CONVENTIONAL SEMANTIC APPROACHES

Grigore Rosu
CS522 – Programming Language Semantics
Conventional Semantic Approaches

A language designer should understand the existing design approaches, techniques and tools, to know what is possible and how, or to come up with better ones. This part of the course will cover the major PL semantic approaches, such as:

- Big-step structural operational semantics (Big-step SOS)
- Small-step structural operational semantics (Small-step SOS)
- Denotational semantics
- Modular structural operational semantics (Modular SOS)
- Reduction semantics with evaluation contexts
- Abstract Machines
- The chemical abstract machine
- Axiomatic semantics
IMP

A simple imperative language
We will exemplify the conventional semantic approaches by means of IMP, a very simple non-procedural imperative language, with

- Arithmetic expressions
- Boolean expressions
- Assignment statements
- Conditional statements
- While loop statements
- Blocks
\textbf{IMP Syntax}

\begin{align*}
\text{Int} & ::= \text{the domain of (unbounded) integer numbers, with usual operations on them} \\
\text{Bool} & ::= \text{the domain of Booleans} \\
\text{Id} & ::= \text{standard identifiers} \\
\text{AExp} & ::= \text{Int} \\
& \quad \text{Id} \\
& \quad \text{AExp} + \text{AExp} \\
& \quad \text{AExp} / \text{AExp} \\
\text{BExp} & ::= \text{Bool} \\
& \quad \text{AExp} \leq \text{AExp} \\
& \quad \neg \text{BExp} \\
& \quad \text{BExp} \&\& \text{BExp} \\
\text{Block} & ::= \{ \} \\
& \quad \{ \text{Stmt} \} \\
\text{Stmt} & ::= \text{Block} \\
& \quad \text{Id} = \text{AExp} ; \\
& \quad \text{Stmt} \text{ Stmt} \\
& \quad \text{if} ( \text{BExp} ) \text{ Block else Block} \\
& \quad \text{while} ( \text{BExp} ) \text{ Block} \\
\text{Pgm} & ::= \text{int List\{Id\} ; Stmt}
\end{align*}

Suppose that, for demonstration purposes, we want “+” and “/” to be non-deterministically strict, “\leq” to be sequentially strict, and “\&\&” to be short-circuited.
Most semantics need some notion of state. A state holds all the semantic ingredients to fully define the meaning of a given program or fragment of program.

For IMP, a state is a partial finite-domain function from identifiers to integers (i.e., a function defined only on a finite subset of identifiers and undefined on the rest), written using a half-arrow:

\[ \sigma : \text{Id} \rightarrow \text{Int} \]

We let State denote the set of such functions, and may write it

\[ \left[ \text{Id} \rightarrow \text{Int} \right]^{\text{finite}} \]

or

\[ \text{Map}\{\text{Id} \mapsto \text{Int}\} \]
We may write states by enumerating each identifier binding. For example, the following state binds $x$ to 8 and $y$ to 0:

$$\sigma = x \mapsto 8, y \mapsto 0$$

Typical state operations are lookup, update and initialization:

- **Lookup**
  $$(-)_\sigma : State \times Id \rightarrow Int$$

- **Update**
  $$(-)[_/_] : State \times Int \times Id \rightarrow State$$

- **Initialization**
  $$\_ \mapsto \_ : List\{Id\} \times Int \rightarrow State$$
BIG-STEP SOS
Big-Step Structural Operational Semantics (Big-Step SOS)

- Gilles Kahn (1987), under the name natural semantics. Also known as relational semantics, or evaluation semantics. We can regard a big-step SOS as a recursive interpreter, telling for a fragment of code and state what it evaluates to.

- **Configuration**: tuple containing code and semantic ingredients
  - E.g., \( \langle a_1, \sigma \rangle \quad \langle a_1 + a_2, \sigma \rangle \quad \langle i_1 \rangle \quad \langle i_1 + \text{Int} \ i_2 \rangle \quad \langle \sigma \rangle \)

- **Sequent**: Pair of configurations, to be derived or proved
  - E.g., \( \langle a_1, \sigma \rangle \downarrow \langle i_1 \rangle \quad \langle a_1 + a_2, \sigma \rangle \downarrow \langle i_1 + \text{Int} \ i_2 \rangle \)

- **Rule**: Tells how to derive a sequent from others
  - E.g.,

\[
\frac{\langle a_1, \sigma \rangle \downarrow \langle i_1 \rangle \quad \langle a_2, \sigma \rangle \downarrow \langle i_2 \rangle}{\langle a_1 + a_2, \sigma \rangle \downarrow \langle i_1 + \text{Int} \ i_2 \rangle}
\]

- Read “evaluates to”
- May omit line when no premises
Big-Step SOS of IMP - Arithmetic

\[ \langle i, \sigma \rangle \Downarrow \langle i \rangle \]

\[ \langle x, \sigma \rangle \Downarrow \langle \sigma(x) \rangle \quad \text{if } \sigma(x) \neq \bot \]

\[ \frac{\langle a_1, \sigma \rangle \Downarrow \langle i_1 \rangle \quad \langle a_2, \sigma \rangle \Downarrow \langle i_2 \rangle}{\langle a_1 + a_2, \sigma \rangle \Downarrow \langle i_1 + \text{Int} \ i_2 \rangle} \]

\[ \frac{\langle a_1, \sigma \rangle \Downarrow \langle i_1 \rangle \quad \langle a_2, \sigma \rangle \Downarrow \langle i_2 \rangle}{\langle a_1 / a_2, \sigma \rangle \Downarrow \langle i_1 / \text{Int} \ i_2 \rangle} \quad \text{if } i_2 \neq 0 \]

**State lookup**

**BIGSTEP-INT**

**BIGSTEP-LOOKUP**

Read: “provided that \( a_1 \) evaluates to \( i_1 \) in \( \sigma \) and \( a_2 \) evaluates to \( i_2 \) in \( \sigma \), then \( a_1 + a_2 \) evaluates to the integer sum of \( i_1 \) and \( i_2 \) in \( \sigma \)

