Principles of Programming Languages Lectures 6: Big-step Structural Operational Semantics

Andrei Arusoaie¹

¹Department of Computer Science

November 2, 2021

Outline

Structural Operational Semantics

Big-step SOS

Configurations
Arithmetic expressions
Boolean expressions
Statements

Structural Operational Semantics

A language designer should understand the existing design approaches:

- Big-step structural operational semantics
- Small-step structural operational semantics
- Denotational Semantics
- Modular operational semantics
- Reduction semantics with evaluation contexts
- Abstract Machines, the chemical abstract machine
- Axiomatic semantics

Slide credits: prof. Grigore Rosu, UIUC



Structural Operational Semantics

A language designer should understand the existing design approaches:

- Big-step structural operational semantics
- Small-step structural operational semantics
- Denotational Semantics
- Modular operational semantics
- Reduction semantics with evaluation contexts
- Abstract Machines, the chemical abstract machine
- Axiomatic semantics

Structural Operational Semantics (SOS)

- SOS was proposed by Plotkin in 1981
- Simple framework to describe the behaviour of the language constructs as inference rules
- Rules specify transitions between program configurations
- Configurations are tuples of various kinds of data structures (e.g., trees, sets, lists)
 - capture various components of program states (environment, memory, stacks, registers, etc.)

Big-step Structural Operational Semantics

- Probably the most natural way of defining structural operational semantics
- A.k.a. natural semantics, relational semantics, evaluation semantics
- ▶ Big-step SOS for a PL: a set of inference rules

IMP

The language that we use to illustrate SOS is IMP. It includes:

- Arithmetic expressions
- Boolean Expressions
- Statements: assignments, (possibly empty)
 blocks/sequences of statements, decisionals (if-then-else),
 loops (while)

Outline

Structural Operational Semantics

Big-step SOS

Configurations

Arithmetic expressions Boolean expressions Statements

Big-step SOS configurations

Configurations are tuples: $\langle code, environment, ... \rangle$.

They hold various information needed to execute a program:

- the code that needs to be executed:
- the current environment;
- the program stack;
- the current program counter;
- **...**

Big-step SOS configurations

Configurations are tuples: $\langle code, environment, ... \rangle$.

They hold various information needed to execute a program:

- the code that needs to be executed;
- the current environment;
- the program stack;
- the current program counter;
- **.**..

IMP configurations

For instance, IMP configurations can store the following information:

- \triangleright $\langle A, E \rangle$, where A has type AExp and E has type Env;
- ► ⟨N⟩, where N has type nat;
- \triangleright $\langle B, E \rangle$, where B has type BExp and E has type Env;
- $\triangleright \langle B \rangle$, where A has type bool;
- $\triangleright \langle E \rangle$, where E has type Env;
- \triangleright $\langle S, E \rangle$, where S has type Stmt and E has type Env;
- \triangleright $\langle S \rangle$, where S has type Stmt.

Big-step SOS sequents: relations over configurations of the form $C \downarrow R$, where

- C is a configuration
- ▶ *R* is a *result configuration* that is obtained after the complete evaluation of *C*.

Result configurations are also called irreducible configurations.

Informally, a sequent says that a configuration *C* evaluates/executes/transitions to a configuration *R*.



Big-step SOS sequents: relations over configurations of the form $C \downarrow R$, where

- C is a configuration
- ▶ *R* is a *result configuration* that is obtained after the complete evaluation of *C*.

Result configurations are also called *irreducible configurations*.

Informally, a sequent says that a configuration *C* evaluates/executes/transitions to a configuration *R*.



Big-step SOS sequents: relations over configurations of the form $C \Downarrow R$, where

- C is a configuration
- R is a result configuration that is obtained after the complete evaluation of C.

Result configurations are also called *irreducible configurations*.

Informally, a sequent says that a configuration *C* evaluates/executes/transitions to a configuration *R*.



