15-411: Dynamic Semantics

Jan Hoffmann

Code review.
Sign up today.

Dynamic Semantics

- **Static semantics:** definition of valid programs
- **Dynamic semantics:** definition of how programs are executed
- So far: Dynamic semantics is given in English on lab handouts
  - This only works since you know how C programs should behave
  - Sometimes needed to consult the reference compiler
- A description in English will always be ambiguous

➡ **Need precise ways of defining the meaning of programs**

Types of (Formal) Dynamic Semantics

- **Denotational Semantics:** Abstract and elegant. Dana Scott
  - Each part of a program is associated with a denotation (math. object)
  - For example: a procedure is associated with a mathematical function
- **Axiomatic Semantics:** Strongly related to program logic. Tony Hoare
  - Gives meaning to phrases using logical axioms
  - The meaning is identical to the set of properties that can be proved
- **Operational Semantics:** Describes how programs are executed
  - Related to interpreters and abstract machines Gordon Plotkin
  - Most popular and flexible form of semantics

Operational Semantics

- **Many different styles**
  - Natural semantics (or big-step semantics or evaluation dynamics)
  - Structural operational semantics
  - Substructural operational semantics
  - Abstract machine (or small-step with continuation)
- **We will use an abstract machine**
  - Very general: can describe non-termination, concurrency, …
  - Low-level and elaborate

Frank Pfenning

How to pick the right dynamic semantics?

Evaluating Expressions

Continuations

Want to model a single evaluation step

$$e \to e'$$

For example: $((4 + 5) * 10 + 2) \to (9 * 10 + 2)$

How can we find the right place at which to make the step?

Use a continuation K:

A stack of partial computations.

$$e \triangleright K$$

“Evaluate expression e and pass the result to K”

The continuation has a ‘hole’ for the result value of e.

Evaluation Rules: Addition

First evaluate e1.

Plug the result here.

$$e_1 + e_2 \rhd K \quad \longrightarrow \quad e_1 \rhd (\_ + e_2\ ,\ K)$$

A constant.

Continue with evaluating e2.

Plug the result here.

$$c_1 \rhd (\_ + e_2\ ,\ K) \quad \longrightarrow \quad e_2 \rhd (c_1 + \_\ ,\ K)$$

Continuation is an addition.

$$c_2 \rhd (c_1 + \_\ ,\ K) \quad \longrightarrow \quad c \rhd K \qquad (c = c_1 + c_2 \bmod 2^{32})$$

Two constants

Actual addition.

Evaluation Rules: Binary Operations

Arithmetic operations are treated like addition

$$e_1 \oplus e_2 \triangleright K \quad \longrightarrow \quad e_1 \triangleright (\_ \oplus e_2\ ,\ K)$$

$$c_1 \triangleright (\_ \oplus e_2\ ,\ K) \quad \longrightarrow \quad e_2 \triangleright (c_1 \oplus \_\ ,\ K)$$

$$c_2 \triangleright (c_1 \oplus \_\ ,\ K) \quad \longrightarrow \quad c \triangleright K \qquad (c = c_1 \oplus c_2 \bmod 2^{32})$$

Arithmetic is modulo 2^{32} to match our x86 architecture

What about effects?

Evaluation Rules: Binops with Effects

In case of an arithmetic exception: Abort the computation and report and error

$$
\begin{array}{llll}
e_1 \oslash e_2 \rhd K & \longrightarrow & e_1 \rhd (\_ \oslash e_2\ ,\ K) & \\
c_1 \rhd (\_ \oslash e_2\ ,\ K) & \longrightarrow & e_2 \rhd (c_1 \oslash \_\ ,\ K) & \\
c_2 \rhd (c_1 \oslash \_\ ,\ K) & \longrightarrow & c \rhd K & (c = c_1 \oslash c_2) \\
c_2 \rhd (c_1 \oslash \_\ ,\ K) & \longrightarrow & \mathsf{exception}(\mathsf{arith}) & (c_1 \oslash c_2 \text{ undefined})
\end{array}
$$

There is no rule for further evaluating an exception.