**BIGSTEP-ADD**

**BIGSTEP-DIV**

Side condition ensures rule will never apply when \( a_2 \) evaluates to 0
Big-Step SOS of IMP - Boolean

\[
\langle t, \sigma \rangle \Downarrow \langle t \rangle
\]

\[
\frac{\langle a_1, \sigma \rangle \Downarrow \langle i_1 \rangle \quad \langle a_2, \sigma \rangle \Downarrow \langle i_2 \rangle}{\langle a_1 \leq_{Int} a_2, \sigma \rangle \Downarrow \langle i_1 \leq_{Int} i_2 \rangle}
\]

\[
\frac{\langle b, \sigma \rangle \Downarrow \langle \text{true} \rangle}{\langle \neg b, \sigma \rangle \Downarrow \langle \text{false} \rangle}
\]

\[
\frac{\langle b, \sigma \rangle \Downarrow \langle \text{false} \rangle}{\langle \neg b, \sigma \rangle \Downarrow \langle \text{true} \rangle}
\]

\[
\frac{\langle b_1, \sigma \rangle \Downarrow \langle \text{false} \rangle}{\langle b_1 \&\& b_2, \sigma \rangle \Downarrow \langle \text{false} \rangle}
\]

\[
\frac{\langle b_1, \sigma \rangle \Downarrow \langle \text{true} \rangle \quad \langle b_2, \sigma \rangle \Downarrow \langle t \rangle}{\langle b_1 \&\& b_2, \sigma \rangle \Downarrow \langle t \rangle}
\]
Big-Step SOS of IMP - Statements

\[
\begin{align*}
\langle \{\}, \sigma \rangle & \downarrow \langle \sigma \rangle & \text{(BigStep-Empty-Block)} \\
\langle s, \sigma \rangle & \downarrow \langle \sigma' \rangle & \langle \{ s \}, \sigma \rangle & \downarrow \langle \sigma' \rangle & \text{(BigStep-Block)} \\
\langle a, \sigma \rangle & \downarrow \langle i \rangle & \langle x = a;, \sigma \rangle & \downarrow \langle \sigma[i/x] \rangle & \text{if } \sigma(x) \neq \bot & \text{(BigStep-Asgn)} \\
\langle s_1, \sigma \rangle & \downarrow \langle \sigma_1 \rangle & \langle s_2, \sigma_1 \rangle & \downarrow \langle \sigma_2 \rangle & \langle s_1 \ s_2, \sigma \rangle & \downarrow \langle \sigma_2 \rangle & \text{(BigStep-Seq)} \\
\langle b, \sigma \rangle & \downarrow \langle \text{true} \rangle & \langle s_1, \sigma \rangle & \downarrow \langle \sigma_1 \rangle & \langle \text{if (b) } s_1 \text{ else } s_2, \sigma \rangle & \downarrow \langle \sigma_1 \rangle & \text{(BigStep-If-True)} \\
\langle b, \sigma \rangle & \downarrow \langle \text{false} \rangle & \langle s_2, \sigma \rangle & \downarrow \langle \sigma_2 \rangle & \langle \text{if (b) } s_1 \text{ else } s_2, \sigma \rangle & \downarrow \langle \sigma_2 \rangle & \text{(BigStep-If-False)} \\
\langle b, \sigma \rangle & \downarrow \langle \text{false} \rangle & \langle \text{while (b) } s, \sigma \rangle & \downarrow \langle \sigma \rangle & \text{(BigStep-While-False)} \\
\langle b, \sigma \rangle & \downarrow \langle \text{true} \rangle & \langle s \text{ while (b) } s, \sigma \rangle & \downarrow \langle \sigma' \rangle & \langle \text{while (b) } s, \sigma \rangle & \downarrow \langle \sigma' \rangle & \text{(BigStep-While-True)}
\end{align*}
\]
Big-Step SOS of IMP - Programs

\[
\frac{\langle s, x_l \leftarrow 0 \rangle \Downarrow \langle \sigma \rangle}{\langle \text{int } x_l; s \rangle \Downarrow \langle \sigma \rangle}
\]

(State initialization)
Rules are schemas, allowing recursively enumerable many instances; side conditions filter out instances

- E.g., these are correct instances of the rule for division

\[
\begin{align*}
\langle x, (x \mapsto 8, y \mapsto 0) \rangle & \downarrow \langle 8 \rangle & \langle 2, (x \mapsto 8, y \mapsto 0) \rangle & \downarrow \langle 2 \rangle \\
\langle x/2, (x \mapsto 8, y \mapsto 0) \rangle & \downarrow \langle 4 \rangle \\
\end{align*}
\]

The second may look suspicious, but it is not. Normally, one should never be able to apply it, because one cannot prove its hypotheses

- However, the following is not a correct instance (no matter what \( ? \) is):

\[
\begin{align*}
\langle x, (x \mapsto 8, y \mapsto 0) \rangle & \downarrow \langle 8 \rangle & \langle y, (x \mapsto 8, y \mapsto 0) \rangle & \downarrow \langle 0 \rangle \\
\langle x/y, (x \mapsto 8, y \mapsto 0) \rangle & \downarrow \langle ? \rangle \\
\end{align*}
\]
The following is a valid proof derivation, or proof tree, using the big-step SOS proof system of IMP above.

Suppose that \( x \) and \( y \) are identifiers and \( \sigma(x) = 8 \) and \( \sigma(y) = 0 \).
Big-Step SOS for Type Systems

- Big-Step SOS is routinely used to define type systems for programming languages.
- The idea is that a fragment of code $c$, in a given type environment $\Gamma$, can be assigned a certain type $\tau$. We typically write

$$\Gamma \vdash c : \tau$$

instead of

$$\langle c, \Gamma \rangle \Downarrow \langle \tau \rangle$$

- Since all variables in IMP have integer type, $\Gamma$ can be replaced by a list of untyped variables in our case. In general, however, a type environment $\Gamma$ contains typed variables, that is, pairs “$x : \tau$”.

Typing Arithmetic Expressions

\[
\begin{align*}
xl & \leftarrow i : \text{int} \\
xl & \leftarrow x : \text{int} \quad \text{if } x \in xl \\
xl & \leftarrow a_1 : \text{int} \quad xl & \leftarrow a_2 : \text{int} \\
\frac{xl \leftarrow a_1 + a_2 : \text{int}}{xl \leftarrow a_1 + a_2 : \text{int}} \\
\frac{xl \leftarrow a_1 : \text{int} \quad xl \leftarrow a_2 : \text{int}}{xl \leftarrow a_1 / a_2 : \text{int}} \\
xl & \leftarrow t : \text{bool} \quad \text{if } t \in \{\text{true, false}\}
\end{align*}
\]
Typing Boolean Expressions

\[
\frac{x_l \leftarrow a_1 : \text{int} \quad x_l \leftarrow a_2 : \text{int}}{x_l \leftarrow a_1 \leq a_2 : \text{bool}} \\
\frac{x_l \leftarrow b : \text{bool}}{x_l \leftarrow \neg b : \text{bool}} \\
\frac{x_l \leftarrow b_1 : \text{bool} \quad x_l \leftarrow b_2 : \text{bool}}{x_l \leftarrow b_1 \& \& b_2 : \text{bool}} \\
x_l \leftarrow \{\} : \text{block}
\]