Big-step SOS sequents: relations over configurations of the form $C \Downarrow R$, where

- C is a configuration
- R is a result configuration that is obtained after the complete evaluation of C.

Result configurations are also called *irreducible configurations*.

Informally, a sequent says that a configuration *C* evaluates/executes/transitions to a configuration *R*.



Big-step SOS sequents: relations over configurations of the form $C \downarrow R$, where

- C is a configuration
- R is a result configuration that is obtained after the complete evaluation of C.

Result configurations are also called *irreducible configurations*.

Informally, a sequent says that a configuration *C* evaluates/executes/transitions to a configuration *R*.



Big-step SOS sequents: relations over configurations of the form $C \downarrow R$, where

- C is a configuration
- R is a result configuration that is obtained after the complete evaluation of C.

Result configurations are also called *irreducible configurations*.

Informally, a sequent says that a configuration *C evaluates/executes/transitions* to a configuration *R*.



Rule schemata

The *Big-step SOS rules* have this form:

$$\frac{C_1 \Downarrow R_1 \qquad C_2 \Downarrow R_2 \qquad \cdots \qquad C_n \Downarrow R_n}{C_0 \Downarrow R_0},$$

where C_0, C_1, \ldots, C_n are configurations and R_0, R_1, \ldots, R_n are result configurations.

Here is an example of a big-step rule for IMP:

$$\frac{\langle a_1, \, \sigma \rangle \Downarrow \langle i_1 \rangle \quad \langle a_2, \, \sigma \rangle \Downarrow \langle i_2 \rangle}{\langle a_1 + a_2, \, \sigma \rangle \Downarrow \langle i_{1+nat} i_{2} \rangle}$$

Rule schemata

The *Big-step SOS rules* have this form:

$$\frac{C_1 \Downarrow R_1 \qquad C_2 \Downarrow R_2 \qquad \cdots \qquad C_n \Downarrow R_n}{C_0 \Downarrow R_0},$$

where C_0, C_1, \ldots, C_n are configurations and R_0, R_1, \ldots, R_n are result configurations.

Here is an example of a big-step rule for IMP:

$$\frac{\langle a_1, \, \sigma \rangle \Downarrow \langle i_1 \rangle \quad \langle a_2, \, \sigma \rangle \Downarrow \langle i_2 \rangle}{\langle a_1 + a_2, \, \sigma \rangle \Downarrow \langle i_1 +_{nat} i_2 \rangle}$$

Outline

Structural Operational Semantics

Big-step SOS

Configurations

Arithmetic expressions

Boolean expressions

Statements

BIGSTEP-CONST:
$$\langle i, \sigma \rangle \Downarrow \langle i \rangle$$

BIGSTEP-LOOKUP: $\langle x, \sigma \rangle \Downarrow \langle \sigma(x) \rangle$ if $\sigma(x) \neq \bot$

BIGSTEP-ADD: $\frac{\langle a_1, \sigma \rangle \Downarrow \langle i_1 \rangle}{\langle a_1 + a_2, \sigma \rangle \Downarrow \langle i_1 + nat i_2 \rangle}$

BIGSTEP-MUL: $\frac{\langle a_1, \sigma \rangle \Downarrow \langle i_1 \rangle}{\langle a_1^* a_2, \sigma \rangle \Downarrow \langle i_1^* nat i_2 \rangle}$

BIGSTEP-CONST: $\langle i, \sigma \rangle \Downarrow \langle i \rangle$

BIGSTEP-LOOKUP: $\langle x,\,\sigma\rangle\Downarrow\langle\sigma(x)\rangle$ if $\sigma(x)\neq \bot$

BIGSTEP-ADD: $\frac{\langle a_1, \sigma \rangle \psi \langle i_1 \rangle \qquad \langle a_2, \sigma \rangle \psi \langle i_2 \rangle}{\langle a_1 + a_2, \sigma \rangle \psi \langle i_1 + nat i_2 \rangle}$

