / $\dots \vdash \dots \rightarrow \dots$

- language specification and proving properties of languages.

- 1 Generalities on semantics
- 2 Operational semantics for mini-while
- 3 Comparing the different semantics

mini-while

$$e \in \mathcal{A} ::= x \mid n \mid e + e \mid e * e \mid \dots$$

(abstract) grammar:

$S(\text{Smt}) ::=$	$x := e$	assign
	$ \text{ skip}$	do nothing
	$ S_1; S_2$	sequence
	$ \text{ if } b \text{ then } S_1 \text{ else } S_2$	test
	$ \text{ while } b \text{ do } S \text{ done}$	loop

Semantics of expressions

We denote $State = Var \rightarrow \mathbf{Z}$.

This kind of state is sometimes called “store”. We denote them by σ .

Access is denoted $\sigma(x)$. Update is denoted by $\sigma[y \mapsto n]$.

Semantics of arithmetic expressions – Val: $\mathcal{A} \rightarrow State \rightarrow \mathbf{Z}$ (in each state an integer value): **On board**

$$Val(n, \sigma) = \mathcal{N}(n)$$

$$Val(x, \sigma) =$$

$$Val(e + e', \sigma) =$$

$$Val(e \times e', \sigma) =$$

Note: what kind of semantics is this? big step? denotational?

Semantics of boolean expressions

$Val : \mathcal{B} \rightarrow State \rightarrow \mathbf{Z}$ **Not now**

$(b ::= tt \mid ff \mid x \mid b \wedge b \mid \dots \mid e < e \mid \dots)$

First properties and exercise

Semantics of arithmetic expressions

Substitution is denoted $e[e'/x]$.

Show the two following properties (first one at the end of the course):

- 1 Let $e \in \mathcal{A}$ a given arithmetic expression. Let σ, σ' be two states. Show that if $(\forall x \in Vars(e), \sigma(x) = \sigma'(x))$, then $Val(e, \sigma) = Val(e, \sigma')$. **At the end of course?**
- 2 Let $e, e' \in \mathcal{A}$, show that:

$$Val(e[e'/x], \sigma) = Val(e, \sigma[x \mapsto Val(e', \sigma)])$$

now

Natural semantics (big step) for mini-while 1/2

In one step from the source program to the final result.

$\Downarrow: Stm \times State \rightarrow State$

$$(x := e, \sigma) \Downarrow \sigma[x \mapsto Val(e, \sigma)]$$

$$(skip, \sigma) \Downarrow \sigma$$

$$\frac{(S_1, \sigma) \Downarrow \sigma' \quad (S_2, \sigma') \Downarrow \sigma''}{((S_1; S_2), \sigma) \Downarrow \sigma''}$$

Natural semantics (big step) for mini-while 2/2

$$\frac{Val(b, \sigma) = tt \quad (S_1, \sigma) \Downarrow \sigma'}{(if\ b\ then\ S_1\ else\ S_2, \sigma) \Downarrow \sigma'}$$

$$\frac{Val(b, \sigma) = ff \quad (S_2, \sigma) \Downarrow \sigma'}{(if\ b\ then\ S_1\ else\ S_2, \sigma) \Downarrow \sigma'}$$

$$\frac{Val(b, \sigma) = tt \quad (S, \sigma) \Downarrow \sigma' \quad (while\ b\ do\ S\ done, \sigma') \Downarrow \sigma''}{(while\ b\ do\ S\ done, \sigma) \Downarrow \sigma''}$$

$$\frac{Val(b, \sigma) = ff}{(while\ b\ do\ S\ done, \sigma) \Downarrow \sigma}$$

Example

Compute the semantics (leaves are axioms, nodes are rules) of:

- $x := 2; \text{while } x > 0 \text{ do } x := x - 1 \text{ done}$
- $x := 2; \text{while } x > 0 \text{ do } x := x + 1 \text{ done}$

Using the semantics to prove properties

Example: determinism

In mini-while there is a single way to evaluate a program.

Theorem: Determinism

For all S , for all $\sigma, \sigma', \sigma''$:

- If $(S, \sigma) \Downarrow \sigma'$ and $(S, \sigma) \Downarrow \sigma''$ then $\sigma' = \sigma''$.
- If $(S, \sigma) \Downarrow \sigma'$, there is no infinite derivation.