- \(\text{(BigStepTypeSystem-Leq)}\)
- \(\text{(BigStepTypeSystem-Not)}\)
- \(\text{(BigStepTypeSystem-And)}\)
- \(\text{(BigStepTypeSystem-Empty-Block)}\)
Typing Statements

The type of $s$ can be either block or stmt

$$
\frac{xl \vdash s : T}{xl \vdash \{ s \} : \text{block}} \quad \text{if } T \in \{ \text{block}, \text{stmt} \}
$$

$$
\frac{xl \vdash a : \text{int}}{xl \vdash x = a : \text{stmt}} \quad \text{if } x \in xl
$$

$$
\frac{xl \vdash s_1 : \tau_1 \quad xl \vdash s_2 : \tau_2}{xl \vdash s_1 \ s_2 : \text{stmt}} \quad \text{if } \tau_1, \tau_2 \in \{ \text{block}, \text{stmt} \}
$$

$$
\frac{xl \vdash b : \text{bool} \quad xl \vdash s_1 : \text{block} \quad xl \vdash s_2 : \text{block}}{xl \vdash \text{if} \ (b) \ s_1 \ \text{else} \ s_2 : \text{stmt}}
$$

$$
\frac{xl \vdash b : \text{bool} \quad xl \vdash s : \text{block}}{xl \vdash \text{while} \ (b) \ s : \text{stmt}}
$$

(BigStepTypeSystem-Block)

(BigStepTypeSystem-Asgn)

(BigStepTypeSystem-Seq)

(BigStepTypeSystem-If)

(BigStepTypeSystem-While)
Typing Programs

\[
\frac{\text{int } x_1; \; s : \text{pgm}}{\tau \quad \text{if } \tau \in \{\text{block, stmt}\} \quad \text{(BigStepTypeSystem-Pgm)}}
\]
Big-Step SOS Type Derivation

Like the big-step rules for the concrete semantics of IMP, the ones for its type system are also rule schemas. We next show a proof derivation for the well-typed-ness of an IMP program that adds all the numbers from 1 to 100:

\[
\frac{\text{tree}_2}{n,s \vdash (\text{while} \ (\neg (n \leq 0)) \ {s = s + n; \ n = n + -1;}) : \text{stmt}}
\]

\[
\frac{\text{tree}_1 \quad \text{tree}_3}{n,s \vdash (n = 100; \ s = 0; \ \text{while} \ (\neg (n \leq 0)) \ {s = s + n; \ n = n + -1;}) : \text{stmt}}
\]

\[
\vdash (\text{int} \ n,s; \ n = 100; \ s = 0; \ \text{while} \ (\neg (n \leq 0)) \ {s = s + n; \ n = n + -1;}) : \text{pgm}
\]

where
Big-Step SOS Type Derivation

\[ \text{tree}_1 = \begin{cases} \quad \cdot \\\\\\\\\\\cdot \\
, s \vdash 100 : \text{int} & n, s \vdash 0 : \text{int} \\
, s \vdash (n = 100;) : \text{stmt} & n, s \vdash (s = 0;) : \text{stmt} \\
, s \vdash (n = 100; s = 0;) : \text{stmt} \end{cases} \]

\[ \text{tree}_2 = \begin{cases} \quad \cdot \\\\\\\\\\\cdot \\
, s \vdash n : \text{int} & n, s \vdash 0 : \text{int} \\
, s \vdash (n <= 0) : \text{bool} \end{cases} \]

\[ n, s \vdash (! (n <= 0)) : \text{bool} \]
Big-Step SOS Type Derivation

\[ \text{tree}_3 = \{ \begin{array}{c}
\text{n, s} \vdash \text{s : int} \\
\text{n, s} \vdash \text{s + n : int} \\
\text{n, s} \vdash \text{stmt} \\
\text{n, s} \vdash \{ \text{s = s + n; n = n + -1;} \} : \text{block}
\end{array} \} \]
SMALL-STEP SOS

Small-step structural operational semantics
Gordon Plotkin (1981)

Also known as transitional semantics, or reduction semantics

One can regard a small-step SOS as a device capable of executing a program step-by-step

**Configuration:** tuple containing code and semantic ingredients

- E.g., $\langle a, \sigma \rangle$, $\langle b, \sigma \rangle$, $\langle s, \sigma \rangle$, $\langle p \rangle$

**Sequent (transition):** Pair of configurations, to be derived (proved)

- E.g., $\langle a_1, \sigma \rangle \rightarrow \langle a_1', \sigma \rangle$, $\langle a_1 + a_2, \sigma \rangle \rightarrow \langle a_1' + a_2, \sigma \rangle$

**Rule:** Tells how to derive a sequent from others

- E.g.,

\[
\frac{\langle a_1, \sigma \rangle \rightarrow \langle a_1', \sigma \rangle}{\langle a_1 + a_2, \sigma \rangle \rightarrow \langle a_1' + a_2, \sigma \rangle}
\]
Small-Step SOS of IMP - Arithmetic

\[ \langle x, \sigma \rangle \rightarrow \langle \sigma(x), \sigma \rangle \quad \text{if } \sigma(x) \neq \bot \]

\[ \frac{\langle a_1, \sigma \rangle \rightarrow \langle a'_1, \sigma \rangle}{\langle a_1 + a_2, \sigma \rangle \rightarrow \langle a'_1 + a_2, \sigma \rangle} \quad (\text{SMALL\textsc{STEP-ADD-ARG1}}) \]

\[ \frac{\langle a_2, \sigma \rangle \rightarrow \langle a'_2, \sigma \rangle}{\langle a_1 + a_2, \sigma \rangle \rightarrow \langle a_1 + a'_2, \sigma \rangle} \quad (\text{SMALL\textsc{STEP-ADD-ARG2}}) \]

\[ \langle i_1 + i_2, \sigma \rangle \rightarrow \langle i_1 + \text{Int} \ i_2, \sigma \rangle \quad (\text{SMALL\textsc{STEP-ADD}}) \]