BIGSTEP-MUL: $\frac{\langle a_1, \sigma \rangle \psi \langle l_1 \rangle}{\langle a_1^* a_2, \sigma \rangle \psi \langle i_2 \rangle} \langle a_2, \sigma \rangle \psi \langle i_2 \rangle$

BIGSTEP-CONST: $\langle i, \sigma \rangle \Downarrow \langle i \rangle$

 $\mathsf{BIGSTEP\text{-}LOOKUP:} \qquad \qquad \langle x,\,\sigma\rangle \Downarrow \langle \sigma(x)\rangle \qquad \qquad \mathsf{if}\,\,\sigma(x) \neq \bot$

BIGSTEP-ADD: $\frac{\langle a_1, \sigma \rangle \psi \langle i_1 \rangle \quad \langle a_2, \sigma \rangle \psi \langle i_2 \rangle}{\langle a_1 + a_2, \sigma \rangle \psi \langle i_1 + n_{at} i_2 \rangle}$

BIGSTEP-MUL: $\frac{\langle a_1,\sigma\rangle \psi\langle i_1\rangle \qquad \langle a_2,\sigma\rangle \psi\langle i_2\rangle}{\langle a_1^*a_2,\sigma\rangle \psi\langle i_1^*_{nat}i_2\rangle}$

BIGSTEP-CONST: $\langle i, \sigma \rangle \Downarrow \langle i \rangle$

 $\mathsf{BIGSTEP\text{-}LOOKUP:} \qquad \qquad \langle x,\,\sigma\rangle \Downarrow \langle \sigma(x)\rangle \qquad \qquad \mathsf{if} \ \sigma(x) \neq \bot$

BIGSTEP-ADD: $\frac{\langle a_1, \sigma \rangle \psi \langle i_1 \rangle \langle a_2, \sigma \rangle \psi \langle i_2 \rangle}{\langle a_1 + a_2, \sigma \rangle \psi \langle i_1 + n_{at} i_2 \rangle}$

BIGSTEP-MUL: $\frac{\langle a_1, \sigma \rangle \psi \langle i_1 \rangle \quad \langle a_2, \sigma \rangle \psi \langle i_2 \rangle}{\langle a_1^* a_2, \sigma \rangle \psi \langle i_1^*_{nat} i_2 \rangle}$

Derivation

▶ A derivation example where $\sigma(n) = 10$:

$$\frac{\frac{\cdot}{\langle 2,\sigma\rangle \Downarrow \langle 2\rangle} \text{ BigStep-Const} \quad \frac{\cdot}{\langle n,\sigma\rangle \Downarrow \langle 10\rangle}}{\langle 2+n,\sigma\rangle \Downarrow \langle 2+_{nat}10\rangle} \text{ BigStep-Lookup}} \text{ BigStep-Add}$$

rule schema vs. rule?

The following *rule*:

$$\frac{\langle 2, \sigma \rangle \Downarrow \langle 2 \rangle \quad \langle n, \sigma \rangle \Downarrow \langle 10 \rangle}{\langle 2 +' \quad n, \sigma \rangle \Downarrow \langle 2 +_{nat} 10 \rangle}$$

is a an instance of the rule schema:

$$\frac{\langle a_1,\,\sigma\rangle \Downarrow \langle i_1\rangle \quad \langle a_2,\,\sigma\rangle \Downarrow \langle i_2\rangle}{\langle a_1 + \prime \quad a_2,\,\sigma\rangle \Downarrow \langle i_1 +_{nat} i_2\rangle} \text{ BIGSTEP-ADD}.$$

On the other hand, this is not an instance of BIGSTEP-ADD

$$\frac{\langle 2, \sigma \rangle \Downarrow \langle 2 \rangle \quad \langle n, \sigma \rangle \Downarrow \langle 10 \rangle}{\langle 2 +' \quad n, \sigma \rangle \Downarrow \langle 2 +_{nat} 10, \sigma \rangle} ?$$

because $\langle 2 +' n, \sigma \rangle \Downarrow \langle 2 +_{nat} 10, \sigma \rangle$ is not an instance of the conclusion of BIGSTEP-ADD.



rule schema vs. rule?