The Proof is by induction on the structure of the derivation tree.

WE do a proof sketch

Structural Op. Semantics (SOS = small step) for mini-while 1/2

Evaluating one statement at a time.

$\Rightarrow: Stm \times State \rightarrow Stm \times State$ OR $Stm \times State \rightarrow State$ (we could have a **done** statement to avoid the two cases).

$$(x := e, \sigma) \rightarrow \sigma[x \mapsto Val(e, \sigma)] \quad (\text{skip}, \sigma) \rightarrow \sigma$$

$$\frac{(S_1, \sigma) \rightarrow \sigma'}{((S_1; S_2), \sigma) \rightarrow (S_2, \sigma')}$$

$$\frac{(S_1, \sigma) \rightarrow (S'_1, \sigma')}{((S_1; S_2), \sigma) \rightarrow (S'_1; S_2, \sigma')}$$

Structural Op. Semantics (SOS = small step) for mini-while 2/2

$$\frac{Val(b, \sigma) = tt}{(\text{if } b \text{ then } S_1 \text{ else } S_2, \sigma) \rightarrow (S_1, \sigma)}$$

$$\frac{Val(b, \sigma) = ff}{(\text{if } b \text{ then } S_1 \text{ else } S_2, \sigma) \rightarrow (S_2, \sigma)}$$

$$(\text{while } b \text{ do } S \text{ done}, \sigma) \rightarrow$$

$$(\text{if } b \text{ then } (S; \text{while } b \text{ do } S \text{ done}) \text{ else skip}, \sigma)$$

Exercises

Compute the semantics of:

- $x := 2; \text{while } x > 0 \text{ do } x := x - 1 \text{ done}$
- $x := 2; \text{while } x > 0 \text{ do } x := x + 1 \text{ done}$

How to prove determinism for the SOS semantics? What is the structure of the proof? **do the proof**

- 1 Generalities on semantics
- 2 Operational semantics for mini-while
- 3 Comparing the different semantics

Comparison: divergence

In general a program diverges if it runs forever.

In mini-while, a program diverges in state σ iff:

- NAT: no successor to (S, σ) .
- SOS: infinite sequence beginning with (S, σ) .

In other languages/semantics there might be other reasons to have no successor (see later in course), and you could have no successor in the SOS without reaching a final state.

Comparison: equivalence of programs

Semantics is also useful for defining program equivalence, in mini-while it is quite simple:

Two mini-while programs S_1 and S_2 are semantically equivalent if:

- NAT: $\forall \sigma, \sigma', (S_1, \sigma) \Downarrow \sigma' \text{ iff } (S_2, \sigma) \Downarrow \sigma'$
- SOS: $\forall \sigma$:
 - for all config (blocking or not): $(S_1, \sigma) \rightarrow^* \sigma' \text{ iff } (S_2, \sigma) \rightarrow^* \sigma'$
 - (S_1, σ) diverges iff (S_2, σ) diverges

Are the two semantics equivalent?

$$\mathcal{S}_{NS}[S]\sigma = \begin{cases} \sigma' & \text{If } (S, \sigma) \Downarrow \sigma' \\ \text{undef} & \text{else} \end{cases}$$

$$\mathcal{S}_{SOS}[S]\sigma = \begin{cases} \sigma' & \text{If } (S, \sigma) \Rightarrow^* \sigma' \\ \text{undef} & \text{else} \end{cases}$$

Theorem

$$\mathcal{S}_{NS} = \mathcal{S}_{SOS}$$

Proof: see next slides ...

Equivalence of semantics 1/2

Proposition

If $(S, \sigma) \Downarrow \sigma'$ then $(S, \sigma) \rightarrow^* \sigma'$.

Proof relies on:

Lemma

If $(S_1, \sigma) \rightarrow^k \sigma'$ then $((S_1; S_2), \sigma) \rightarrow^k (S_2, \sigma')$

Proof: structural induction on the derivation tree for $(S, \sigma) \Downarrow$.