\( + \) is non-deterministic (its arguments can evaluate in any order, and interleaved)
Small-Step SOS of IMP - Arithmetic

\[
\begin{align*}
\langle a_1, \sigma \rangle \to \langle a'_1, \sigma \rangle \\
\langle a_1 / a_2, \sigma \rangle \to \langle a'_1 / a_2, \sigma \rangle \\
\langle a_2, \sigma \rangle \to \langle a'_2, \sigma \rangle \\
\langle a_1 / a_2, \sigma \rangle \to \langle a_1 / a'_2, \sigma \rangle \\
\langle i_1 / i_2, \sigma \rangle \to \langle \text{int } i_1 / i_2, \sigma \rangle & \text{ if } i_2 \neq 0
\end{align*}
\]

(SMALLSTEP-DIV-ARG1) 
(SMALLSTEP-DIV-ARG2) 
(SMALLSTEP-DIV)

// is also non-deterministic

Side condition ensures rule will never apply when denominator is 0
Small-Step SOS of IMP - Boolean

\[
\begin{align*}
\langle a_1, \sigma \rangle & \rightarrow \langle a_1', \sigma \rangle \\
\langle a_1 \leq a_2, \sigma \rangle & \rightarrow \langle a_1', \leq a_2, \sigma \rangle \\
\langle a_2, \sigma \rangle & \rightarrow \langle a_2', \sigma \rangle \\
\langle i_1 \leq a_2, \sigma \rangle & \rightarrow \langle i_1 \leq a_2', \sigma \rangle \\
\langle i_1 \leq i_2, \sigma \rangle & \rightarrow \langle i_1 \leq_{int} i_2, \sigma \rangle
\end{align*}
\]

(SMALLSTEP-LEQ-ARG1)

(SMALLSTEP-LEQ-ARG2)

(SMALLSTEP-LEQ)

Ensures sequential strictness
Small-Step SOS of IMP - Boolean

\[
\begin{align*}
\langle b, \sigma \rangle &\rightarrow \langle b', \sigma \rangle \\
\langle ! b, \sigma \rangle &\rightarrow \langle ! b', \sigma \rangle \\
\langle ! \text{true}, \sigma \rangle &\rightarrow \langle \text{false}, \sigma \rangle \\
\langle ! \text{false}, \sigma \rangle &\rightarrow \langle \text{true}, \sigma \rangle \\
\langle b_1, \sigma \rangle &\rightarrow \langle b'_1, \sigma \rangle \\
\langle b_1 \&\& b_2, \sigma \rangle &\rightarrow \langle b'_1 \&\& b_2, \sigma \rangle \\
\langle \text{false} \&\& b_2, \sigma \rangle &\rightarrow \langle \text{false}, \sigma \rangle \\
\langle \text{true} \&\& b_2, \sigma \rangle &\rightarrow \langle b_2, \sigma \rangle
\end{align*}
\]

(SMALLSTEP-NOT-ARG)  
(SMALLSTEP-NOT-TRUE)  
(SMALLSTEP-NOT-FALSE)  
(SMALLSTEP-AND-ARG1)  
(SMALLSTEP-AND-FALSE)  
(SMALLSTEP-AND-TRUE)

Short-circuit semantics
Small-Step SOS Derivation

The following is a valid proof derivation, or proof tree, using the small-step SOS proof system for expressions of IMP above. Suppose that $x$ and $y$ are identifiers and $\sigma(x) = 1$. 

\[
\frac{
\langle x, \sigma \rangle \rightarrow \langle 1, \sigma \rangle}
{\langle y / x, \sigma \rangle \rightarrow \langle y / 1, \sigma \rangle}
\frac{
\langle x + (y / x), \sigma \rangle \rightarrow \langle x + (y / 1), \sigma \rangle}
{\langle (x + (y / x)) \leq x, \sigma \rangle \rightarrow \langle (x + (y / 1)) \leq x, \sigma \rangle}
\]
Small-Step SOS of IMP - Statements

\[
\langle \{ s \}, \sigma \rangle \rightarrow \langle s, \sigma \rangle
\]

\[
\langle a, \sigma \rangle \rightarrow \langle a', \sigma \rangle
\]

\[
\langle x = a ; , \sigma \rangle \rightarrow \langle x = a' ; , \sigma \rangle
\]

\[
\langle x = i ; , \sigma \rangle \rightarrow \langle \{ \}, \sigma[i/x] \rangle \quad \text{if } \sigma(x) \neq \bot
\]

\[
\langle s_1, \sigma \rangle \rightarrow \langle s'_1, \sigma' \rangle
\]

\[
\langle s_1 s_2, \sigma \rangle \rightarrow \langle s'_1 s'_2, \sigma' \rangle
\]

\[
\langle \{ \} s_2, \sigma \rangle \rightarrow \langle s_2, \sigma \rangle
\]

\text{(SMALLSTEP-BLOCK)}

\text{(SMALLSTEP-ASGN-ARG2)}

\text{(SMALLSTEP-ASGN)}

\text{(SMALLSTEP-SEQ-ARG1)}

\text{(SMALLSTEP-SEQ-EMPTY-BLOCK)}
Small-Step SOS of IMP - Statements

\[
\begin{align*}
\langle b, \sigma \rangle &\rightarrow \langle b', \sigma \rangle \\
\frac{\text{if} (b) s_1 \text{ else } s_2, \sigma}{\text{if} (b') s_1 \text{ else } s_2, \sigma} \\
\langle \text{if} (\text{true}) s_1 \text{ else } s_2, \sigma \rangle &\rightarrow \langle s_1, \sigma \rangle \\
\langle \text{if} (\text{false}) s_1 \text{ else } s_2, \sigma \rangle &\rightarrow \langle s_2, \sigma \rangle \\
\langle \text{while} (b) s, \sigma \rangle &\rightarrow \langle \text{if} (b) \{ s \text{ while} (b) s \} \text{ else } \{ \}, \sigma \rangle \\
\langle \text{int } xl; s \rangle &\rightarrow \langle s, xl \leftrightarrow 0 \rangle
\end{align*}
\]

(State initialization)
Small-Step SOS in Rewriting Logic

- Any small-step SOS can be associated a rewrite logic theory (or, equivalently, a Maude module)
- The idea is to associate to each small-step SOS rule a rewrite rule

\[
\begin{align*}
C_1 &\rightarrow C'_1 \\
C_2 &\rightarrow C'_2 \\
\vdots \\
C_n &\rightarrow C'_n
\end{align*}
\]

\[
C \rightarrow C' \quad \text{[if condition]}
\]

a rewrite rule

\[
\bar{o}C \rightarrow \bar{C'} \quad \text{if} \quad \bar{o}C_1 \rightarrow \bar{C'}_1 \wedge \bar{o}C_2 \rightarrow \bar{C'}_2 \wedge \ldots \wedge \bar{o}C_n \rightarrow \bar{C'}_n \quad \text{[\wedge condition]}
\]