The following *rule*:

$$\frac{\langle 2, \sigma \rangle \Downarrow \langle 2 \rangle \quad \langle n, \sigma \rangle \Downarrow \langle 10 \rangle}{\langle 2 + ' \quad n, \sigma \rangle \Downarrow \langle 2 +_{nat} 10 \rangle}$$

is a an instance of the rule schema:

$$\frac{\langle a_1, \, \sigma \rangle \Downarrow \langle i_1 \rangle \quad \langle a_2, \, \sigma \rangle \Downarrow \langle i_2 \rangle}{\langle a_1 + a_2, \, \sigma \rangle \Downarrow \langle i_1 +_{nat} i_2 \rangle} \text{ BIGSTEP-ADD.}$$

On the other hand, this is not an instance of BIGSTEP-ADD:

$$\frac{\langle 2, \sigma \rangle \Downarrow \langle 2 \rangle \quad \langle n, \sigma \rangle \Downarrow \langle 10 \rangle}{\langle 2 +' \quad n, \sigma \rangle \Downarrow \langle 2 +_{nat} 10, \frac{\sigma}{\sigma} \rangle} ?$$

because $\langle 2 +' n, \sigma \rangle \Downarrow \langle 2 +_{nat} 10, \sigma \rangle$ is not an instance of the conclusion of BIGSTEP-ADD.

Outline

Structural Operational Semantics

Big-step SOS

Configurations

Arithmetic expressions

Boolean expressions

Statements

Big-step rules for IMP – 2: boolean expressions

 $\mathsf{BigStep\text{-}True} : \qquad \langle \mathsf{btrue}, \sigma \rangle \Downarrow \langle \mathit{true} \rangle$

 ${\tt BigStep-False:} \qquad \langle {\tt bfalse}, \sigma \rangle \Downarrow \langle \textit{false} \rangle$

 $\frac{\langle b, \sigma \rangle \Downarrow \langle true \rangle}{\langle b, \sigma \rangle \Downarrow \langle falso \rangle}$

BIGSTEP-NOTTRUE: $\overline{\langle\,!\,b,\,\sigma\rangle\Downarrow\langle\mathit{false}\rangle}$

 $\langle b, \sigma \rangle \Downarrow \langle false \rangle$

BIGSTEP-NOTFALSE: $\overline{\langle\,!\,b,\,\sigma\rangle\,\Downarrow\langle\mathit{true}\rangle}$

Big-step rules for IMP – 2: boolean expressions

$$\frac{\langle b_1, \, \sigma \rangle \Downarrow \langle \textit{false} \rangle}{\langle b_1 \, \text{ and } b_2, \, \sigma \rangle \Downarrow \langle \textit{false} \rangle}$$

BIGSTEP-LT:
$$\frac{\langle a_1, \, \sigma \rangle \Downarrow \langle i_1 \rangle \quad \langle a_2, \, \sigma \rangle \Downarrow \langle i_2 \rangle}{\langle a_1 \, < \, a_2, \, \sigma \rangle \Downarrow \langle i_1 <_{\textit{nat}} \, i_2 \rangle}$$

BIGSTEP-GT:
$$\frac{\langle a_1, \, \sigma \rangle \Downarrow \langle i_1 \rangle \quad \langle a_2, \, \sigma \rangle \Downarrow \langle i_2 \rangle}{\langle a_1 \rangle \langle a_2, \, \sigma \rangle \Downarrow \langle i_1 \rangle_{nat} \langle i_2 \rangle}$$