Example Evaluation

$$((4 + 5) * 10) + 2 \quad \triangleright \quad \cdot$$

Evaluation Rules: End of and Evaluation

If we reach a constant and the empty continuation then we stop

$$c \rhd \cdot \quad \longrightarrow \quad \mathsf{value}(c)$$

Evaluation Rules: Boolean Expressions

$$e_1 \;\texttt{\&\&}\; e_2 \rhd K \quad\longrightarrow\quad e_1 \rhd (\_ \;\texttt{\&\&}\; e_2\;,\; K)$$

$$\mathsf{false} \rhd (\_ \;\texttt{\&\&}\; e_2\;,\; K) \quad\longrightarrow\quad \mathsf{false} \rhd K$$

Notice the short-cutting.

$$\mathsf{true} \rhd (\_ \;\texttt{\&\&}\; e_2\;,\; K) \quad\longrightarrow\quad e_2 \rhd K$$

true and *false* are also values

(We could also use 1 and 0 but distinguishing helps detect errors.)

Variables and Environments

How do we evaluate variable?

$$x \triangleright K \quad \longrightarrow \quad ?$$

Integers or booleans.

Need to have an environment that maps variables to values

$$\eta ::= \cdot \mid \eta, x \mapsto v$$

The machine state consists now of an expression, a continuation, and an environment

$$\eta \vdash e \triangleright K$$

Variables and Environments II

The rules we have seen so far just carry over

$$\eta \vdash e_1 \oplus e_2 \triangleright K \quad \longrightarrow \quad \eta \vdash e_1 \triangleright (\_ \oplus e_2\ ,\ K)$$

$$\eta \vdash c_1 \triangleright (\_ \oplus e_2\ ,\ K) \quad \longrightarrow \quad \eta \vdash e_2 \triangleright (c_1 \oplus \_\ ,\ K)$$

$$\eta \vdash c_2 \triangleright (c_1 \oplus \_\ ,\ K) \quad \longrightarrow \quad \eta \vdash c \triangleright K \qquad (c = c_1 \oplus c_2 \bmod 2^{32})$$

Variables are simply looked up

$$\eta \vdash x \triangleright K \quad \longrightarrow \quad \eta \vdash \eta(x) \triangleright K$$

The environment never changes when evaluating expressions

Executing Statements

Executing Statements I

Executions of statements don't pass values to the continuation

Statements have usually an effect on the environment

Machine configurations:

$$\eta \vdash s \blacktriangleright K$$

Continuations contain statements.

Sequences:

$$\eta \vdash \mathsf{seq}(s_1, s_2) \blacktriangleright K \quad \longrightarrow \quad \eta \vdash s_1 \blacktriangleright (s_2 \ , \ K)$$

No ops:

$$\eta \vdash \mathsf{nop} \blacktriangleright (s \ , \ K) \quad \longrightarrow \quad \eta \vdash s \blacktriangleright K$$

A terminating execution ends with a nop.

Executing Statements II

Interaction with expressions is straightforward

Assignments:

$$\eta \vdash \mathsf{assign}(x, e) \blacktriangleright K \quad \longrightarrow \quad \eta \vdash e \rhd (\mathsf{assign}(x, \_)\ ,\ K)$$

$$\eta \vdash v \rhd (\mathsf{assign}(x, \_)\ ,\ K) \quad \longrightarrow$$

Update the environment with new mapping.

Executing Statements III

Conditionals:

$$\eta \vdash \mathsf{if}(e, s_1, s_2) \blacktriangleright K \qquad\qquad \longrightarrow$$

Executing Statements IV

Loops:

$$\eta \vdash \mathsf{while}(e, s) \blacktriangleright K \quad \longrightarrow \quad ?$$

Not that the following statements are equivalent:

$$\mathsf{while}(e, s) \equiv \mathsf{if}(e, \mathsf{seq}(s, \mathsf{while}(e, s)), \mathsf{nop})$$

$$\eta \vdash \mathsf{while}(e, s) \blacktriangleright K \quad \longrightarrow \quad \eta \vdash \mathsf{if}(e, \mathsf{seq}(s, \mathsf{while}(e, s)), \mathsf{nop}) \blacktriangleright K$$

Non-termination:

$$s_0 \longrightarrow s_1 \longrightarrow s_2 \longrightarrow \cdots$$

We can make an infinite number of steps without reaching a final state

Executing Statements V

Assertions:

$$\eta \vdash \mathsf{assert}(e) \blacktriangleright K \quad \longrightarrow \quad \eta \vdash e \triangleright (\mathsf{assert}(\_), K)$$

$$\eta \vdash \mathsf{true} \triangleright (\mathsf{assert}(\_), K) \quad \longrightarrow \quad \eta \vdash \mathsf{nop} \blacktriangleright K$$

$$\eta \vdash \mathsf{false} \triangleright (\mathsf{assert}(\_), K) \quad \longrightarrow \quad \mathsf{exception}(\mathsf{abort})$$

Declarations:

$$\eta \vdash \mathsf{decl}(x, \tau, s) \blacktriangleright K \quad \longrightarrow \quad \eta[x \mapsto \mathsf{nothing}] \vdash s \blacktriangleright K$$

If C0 had shadowing then we would have to be careful here.