(the circle means “ready for one step”)
DENOTATIONAL

Denotational or fixed-point semantics
Denotational Semantics

- Christopher Strachey and Dana Scott (1970)

- Associate *denotation*, or meaning, to (fragments of) programs into *mathematical domains*; for example,
  - The denotation of an arithmetic expression in IMP is a *partial function from states to integer numbers*
    \[
    [-] : \text{AExp} \rightarrow (\text{State} \rightarrow \text{Int})
    \]
  - The denotation of a statement in IMP is a *partial function from states to states*
    \[
    [-] : \text{Stmt} \rightarrow (\text{State} \rightarrow \text{State})
    \]
Denotational Semantics

Christopher Strachey and Dana Scott (1970)

Denotation, or meaning, to (fragments of) programs into mathematical domains; for example,

- The denotation of an arithmetic expression in IMP is a partial function from states to integer numbers:

  \[
  [-] : AExp \rightarrow (State \rightarrow Int)
  \]

- The denotation of a statement in IMP is a partial function from states to states:

  \[
  [-] : Stmt \rightarrow (State \rightarrow State)
  \]

Partial because some expressions may be undefined in some states (e.g., division by zero)

Partial because some statements in some states may use undefined expressions, or may not terminate.
Denotational Semantics - Terminology

Denotation (function)

Fragment of program

Mathematical domain

\([\_]: \text{AExp} \rightarrow (\text{State} \rightarrow \text{Int})\)

\([\_]: \text{Stmt} \rightarrow (\text{State} \rightarrow \text{State})\)
Once the right mathematical domains are chosen, giving a denotational semantics to a language should be a straightforward and *compositional* process; e.g.

\[
[a_1 + a_2]\sigma = [a_1]\sigma +_{\text{Int}} [a_2]\sigma
\]

\[
[a_1 / a_2]\sigma = \begin{cases} 
[a_1]\sigma /_{\text{Int}} [a_2]\sigma & \text{if } [a_2]\sigma \neq 0 \\
\bot & \text{if } [a_2]\sigma = 0
\end{cases}
\]

The hardest part is to give semantics to recursion. This is done using *fixed-points*. 
Mathematical Domains

- Mathematical domains can be anything; it is common though that they are organized as complete partial orders with bottom element.

- The partial order structure captures the intuition of informativeness: \( a \leq b \) means \( a \) is less informative than \( b \). E.g., as a loop is executed, we get more and more information about its semantics.

- Completeness means that chains of more and more informative elements have a limit.

- The bottom element, written \( \bot \), stands for undefined, or no information at all.
Partial Orders

- Partial order \((D, \leq)\) is set \(D\) and binary rel. \(\leq\) which is
  - Reflexive: \(x \leq x\)
  - Transitive: \(x \leq y\) and \(y \leq z\) imply \(x \leq z\)
  - Anti-symmetric: \(x \leq y\) and \(y \leq x\) imply \(x = y\)

- Total order = partial order with \(x \leq y\) or \(y \leq x\)

- Important example: domains of partial functions

\[(A \rightarrow B, \leq)\]

\(f \leq g\) iff \(g\) defined everywhere \(f\) is defined and
\(f(a) = g(a)\) whenever \(f(a)\) is defined
(Least) Upper Bounds

- An upper bound (u.b.) of $X \subseteq D$ is any element $p \in D$ such that $x \subseteq p$ for any $x \in X$.

- The least upper bound (l.u.b.) of $X \subseteq D$, written $\sqcup X$, is an upper bound with $\sqcup X \subseteq q$ for any u.b. $q$.
  - When they exist, least upper bounds are unique.

- The domains of partial functions, $(A \rightarrow B, \leq)$, admit upper bounds and least upper bounds if and only if all the partial functions in the considered set are compatible: any two agree on any element in which both are defined.
Complete Partial Orders (CPO)

- A chain in \((D, \subseteq)\) is an infinite sequence
  \[ d_0 \subseteq d_1 \subseteq d_2 \subseteq \ldots \subseteq d_n \subseteq \ldots \]
  also written
  \[ \{d_n \mid n \in \mathbb{N}\} \]

- Partial order \((D, \subseteq)\) is a complete partial order (CPO) iff any of its chains admits a least upper bound

- \((D, \subseteq, \bot)\) is a bottomed CPO (BCPO) iff \(\bot\) is a minimal element of \(D\), also called its bottom

- The domain of partial functions \((A \to B, \leq, \bot)\) is a BCPO, where \(\bot\) is the partial function undefined everywhere
Monotone and Continuous Functions

- \( \mathcal{F} : (D, \sqsubseteq) \to (D', \sqsubseteq') \) monotone iff
  \[ x \sqsubseteq y \quad \text{implies} \quad \mathcal{F}(x) \sqsubseteq' \mathcal{F}(y) \]

- Monotone functions preserve chains:
  \[ \{d_n \mid n \in \mathbb{N}\} \quad \text{chain implies} \quad \{\mathcal{F}(d_n) \mid n \in \mathbb{N}\} \quad \text{chain} \]
  However, they do not always preserve l.u.b. of chains

- \( \mathcal{F} : (D, \sqsubseteq) \to (D', \sqsubseteq') \) continuous iff monotone and preserves l.u.b. of chains:
  \[ \sqcup \mathcal{F}(d_n) = \mathcal{F}(\sqcup d_n) \]

- \( \text{Cont}((D, \sqsubseteq, \bot), (D', \sqsubseteq', \bot')) \), the domain of continuous functions between two BCPOs, is itself a BCPO
Fixed-Point Theorem

- Let \((D, \sqsubseteq, \bot)\) be a BCPO and \(\mathcal{F} : (D, \sqsubseteq, \bot) \to (D, \sqsubseteq, \bot)\) be a continuous function. Then the l.u.b. of the chain \(\{\mathcal{F}^n(\bot) \mid n \in \mathbb{N}\}\) is the least fixed-point of \(\mathcal{F}\).