Derivation example

A derivation tree for $\langle 2 + n < 10, \sigma \rangle \Downarrow \langle (2 +_{nat} 10) <_{nat} 10 \rangle$:

$$\begin{array}{c|c} \frac{\cdot}{\langle 2,\sigma\rangle \Downarrow \langle 2\rangle} \text{ Const} & \frac{\cdot}{\langle n,\sigma\rangle \Downarrow \langle 10\rangle} \text{ Lookup} \\ \frac{\langle 2+n,\sigma\rangle \Downarrow \langle 2+_{\textit{nat}} 10\rangle}{\langle 2+n < 10,\sigma_1\rangle \Downarrow \langle (2+_{\textit{nat}} 10) <_{\textit{nat}} 10\rangle} \text{ Const} \\ \hline \end{pmatrix} \text{ LT}$$

Outline

Structural Operational Semantics

Big-step SOS

Configurations Arithmetic expressions Boolean expressions

Statements

Big-step rules for IMP – 3: statements

BIGSTEP-ASSIGN:
$$\frac{\langle a,\,\sigma\rangle \Downarrow \langle i\rangle}{\langle x\,\,::=\,\,a,\,\sigma\rangle \Downarrow \langle \sigma[i/x]\rangle} \quad \textit{if } \sigma(x) \neq \bot$$

BIGSTEP-SEQ:
$$\frac{\langle s_1, \, \sigma \rangle \Downarrow \langle \sigma_1 \rangle \quad \langle s_2, \, \sigma_1 \rangle \Downarrow \langle \sigma_2 \rangle}{\langle s_1 \; ; ; \; s_2, \, \sigma \rangle \Downarrow \langle \sigma_2 \rangle}$$

BIGSTEP-SKIP:
$$\langle \text{skip}, \sigma \rangle \Downarrow \langle \sigma \rangle$$

$$\frac{\langle \textbf{b}, \sigma \rangle \Downarrow \langle \textit{false} \rangle}{\langle \texttt{while} \, \textbf{b} \, \textbf{s}, \, \sigma \rangle \Downarrow \langle \sigma \rangle}$$

$$\frac{\langle b,\,\sigma\rangle \Downarrow \langle \textit{true}\rangle \qquad \langle \textit{s}\;\;\textit{;}\;\; \text{while}\, \textit{b}\, \textit{s},\,\sigma\rangle \Downarrow \langle \sigma'\rangle}{\langle \text{while}\, \textit{b}\, \textit{s},\,\sigma\rangle \Downarrow \langle \sigma'\rangle}$$

Exercise

Write a derivation for the sequent:

$$\langle i::=0;$$
; while $i<1$ $(i::=i+1), \sigma \rangle \Downarrow \langle (\sigma[0/i])[1/i] \rangle$.

The derivation is shown in the next slides

Derivation

$$\frac{\frac{\cdot}{\langle i : :=0,\sigma\rangle \Downarrow \langle \sigma[0/i]\rangle} \text{ ASSIGN }}{\langle i : :=0; \text{; while } i < 1 \text{ } (i : :=i+1),\sigma\rangle \Downarrow \langle (\sigma[0/i])[1/i]\rangle} \text{ SEQ}$$

Derivation for A

$$\frac{\frac{\cdot}{\langle i,\sigma[0/i]\rangle \Downarrow \langle 0\rangle} \text{ LOOKUP} \qquad \frac{\cdot}{\langle 1,\sigma[0/i]\rangle \Downarrow \langle 1\rangle} \text{ Const}}{\frac{\langle i<1,\sigma[0/i]\rangle \Downarrow \langle \textit{true}\rangle}{\langle \text{while } i<1 \text{ } (i::=i+1),\sigma[0/i]\rangle \Downarrow \langle (\sigma[0/i])[1/i]\rangle}} \text{ WhileTrue}$$