Equivalence of semantics 2/2

Proposition

If $(S, \sigma) \rightarrow^k \sigma'$ then $(S, \sigma) \Downarrow \sigma'$.

Proof relies on:

Lemma

If $(S_1; S_2, \sigma) \rightarrow^k \sigma''$ then there exists σ', k_1 such that $(S_1, \sigma) \rightarrow^{k_1} \sigma'$ and $(S_2, \sigma') \rightarrow^{k-k_1} \sigma''$

Proof: induction on k .

Expressing parallelism

SOS can express interleaving, NAT cannot:

$$\frac{(S_1, \sigma) \rightarrow (S'_1, \sigma')}{((S_1 || S_2), \sigma) \rightarrow (S'_1 || S_2, \sigma')} \quad \frac{(S_2, \sigma) \rightarrow (S'_2, \sigma')}{((S_1 || S_2), \sigma) \rightarrow (S_1 || S'_2, \sigma')}$$

... more later in the course.

Mini-while is not exactly mini-C

variable initialisation!

- **variable declarations**

- Main problem is scope of variables (x may not refer to the same variable depending on the point in the program)
- see course on typing

- Expression **evaluation**

restricted to expressions without side-effect, the val function has to be encoded as a set of instructions (a more precise semantics would define several reduction steps)

- **print-int and print-string** (operational semantics not much interesting)
- Mini-C will have **functions** ... defined later in the course

Conclusion

We have seen different kinds of semantics and compared them briefly.

We have shown how to define operational semantics.

- For expression evaluation
- On mini-while

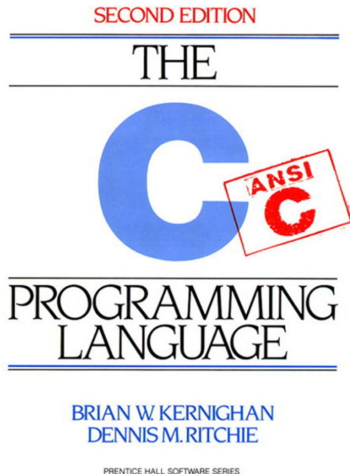
And how to reason on them to derive language properties (or at least properties of the semantics).

Next course on typing will illustrate more properties.

Additional exercise: **repeat**.

Final words: Different degrees of precision

Semi-formal specification in natural language



Final words: Different degrees of precision

Formal semantics

(store)	s	\models	$\{inst\ inst^*,\ tab\ tabinst^*,\ mem\ meminst^*\}$
(instances)	$inst$	\models	$\{func\ clob,\ glob\ v^*,\ tab\ i^*,\ mem\ i^*\}$
	$tabinst$	\models	cl^*
	$meminst$	\models	b^*
(closures)	cl	\models	$\{inst\ i,\ code\ f\}$ (where f is not an import and has all exports ex^* erased)
(values)	v	\models	i^*
(administrative operators)	e	\models	$\dots, [trap\ call\ cl\ label_{i_0}(e^*)\ e^*\ end\ local_{i_1}(i_1\ v^*)\ e^*\ end\$
(local contexts)	L^k	\models	$e^* [i]\ e^*$
	L^{k+1}	\models	$v^* label_{i_0}(e^*)\ L^k\ end\ e^*$