  - Typically written \(\text{fix}(\mathcal{F})\)

- Proof sketch:

\[
\begin{align*}
\mathcal{F}(\text{fix}(\mathcal{F})) & = \mathcal{F}(\bigsqcup_{n \in \mathbb{N}} \mathcal{F}^n(\bot)) \\
& = \bigsqcup_{n \in \mathbb{N}} \mathcal{F}^{n+1}(\bot) \\
& = \bigsqcup_{n \in \mathbb{N}} \mathcal{F}^n(\bot) \\
& = \text{fix}(\mathcal{F}).
\end{align*}
\]
Applications of Fixed-Point Theorem

- Consider the following “definition” of the factorial:

\[ f(n) = \begin{cases} 
1 & \text{, if } n = 0 \\
 n \times f(n - 1) & \text{, if } n > 0 
\end{cases} \]

- This is a recursive definition

- Is it well-defined? Why?

- Yes. Because it is the least fixed-point of the following continuous (prove it!) function from \( \mathbb{N} \to \mathbb{N} \) to itself

\[ \mathcal{F}(g)(n) = \begin{cases} 
1 & \text{, if } n = 0 \\
 n \times g(n - 1) & \text{, if } n > 0 \text{ and } g(n - 1) \text{ defined} \\
 \text{undefined} & \text{, if } n > 0 \text{ and } g(n - 1) \text{ undefined} 
\end{cases} \]
Denotational Semantics of IMP

Arithmetic Expressions

\[
\llbracket - \rrbracket : AExp \rightarrow (\text{State} \rightarrow \text{Int})
\]

\[
\llbracket i \rrbracket_\sigma = i
\]

\[
\llbracket x \rrbracket_\sigma = \sigma(x)
\]

\[
\llbracket a_1 + a_2 \rrbracket_\sigma = \llbracket a_1 \rrbracket_\sigma +_{\text{Int}} \llbracket a_2 \rrbracket_\sigma
\]

\[
\llbracket a_1 / a_2 \rrbracket_\sigma = \begin{cases} 
\llbracket a_1 \rrbracket_\sigma /_{\text{Int}} \llbracket a_2 \rrbracket_\sigma & \text{if} \quad \llbracket a_2 \rrbracket_\sigma \neq 0 \\
\bot & \text{if} \quad \llbracket a_2 \rrbracket_\sigma = 0
\end{cases}
\]
Denotational Semantics of IMP

Boolean Expressions

\[ [\_ ] : BExp \rightarrow (State \rightarrow Bool) \]

\[ [t]_\sigma = t \]

\[ [a_1 <= a_2]_\sigma = [a_1]_\sigma \leq_{int} [a_2]_\sigma \]

\[ [\neg b]_\sigma = \neg_{Bool}( [b]_\sigma ) \]

\[ [b_1 \&\& b_2]_\sigma = \begin{cases} 
[b_2]_\sigma & \text{if } [b_1]_\sigma = \text{true} \\
\text{false} & \text{if } [b_1]_\sigma = \text{false} \\
\bot & \text{if } [b_1]_\sigma = \bot 
\end{cases} \]
Denotational Semantics of IMP Statements (without loops)

\[ [] : Stmt \rightarrow (State \rightarrow State) \]

\[ [\{\}]\sigma = \sigma \]

\[ [\{ s \}]\sigma = [s]\sigma \]

\[ [x = a ; ]\sigma = \begin{cases} 
\sigma[[a]\sigma/x] & \text{if } \sigma(x) \neq \bot \\
\bot & \text{if } \sigma(x) = \bot 
\end{cases} \]

\[ [s_1 \; s_2]\sigma = [s_2][s_1]\sigma \]

\[ [\text{if}(b) \; s_1 \; \text{else} \; s_2]\sigma = \begin{cases} 
[s_1]\sigma & \text{if } [b]\sigma = \text{true} \\
[s_2]\sigma & \text{if } [b]\sigma = \text{false} \\
\bot & \text{if } [b]\sigma = \bot 
\end{cases} \]
Denotational Semantics of IMP

While

- We first define a continuous function as follows

\[ F : (\text{State} \rightarrow \text{State}) \rightarrow (\text{State} \rightarrow \text{State}) \]

\[ F(\alpha)(\sigma) = \begin{cases} 
\alpha([s] \sigma) & \text{if } [b] \sigma = \text{true} \\
\sigma & \text{if } [b] \sigma = \text{false} \\
\bot & \text{if } [b] \sigma = \bot
\end{cases} \]

- Then we define the denotational semantics of while

\[ [[\text{while } (b) s]] = \text{fix}(F) \]
MSOS

Modular structural operational semantics
Modular Structural Operational Semantics (Modular SOS, or MSOS)

- Peter Mosses (1999)
- Addresses the non-modularity aspects of SOS
  - A definitional framework is non-modular when, in order to add a new feature to an existing language, one needs to revisit and change some of the already defined, unrelated language features
  - The non-modularity of SOS becomes clear when we define IMP++
- Why modularity is important
  - Modifying existing rules when new rules are added is error prone
  - When experimenting with language design, one needs to make changes quickly; having to do unrelated changes slows us down
  - Rapid language development, e.g., domain-specific languages
Philosophy of MSOS

- Separate the syntax from configurations and treat it differently
- Transitions go from syntax to syntax, hiding the other configuration components into transition labels
- Labels encode all the non-syntactic configuration changes
- Specialized notation in transition labels, to
  - Say that certain configuration components stay unchanged
  - Say that certain configuration changes are propagated from the premise to the conclusion of a rule
An MSOS transition has the form

\[ P \xrightarrow{\Delta} P' \]

- \( P \) and \( P' \) are programs or fragments of program
- \( \Delta \) is a label describing the changes in the configuration components, defined as a record; primed fields stay for “after” the transition

Example:

\[
x := i \quad \{ \text{state} = \sigma, \text{state}' = \sigma[i/x], \ldots \} \quad \xrightarrow{} \quad \text{skip} \quad \text{if} \quad \sigma(x) \neq \bot
\]

This rule can be automatically “desugared” into the SOS rule

\[
\langle x := i, \sigma \rangle \rightarrow \langle \text{skip}, \sigma[i/x] \rangle \quad \text{if} \quad \sigma(x) \neq \bot
\]

But also into (if the configuration contains more components, like in IMP++)

\[
\langle x := i, \sigma, \omega \rangle \rightarrow \langle \text{skip}, \sigma[i/x], \omega \rangle \quad \text{if} \quad \sigma(x) \neq \bot
\]
Labels are field assignments, or records, and can use “…” for “and so on”, called record comprehension.