Derivation for B

$$\frac{\frac{\cdot}{\langle i,\sigma[0/i]\rangle \Downarrow \langle 0\rangle} \text{ LOOKUP}}{\frac{\langle i,\sigma[0/i]\rangle \Downarrow \langle 1\rangle}{\langle i::=i+1\rangle \Downarrow \langle (\sigma[0/i])[1/i]\rangle}} \frac{\cdot}{\text{ADD}} \\ \frac{\frac{\langle i+1\rangle \Downarrow \langle 1\rangle}{\langle i::=i+1\rangle \Downarrow \langle (\sigma[0/i])[1/i]\rangle}}{\frac{\langle i::=i+1\rangle \Downarrow \langle (\sigma[0/i])[1/i]\rangle}{\langle i::=i+1;; \text{ while } i<2 \ (i::=i+1), \sigma[0/i]\rangle \Downarrow \langle (\sigma[0/i])[1/i]\rangle}} \text{ SEQ}$$

Derivation for C

```
\frac{\frac{\cdot}{\langle i, (\sigma[0/i])[1/i]\rangle \Downarrow \langle 1\rangle} \text{ LOOKUP}}{\frac{\langle i, (\sigma[0/i])[1/i]\rangle \Downarrow \langle 1\rangle}{\langle i < '1, (\sigma[0/i])[1/i]\rangle \Downarrow \langle false\rangle}} \frac{\cdot}{\text{LT}} \text{ LT}} \frac{\langle i < '1, (\sigma[0/i])[1/i]\rangle \Downarrow \langle false\rangle}{\langle \text{while } i < 1 \ (i::=i+1), (\sigma[0/i])[1/i]\rangle \Downarrow \langle (\sigma[0/i])[1/i]\rangle} \text{ WhileFalse}}
```

Complete derivation

It is a real challenge to read the assembled derivation:

			$\langle i, \sigma[0/i] \rangle \Downarrow \langle 0 \rangle$	$\langle 1, \sigma[0/i] \rangle \Downarrow \langle 1 \rangle$	$\langle i, (\sigma[0/i])[1/i] \rangle \Downarrow \langle 1 \rangle$	$\langle 1, (\sigma[0/i])[1/i] \rangle \downarrow \langle 1 \rangle$
			⟨i+'1⟩ ↓ ⟨1⟩		$\langle i < 1, (\sigma[0/i])[1/i] \rangle \Downarrow \langle false \rangle$	
	$\langle i, \sigma[0/i] \rangle \Downarrow \langle 0 \rangle$	$\langle 1, \sigma[0/i] \rangle \Downarrow \langle 1 \rangle$	(i::=i+'1) ↓ ⟨	$(\sigma[0/i])[1/i])$	(while $i < 1$ ($i := i + 1$), ($\sigma[0]$	$(i])[1/i]$ \Downarrow $\langle (\sigma[0/i])[1/i] \rangle$
	$(i < 1, \sigma[0/i]) \Downarrow (true)$		$(i\colon:=i+'1;;while\;i<'2\;\;(i\colon:=i+'1),\sigma[0/i])\;\Downarrow\;\langle(\sigma[0/i])[1/i]\rangle$			
$\langle i::=0, \sigma \rangle \Downarrow \langle \sigma[0/i] \rangle$	(while $i < 1$ $(i := i + 1)$, $\sigma[0/i] \rangle \Downarrow \langle (\sigma[0/i])[1/i] \rangle$					

 $(i\colon :=0\,;\,; \text{while } i<'1\ (i\colon :=i+'1)\,,\,\sigma)\ \Downarrow\ ((\sigma[0/i])[1/i])$

Bibliography

Chapter Simple Imperative Programs, Section: Evaluation as relation in Software Foundations - Volume 1, Benjamin C. Pierce, Arthur Azevedo de Amorim, Chris Casinghino, Marco Gaboardi, Michael Greenberg, Cătălin Hriţcu, Vilhelm Sjöberg, Andrew Tolmach, Brent Yorgey https://softwarefoundations.cis.upenn.edu/lf-current/Imp.html