Final states:

$$\mathsf{exception}(E), \qquad \mathsf{nop} \blacktriangleright \cdot$$

Example: Infinite Loop

$\mathsf{while}(x > 0, \mathsf{assign}(x, x+1))$ $\eta = [x \mapsto 1]$ $\mathsf{s} = \mathsf{assign}(x, x+1)$

$\longrightarrow$
$\longrightarrow$
$\longrightarrow$
$\longrightarrow$
$\longrightarrow$
$\longrightarrow$
$\longrightarrow$
$\longrightarrow$
$\longrightarrow$
$\longrightarrow$
$\longrightarrow$
$\longrightarrow$
$\longrightarrow$
$\longrightarrow$
$\longrightarrow$

Functions

Function Calls

What needs to happen at a function call?

- Evaluate the arguments in left-to-right order
- Save the environment of the caller to continue the execution after the function call
- Save the continuation of the caller
- Execute the body of the callee in a new environment that maps the formal parameters to the argument values
- Pass the return value to the environment of the caller

Call Stack

We need to keep track of continuations and environment in stack frames

Call stack:

Continuation

$$S \quad ::= \quad \cdot \mid S \; , \langle \eta, K \rangle$$

Environment

Configurations:

Evaluation $\qquad S \; ; \eta \vdash e \rhd K$

Execution $\qquad S \; ; \eta \vdash s \blacktriangleright K$

Existing rules can be lifted to the new configurations by passing through the call stack

Rules for Function Calls

n args is similar.

We only show the special case of 0 and 2 arguments

Evaluate s in empty environment.

No arguments:

$$S\ ;\eta \vdash f(\,) \rhd K \quad \longrightarrow \quad (S\ ,\ \langle \eta, K \rangle)\ ;\cdot \vdash s \blacktriangleright \cdot$$

(given that f is defined as $f(\,)\{s\}$*)*

Store callee's stack frame

Two arguments:

$$S\ ;\eta \vdash f(e_1, e_2) \rhd K \quad \longrightarrow \quad S\ ;\eta \vdash e_1 \rhd (f(\_, e_2)\ ,\ K)$$

$$S\ ;\eta \vdash c_1 \rhd (f(\_, e_2)\ ,\ K) \quad \longrightarrow \quad S\ ;\eta \vdash e_2 \rhd (f(c_1, \_)\ ,\ K)$$

$$S\ ;\eta \vdash c_2 \rhd (f(c_1, \_)\ ,\ K) \quad \longrightarrow \quad (S\ ,\ \langle \eta, K \rangle)\ ;[x_1 \mapsto c_1, x_2 \mapsto c_2] \vdash s \blacktriangleright \cdot$$

(given that f is defined as $f(x_1, x_2)\{s\}$*)*

Rules for Returns

Need to restore continuation and environment and pass return value

$$S \; ; \eta \vdash \mathsf{return}(e) \blacktriangleright K$$

$$S \;, \langle \eta', K' \rangle \; ; \eta \vdash v \triangleright (\mathsf{return}(\_) \;, K)$$

Special case: returning void

Will only be reached by functions without return.

$$S \;, \langle \eta', K' \rangle \; ; \eta \vdash \mathsf{nop} \blacktriangleright \cdot \quad \longrightarrow \quad S \; ; \eta' \vdash \mathsf{nothing} \triangleright K'$$

Dummy value

Alternative: elaborate each function that returns void with return(nothing) statements.

Execution of the Main Function

How can we execute a program?

$$\cdot\ ;\ \cdot \vdash \mathsf{main}(\,) \rhd \cdot \qquad \text{(initial state)}$$

$$\cdot\ ;\ \eta \vdash c \rhd \cdot \quad \longrightarrow \quad \mathsf{value}(c) \qquad \text{(final state)}$$

Statics, Dynamics, and Safety

Overview of Machine States (Configurations)

$S\,;\eta \vdash e \rhd K$ – Evaluating the expression e with the continuation K

$S\,;\eta \vdash s \blacktriangleright K$ – Evaluating the statement s with the continuation K

$\mathsf{value}(c)$ – Final state, return a value

$\mathsf{exception}(E)$ – Final state, report an error

What do we expect from the transitions?

There shouldn't be more steps after reaching a final state

The language should be deterministic: there at most one transition per state

Progress

Well-typed programs don't go wrong!