Fields can be primed or not.
- Unprimed = configuration component before the transition is applied
- Primed = configuration component after the transition is applied

Some fields appear both unprimed and primed (called read-write), while others appear only primed (called write-only) or only unprimed (called read-only).
MSOS Labels

- **Field types**
  - **Read/write** = fields which appear both unprimed and unprimed
    \[ x := i \xrightarrow{\{\text{state}=\sigma, \text{state}'=\sigma[i/x], \ldots\}} \text{skip} \quad \text{if } \sigma(x) \neq \bot \]
  - **Write-only** = fields which appear only primed
    \[ \text{print } i \xrightarrow{\{\text{output}'=i, \ldots\}} \text{skip} \]
  - **Read-only** = fields which appear only unprimed
    \[ e_2 \xrightarrow{\{\text{env}=\rho[v_1/x],\ldots\}} e'_2 \]
    \[ \text{let } x = v_1 \text{ in } e_2 \xrightarrow{\{\text{env}=\rho,\ldots\}} \text{let } x = v_1 \text{ in } e'_2 \]
MSOS Rules

- Like in SOS, but using MSOS transitions as sequents
- Same labels or parts of them can be used multiple times in a rule
- Example:

\[
\frac{s_1 \xrightarrow{\Delta} s_1'}{s_1 \; s_2 \xrightarrow{\Delta} s_1' \; s_2}
\]

- Same $\Delta$ means that changes propagate from premise to conclusion
- The author of MSOS now promotes a simplifying notation
  - If the premise and the conclusion repeat the same label or part of it, simply drop that label or part of it. For example:
MSOS of IMP - Arithmetic

\[
x \xrightarrow{\text{state}=\sigma, \ldots} \sigma(x) \quad \text{if } \sigma(x) \neq \bot
\]

\[
\frac{a_1 \rightarrow a'_1}{a_1 + a_2 \rightarrow a'_1 + a_2}
\]

(\text{MSOS-Add-Arg1})

\[
\frac{a_2 \rightarrow a'_2}{a_1 + a_2 \rightarrow a_1 + a'_2}
\]

(\text{MSOS-Add-Arg2})

\[
i_1 + i_2 \rightarrow i_1 + \text{Int } i_2
\]

(\text{MSOS-Add})
MSOS of IMP - Arithmetic

\[
\begin{align*}
\frac{a_1 \rightarrow a'_1}{a_1 / a_2 \rightarrow a'_1 / a_2} & \quad \text{(MSOS-Div-Arg1)} \\
\frac{a_2 \rightarrow a'_2}{a_1 / a_2 \rightarrow a_1 / a'_2} & \quad \text{(MSOS-Div-Arg2)} \\
i_1 / i_2 \rightarrow i_1 / \text{Int } i_2 & \quad \text{if } i_2 \neq 0 \quad \text{(MSOS-Div)}
\end{align*}
\]
**MSOS of IMP - Boolean**

\[
\frac{a_1 \rightarrow a'_1}{a_1 \leq a_2 \rightarrow a'_1 \leq a_2} \quad \text{(MSOS-Leq-Arg1)}
\]

\[
\frac{a_2 \rightarrow a'_2}{i_1 \leq a_2 \rightarrow i_1 \leq a'_2} \quad \text{(MSOS-Leq-Arg2)}
\]

\[
i_1 \leq i_2 \rightarrow i_1 \leq_{\text{Int}} i_2 \quad \text{(MSOS-Leq)}
\]
**MSOS of IMP - Boolean**

\[
\frac{b \rightarrow b'}{\neg b \rightarrow \neg b'} \quad (\text{MSOS-Not-Arg})
\]

\[
\text{not true} \rightarrow \text{false} \quad (\text{MSOS-Not-True})
\]

\[
\text{not false} \rightarrow \text{true} \quad (\text{MSOS-Not-False})
\]

\[
\frac{b_1 \rightarrow b_1'}{b_1 \text{ and } b_2 \rightarrow b_1' \text{ and } b_2} \quad (\text{MSOS-And-Arg1})
\]

\[
\text{false and } b_2 \rightarrow \text{false} \quad (\text{MSOS-And-False})
\]

\[
\text{true and } b_2 \rightarrow b_2 \quad (\text{MSOS-And-True})
\]
MSOS of IMP - Statements

\[
\frac{a \rightarrow a'}{x := a \rightarrow x := a'}
\]

\[
x := i \xrightarrow{\{\text{state}=\sigma, \text{state}'=\sigma[i/x], \ldots\}} \text{skip \hspace{1cm} if } \sigma(x) \neq 1
\]

\[
\frac{s_1 \rightarrow s_1'}{s_1 ; s_2 \rightarrow s_1' ; s_2}
\]

\[
\text{skip } ; s_2 \rightarrow s_2
\]
**MSOS of IMP - Statements**

\[
b \rightarrow b' \\
\frac{\text{if } b \text{ then } s_1 \text{ else } s_2 \rightarrow \text{ if } b' \text{ then } s_1 \text{ else } s_2}{\text{(MSOS-If-Arg1)}}
\]

\[
\frac{\text{if true then } s_1 \text{ else } s_2 \rightarrow s_1}{\text{(MSOS-If-True)}}
\]

\[
\frac{\text{if false then } s_1 \text{ else } s_2 \rightarrow s_2}{\text{(MSOS-If-False)}}
\]

\[
\text{while } b \text{ do } s \rightarrow \text{ if } b \text{ then } (s ; \text{ while } b \text{ do } s) \text{ else skip}
\]

\[
\frac{\text{var } x_l ; s \rightarrow s}{\text{(MSOS-Var)}}
\]
RSEC

Reduction semantics with evaluation contexts
Reduction Semantics with Evaluation Contexts (RSEC)