Reduction	$s; v^*; e^* \mapsto_{\alpha_1} s'; v^*; e'^*$	$s; v^*; e^* \mapsto_{\alpha_2} s'; v^*; e'^*$	$s; v^*; e^* \mapsto_{\alpha_3} s'; v^*; e'^*$
	$s; v^*; L^k[e^*] \mapsto_{\alpha_1} s'; v^*; L^k[e'^*]$	$s; v^*_1; local_{i_0}(i_1\ v^*)\ e^* \end{array} \mapsto_{\alpha_2} s'; v^*_1; local_{i_0}(i_1\ v^*)\ e'^* \end{array}$	
	$L^k[trap]$	\hookrightarrow	$trap$ if $L^k \neq []$
	$(t.\ const\ c_1)\ t.\ unop\ c_2$	\hookrightarrow	$t.\ const\ unop_{i_0}(c_2)$
	$(t.\ const\ c_1)\ (t.\ const\ c_2)\ t.\ binop\ c_3$	\hookrightarrow	$t.\ const\ c$ if $c = binop_{i_0}(c_1, c_2)$
	$(t.\ const\ c_1)\ (t.\ const\ c_2)\ t.\ binop\ c_3$	\hookrightarrow	$trap$ otherwise
	$(t.\ const\ c_1)\ (t.\ const\ c_2)\ t.\ testop\ c_3$	\hookrightarrow	$i32.\ const\ testop_{i_0}(c_3)$
	$(t.\ const\ c_1)\ (t.\ const\ c_2)\ t.\ relop\ c_3$	\hookrightarrow	$i32.\ const\ relop_{i_0}(c_1, c_2)$
	$(t_1.\ const\ c)\ t_2.\ convert\ t_1.\ as^*$	\hookrightarrow	$t_2.\ const\ c'$ if $c' = conv_{i_0}^{as^*}(c)$
	$(t_1.\ const\ c)\ t_2.\ convert\ t_1.\ as^*$	\hookrightarrow	$trap$ otherwise
	$(t_1.\ const\ c)\ t_2.\ reinterpret\ t_1$	\hookrightarrow	$t_2.\ const\ const_{i_0}(bit_{i_0}(c))$
	unreachable	\hookrightarrow	$trap$
	nop	\hookrightarrow	c
	$e.\ drop$	\hookrightarrow	e
	$v_1\ v_2\ (i32.\ const\ 0)\ select$	\hookrightarrow	v_2
	$v_1\ v_2\ (i32.\ const\ k + 1)\ select$	\hookrightarrow	v_1
	$v^* block\ [i_0^* \rightarrow i_1^*]\ v^*\ end$	\hookrightarrow	$label_{i_0}(i_1\ v^*)\ v^*\ end$
	$v^* loop\ [i_0^* \rightarrow i_1^*]\ v^*\ end$	\hookrightarrow	$label_{i_0}(loop\ [i_0^* \rightarrow i_1^*]\ v^*\ end)\ v^*\ end$
	$(i32.\ const\ 0)\ if\ v^*\ else\ v_2^*\ end$	\hookrightarrow	$block\ if\ v_2^*\ end$
	$(i32.\ const\ k + 1)\ if\ v^*\ else\ v_2^*\ end$	\hookrightarrow	$block\ if\ v_1^*\ end$
	$label_{i_0}(e^*)\ v^*\ end$	\hookrightarrow	v^*
	$label_{i_0}(e^*)\ trap\ end$	\hookrightarrow	$trap$
	$label_{i_0}(e^*)\ L^k[v^*]\ (br\ j)\ end$	\hookrightarrow	$v^*\ e^*$
	$(i32.\ const\ 0)\ (br\ if\ j)$	\hookrightarrow	c
	$(i32.\ const\ k + 1)\ (br\ if\ j)$	\hookrightarrow	$br\ j$
	$(i32.\ const\ k)\ (br\ table\ j_0^*\ j_1^*)$	\hookrightarrow	$br\ j$
	$(i32.\ const\ k + n)\ (br\ table\ j_0^*\ j_1^*)$	\hookrightarrow	$br\ j$
	$s;\ call\ j$	$\hookrightarrow_{\alpha_1}$	$call\ s_{i_0}(i, j)$
	$s;\ (i32.\ const\ j)\ call\ indirect\ if$	$\hookrightarrow_{\alpha_1}$	$call\ s_{i_0}(i, j)$
	$s;\ (i32.\ const\ j)\ call\ indirect\ if$	$\hookrightarrow_{\alpha_1}$	$trap$ if $s_{i_0}(i, j)_{i_0} = (func\ if\ local\ i^*\ e^*)$
	$v^*\ (call\ cl)$	\hookrightarrow	$local_{i_0}(cl_{i_0}; v^*\ (t.\ const\ 0)^k)\ block\ (e \rightarrow i_0^*)\ v^*\ end\ end\ \dots$
	$local_{i_0}(i_1\ v^*)\ v^*\ end$	\hookrightarrow	v^* if $e_{i_0} = (func\ (i_1^* \rightarrow i_0^*)\ local\ i^k\ e^*)$
	$local_{i_0}(i_1\ v^*)\ trap\ end$	\hookrightarrow	$trap$
	$local_{i_0}(i_1\ v^*)\ return\ end$	\hookrightarrow	v^*
	$v_1^*\ v_2^*\ get\ local\ j$	\hookrightarrow	v
	$v_1^*\ v_2^*\ set\ local\ j$	\hookrightarrow	$v_1^*\ v_2^*\ i_0^*\ e$
	$v;\ (tee\ local\ j)$	\hookrightarrow	$v;\ (v\ tee\ local\ j)$
	$s;\ get\ global\ j$	$\hookrightarrow_{\alpha_1}$	$g_{i_0}(i, j)$
	$s;\ v\ (set\ global\ j)$	$\hookrightarrow_{\alpha_1}$	$s'; e$
	$s;\ (i32.\ const\ k)\ (t.\ load\ a)\ o$	$\hookrightarrow_{\alpha_1}$	$t.\ const\ const_{i_0}(b^*)$
	$s;\ (i32.\ const\ k)\ (t.\ load\ tp.\ as^*)\ a\ o$	$\hookrightarrow_{\alpha_1}$	$t.\ const\ const_{i_0}^{as^*}(b^*)$
	$s;\ (i32.\ const\ k)\ (t.\ load\ tp.\ as^*)\ a\ o$	$\hookrightarrow_{\alpha_1}$	$trap$
	$s;\ (i32.\ const\ k)\ (t.\ const\ c)\ (t.\ store\ a)\ o$	$\hookrightarrow_{\alpha_1}$	$s'; e$
	$s;\ (i32.\ const\ k)\ (t.\ const\ c)\ (t.\ store\ tp.\ a)\ o$	$\hookrightarrow_{\alpha_1}$	$s'; e$
	$s;\ (i32.\ const\ k)\ (t.\ const\ c)\ (t.\ store\ tp.\ a)\ o$	$\hookrightarrow_{\alpha_1}$	$trap$
	$s;\ current\ memory$	$\hookrightarrow_{\alpha_1}$	$i32.\ const\ [s_{i_0}(i, *)] / 64\ Ki$
	$s;\ (i32.\ const\ k)\ grow\ memory$	$\hookrightarrow_{\alpha_1}$	$s'; i32.\ const\ [s_{i_0}(i, *)] / 64\ Ki$ if $s' = s$ with $mem(i, *) = s_{i_0}(i, *)\ (0)^{64\ Ki}$
	$s;\ (i32.\ const\ k)\ grow\ memory$	$\hookrightarrow_{\alpha_1}$	$i32.\ const\ (-1)$