There are many non-final states that don't have transitions, e.g.

$S; \eta \vdash 42 \rhd (\mathsf{if}(\_, s_1, s_2); K)$ $\qquad$ $\cdot; \cdot \vdash \mathsf{nop} \blacktriangleright \cdot$

Stuck states.

The behavior of these states is undefined.

Central relationship between static and dynamic semantics:

Programs that are well-defined according to the static semantics should be free of undefined behavior.

Theorem 1 (No undefined behavior) *If a program passes all the static semantics, and*

$$\cdot; \cdot \vdash \mathsf{main}() \longrightarrow \mathcal{ST}_1 \longrightarrow \ldots \longrightarrow \mathcal{ST}_n$$

then either $\mathcal{ST}_n$ *is a final state or else* $\mathcal{ST}_n$ *is not-stuck because there exists a state* $\mathcal{ST}'$ *such that* $\mathcal{ST}_n \longrightarrow \mathcal{ST}'$.

How to prove this?

15-312

All ops.

Expressions	e	$::=$	$c \mid e_1 \odot e_2 \mid \mathsf{true} \mid \mathsf{false} \mid e_1 \;\&\&\; e_2 \mid x \mid f(e_1, e_2) \mid f(\,)$
Statements	s	$::=$	$\mathsf{nop} \mid \mathsf{seq}(s_1, s_2) \mid \mathsf{assign}(x, e) \mid \mathsf{decl}(x, \tau, s)$
		$\mid$	$\mathsf{if}(e, s_1, s_2) \mid \mathsf{while}(e, s) \mid \mathsf{return}(e) \mid \mathsf{assert}(e)$
Values	v	$::=$	$c \mid \mathsf{true} \mid \mathsf{false} \mid \mathsf{nothing}$
Environments	η	$::=$	$\cdot \mid \eta, x \mapsto c$
Stacks	S	$::=$	$\cdot \mid S\ ,\ \langle \eta, K \rangle$
Cont. frames	ϕ	$::=$	$\_ \odot e \mid c \odot \_ \mid \_ \;\&\&\; e \mid f(\_, e) \mid f(c, \_)$
		$\mid$	$s \mid \mathsf{assign}(x, \_) \mid \mathsf{if}(\_, s_1, s_2) \mid \mathsf{return}(\_) \mid \mathsf{assert}(\_)$
Continuations	K	$::=$	$\cdot \mid \phi\ ,\ K$
Exceptions	E	$::=$	$\mathsf{arith} \mid \mathsf{abort} \mid \mathsf{mem}$

Summary I

$$
\begin{array}{lcl}
S\ ;\eta \vdash e_1 \odot e_2 \rhd K & \longrightarrow & S\ ;\eta \vdash e_1 \rhd (\_ \odot e_2\ ,\ K) \\
S\ ;\eta \vdash c_1 \rhd (\_ \odot e_2\ ,\ K) & \longrightarrow & S\ ;\eta \vdash e_2 \rhd (c_1 \odot \_\ ,\ K) \\
S\ ;\eta \vdash c_2 \rhd (c_1 \odot \_\ ,\ K) & \longrightarrow & S\ ;\eta \vdash c \rhd K \qquad (c = c_1 \odot c_2) \\
S\ ;\eta \vdash c_2 \rhd (c_1 \odot \_\ ,\ K) & \longrightarrow & \mathsf{exception}(\mathsf{arith}) \qquad (c_1 \odot c_2 \text{ undefined}) \\
\\
S\ ;\eta \vdash e_1 \ \texttt{\&\&}\ e_2 \rhd K & \longrightarrow & S\ ;\eta \vdash e_1 \rhd (\_ \ \texttt{\&\&}\ e_2\ ,\ K) \\
S\ ;\eta \vdash \mathsf{false} \rhd (\_ \ \texttt{\&\&}\ e_2\ ,\ K) & \longrightarrow & S\ ;\eta \vdash \mathsf{false} \rhd K \\
S\ ;\eta \vdash \mathsf{true} \rhd (\_ \ \texttt{\&\&}\ e_2\ ,\ K) & \longrightarrow & S\ ;\eta \vdash e_2 \rhd K \\
\\
S\ ;\eta \vdash x \rhd K & \longrightarrow & S\ ;\eta \vdash \eta(x) \rhd K
\end{array}
$$