Figure 2. Small-step reduction rules

Final words: Different degrees of precision

Mechanized formal semantics in a proof assistant

```
(** One step of execution *)

Inductive step: state -> trace -> state -> Prop :=

| step_skip_seq: forall f s k sp e m,
  step (State f Sskip (Kseq s k) sp e m)
  E0 (State f s k sp e m)

| step_skip_block: forall f k sp e m,
  step (State f Sskip (Kblock k) sp e m)
  E0 (State f Sskip k sp e m)

| step_skip_call: forall f k sp e m m',
  is_call_cont k ->
  Mem.free m sp 0 f.(fn_stackspace) = Some m' ->
  step (State f Sskip k (Vptr sp Ptrofs.zero) e m)
  E0 (Returnstate Vundef k m')

| step_assign: forall f id a k sp e m v,
  eval_expr sp e m a v ->
  step (State f (Sassign id a) k sp e m)
  E0 (State f Sskip k sp (PTree.set id v e) m)

| step_store: forall f chunk addr a k sp e m vaddr v m',
  eval_expr sp e m addr vaddr ->
  eval_expr sp e m a v ->
  Mem.storev chunk m vaddr v = Some m' ->
  step (State f (Sstore chunk addr a) k sp e m)
  E0 (State f Sskip k sp e m')

| step_call: forall f optid sig a bl k sp e m vf vargs fd,
  eval_expr sp e m a vf ->
  eval_explist sp e m bl vargs ->
  Genv.find_funct ge vf = Some fd ->
  funsig fd = sig ->
  step (State f (Scall optid sig a bl) k sp e m)
```