Summary: Expressions

$$S \; ; \eta \vdash \mathsf{seq}(s_1, s_2) \blacktriangleright K \quad \longrightarrow \quad S \; ; \eta \vdash s_1 \blacktriangleright (s_2 \; , K)$$
$$S \; ; \eta \vdash \mathsf{nop} \blacktriangleright (s \; , K) \quad \longrightarrow \quad S \; ; \eta \vdash s \blacktriangleright K$$
$$S \; ; \eta \vdash \mathsf{assign}(x, e) \blacktriangleright K \quad \longrightarrow \quad S \; ; \eta \vdash e \triangleright (\mathsf{assign}(x, \_) \; , K)$$
$$S \; ; \eta \vdash c \triangleright (\mathsf{assign}(x, \_) \; , K) \quad \longrightarrow \quad S \; ; \eta[x \mapsto c] \vdash \mathsf{nop} \blacktriangleright K$$

$$S \; ; \eta \vdash \mathsf{decl}(x, \tau, s) \blacktriangleright K \quad \longrightarrow \quad S \; ; \eta[x \mapsto \mathsf{nothing}] \vdash s \blacktriangleright K$$

$$S \; ; \eta \vdash \mathsf{assert}(e) \blacktriangleright K \quad \longrightarrow \quad S \; ; \eta \vdash e \triangleright (\mathsf{assert}(\_) \; , K)$$
$$S \; ; \eta \vdash \mathsf{true} \triangleright (\mathsf{assert}(\_) \; , K) \quad \longrightarrow \quad S \; ; \eta \vdash \mathsf{nop} \blacktriangleright K$$
$$S \; ; \eta \vdash \mathsf{false} \triangleright (\mathsf{assert}(\_) \; , K) \quad \longrightarrow \quad \mathsf{exception}(\mathsf{abort})$$

$$S \; ; \eta \vdash \mathsf{if}(e, s_1, s_2) \blacktriangleright K \quad \longrightarrow \quad S \; ; \eta \vdash e \triangleright (\mathsf{if}(\_, s_1, s_2) \; , K)$$
$$S \; ; \eta \vdash \mathsf{true} \triangleright (\mathsf{if}(\_, s_1, s_2), K) \quad \longrightarrow \quad S \; ; \eta \vdash s_1 \blacktriangleright K$$
$$S \; ; \eta \vdash \mathsf{false} \triangleright (\mathsf{if}(\_, s_1, s_2), K) \quad \longrightarrow \quad S \; ; \eta \vdash s_2 \blacktriangleright K$$

$$S \; ; \eta \vdash \mathsf{while}(e, s) \blacktriangleright K \quad \longrightarrow \quad S \; ; \eta \vdash \mathsf{if}(e, \mathsf{seq}(s, \mathsf{while}(e, s)), \mathsf{nop}) \blacktriangleright K$$

Summary: Statements

$$S\ ;\ \eta \vdash f(e_1, e_2) \rhd K \quad \longrightarrow \quad S\ ;\ \eta \vdash e_1 \rhd (f(\_, e_2)\ ,\ K)$$

$$S\ ;\ \eta \vdash c_1 \rhd (f(\_, e_2)\ ,\ K) \quad \longrightarrow \quad S\ ;\ \eta \vdash e_2 \rhd (f(c_1, \_)\ ,\ K)$$

$$S\ ;\ \eta \vdash c_2 \rhd (f(c_1, \_)\ ,\ K) \quad \longrightarrow \quad (S\ ,\ \langle \eta, K \rangle)\ ;\ [x_1 \mapsto c_1, x_2 \mapsto c_2] \vdash s \blacktriangleright \cdot$$

(given that f *is defined as* $f(x_1, x_2)\{s\}$*)*

$$S\ ;\ \eta \vdash f(\,) \rhd K \quad \longrightarrow \quad (S\ ,\ \langle \eta, K \rangle)\ ;\ \cdot \vdash s \blacktriangleright \cdot$$

(given that f *is defined as* $f(\,)\{s\}$*)*

$$S\ ;\ \eta \vdash \mathsf{return}(e) \blacktriangleright K \quad \longrightarrow \quad S\ ;\ \eta \vdash e \rhd (\mathsf{return}(\_)\ ,\ K)$$

$$(S\ ,\ \langle \eta', K' \rangle)\ ;\ \eta \vdash v \rhd (\mathsf{return}(\_)\ ,\ K) \quad \longrightarrow \quad S\ ;\ \eta' \vdash v \rhd K'$$

$$\cdot\ ;\ \eta \vdash c \rhd (\mathsf{return}(\_)\ ,\ K) \quad \longrightarrow \quad \mathsf{value}(c)$$

Summary: Functions