- Matthias Felleisen and collaborators (1992)
- Previous operational approaches encoded the program execution context as a proof context, by means of rule conditions or premises
  - This has a series of advantages, but makes it hard to define control intensive features, such as abrupt termination, exceptions, call/cc, etc.
- We would like to have the execution context explicit, so that we can easily save it, change it, or even delete it
- Reduction semantics with evaluation contexts does precisely that
  - It allows to formally define evaluation contexts
  - Rules become mostly unconditional
  - Reductions can only happen “in context”
RSEC relies on reversible implicit mechanisms to
- Split syntax into an evaluation context and a redex
- Plug a redex into an evaluation contexts and obtain syntax again

\[ p = c[e] \]
Evaluation Contexts

- Evaluation contexts are typically defined by the same means that we use to define the language syntax, that is, grammars.
- The hole $\Box$ represents the place where redex is to be plugged.
- Example:

```
Context ::= $\Box$
  Context <= AExp
  Int <= Context
  Id := Context
  Context ; Stmt
  if Context then Stmt else Stmt
  ...
```
Correct Evaluation Contexts

\[
\begin{align*}
\quad & 3 \leq \square \\
\quad & \square \leq 3 \\
\quad & \square ; x := 5 \\
\quad & \text{if } \square \text{ then } s_1 \text{ else } s_2
\end{align*}
\]
Wrong Evaluation Contexts

\[ \square \leq \square \]

\[ x \leq 3 \]

\[ x \leq \square \]

\[ x := 5 \ ; \ \square \]

\[ \square := 5 \]

\[ \text{if } x \leq 7 \text{ then } \square \text{ else } x := 5 \]
Splitting/Plugging of Syntax

\[
7 = (\Box)[7]
\]

\[
3 \leq x = (3 \leq \Box)[x] = (\Box \leq x)[3] = (\Box)[3 \leq x]
\]

\[
3 \leq (2 + x) + 7 = (3 \leq \Box + 7)[2 + x] = (\Box \leq (2 + x) + 7)[3] = \ldots
\]
The characteristic rule of RSEC allows us to only define semantic rules stating how redexes are reduced

- This significantly reduces the number of rules
- The semantic rules are mostly unconditional (no premises)
- The overall result is a semantics which is compact and easy to read and understand
RSEC of IMP — Evaluation Contexts

<table>
<thead>
<tr>
<th>IMP evaluation contexts syntax</th>
<th>IMP language syntax</th>
</tr>
</thead>
<tbody>
<tr>
<td><strong>Context</strong> ::=</td>
<td><strong>AExp</strong> ::=</td>
</tr>
<tr>
<td>□</td>
<td>*Int</td>
</tr>
<tr>
<td>Context + AExp</td>
<td>AExp + Context</td>
</tr>
<tr>
<td>Context / AExp</td>
<td>AExp / Context</td>
</tr>
<tr>
<td>Context &lt;= AExp</td>
<td>Int &lt;= Cxt</td>
</tr>
<tr>
<td>not Context</td>
<td></td>
</tr>
<tr>
<td>Context and BExp</td>
<td></td>
</tr>
<tr>
<td>*Id := Context</td>
<td></td>
</tr>
<tr>
<td>Context ; Stmt</td>
<td></td>
</tr>
<tr>
<td>if Context then Stmt else Stmt</td>
<td></td>
</tr>
</tbody>
</table>

\[ Pgm ::= \text{varList}\{Id\} ; Stmt \]
RSEC of IMP — Rules

\[
\text{Context ::= \ldots} \quad | \quad \langle \text{Context}, \text{State} \rangle
\]

\[
e \rightarrow e' \\
\frac{c[e] \rightarrow c[e']}{}
\]

\[
\langle c, \sigma \rangle[x] \rightarrow \langle c, \sigma \rangle[\sigma(x)] \quad \text{if } \sigma(x) \neq \bot
\]
\[
i_1 + i_2 \rightarrow i_1 +_{\text{Int}} i_2
\]
\[
i_1 / i_2 \rightarrow i_1 /_{\text{Int}} i_2 \quad \text{if } i_2 \neq 0
\]
\[
i_1 \leq i_2 \rightarrow i_1 \leq_{\text{Int}} i_2
\]

not true \rightarrow false
not false \rightarrow true
true and b_2 \rightarrow b_2
false and b_2 \rightarrow false
\[
\langle c, \sigma \rangle[x := i] \rightarrow \langle c, \sigma[i/x] \rangle[\text{skip}] \quad \text{if } \sigma(x) \neq \bot
\]
skip ; s_2 \rightarrow s_2
if true then s_1 else s_2 \rightarrow s_1
if false then s_1 else s_2 \rightarrow s_2
while b do s \rightarrow if b then (s ; while b do s) else skip
\[
\langle \text{var } x l \ ; \ s \rangle \rightarrow \langle s, (x l \Rightarrow 0) \rangle
\]
RSEC Derivation

\[
\begin{align*}
\langle x := 1 ; y := 2 ; \text{if } x \leq y \text{ then } x := 0 \text{ else } y := 0 , (x \mapsto 0, y \mapsto 0) \rangle \\
\rightarrow \langle \text{skip} ; y := 2 ; \text{if } x \leq y \text{ then } x := 0 \text{ else } y := 0 , (x \mapsto 1, y \mapsto 0) \rangle \\
\rightarrow \langle y := 2 ; \text{if } x \leq y \text{ then } x := 0 \text{ else } y := 0 , (x \mapsto 1, y \mapsto 0) \rangle \\
\rightarrow \langle \text{skip} ; \text{if } x \leq y \text{ then } x := 0 \text{ else } y := 0 , (x \mapsto 1, y \mapsto 2) \rangle \\
\rightarrow \langle \text{if } x \leq y \text{ then } x := 0 \text{ else } y := 0 , (x \mapsto 1, y \mapsto 2) \rangle \\
\rightarrow \langle \text{if } 1 \leq y \text{ then } x := 0 \text{ else } y := 0 , (x \mapsto 1, y \mapsto 2) \rangle \\
\rightarrow \langle \text{if } 1 \leq 2 \text{ then } x := 0 \text{ else } y := 0 , (x \mapsto 1, y \mapsto 2) \rangle \\
\rightarrow \langle \text{if true then } x := 0 \text{ else } y := 0 , (x \mapsto 1, y \mapsto 2) \rangle \\
\rightarrow \langle x := 0 , (x \mapsto 1, y \mapsto 2) \rangle \\
\rightarrow \langle \text{skip} , (x \mapsto 0, y \mapsto 2) \rangle
\end{align*}
\]
CHAM

The chemical abstract machine
The Chemical Abstract Machine (CHAM)

- Berry and Boudol (1992)
- Both a model of concurrency and a specific semantic style
- Chemical metaphor
  - States regarded as chemical solutions containing floating molecules
  - Molecules can interact with each other by means of reactions
  - Reactions take place concurrently, unrestricted by context
  - Solutions are encapsulated within new molecules, using membranes
    - The following is a solution containing $k$ molecules:

\[
\{m_1, m_2, \ldots, m_k\}
\]
**CHAM Syntax and Rules**

\[
\begin{align*}
Molecule & ::= \text{Solution} \\
& \quad \mid Molecule \triangleleft \text{Solution}
\end{align*}
\]

\[
Solution ::= \{Bag\{Molecule\}\}
\]
CHAM Rules

- Ordinary rewrite rules between solution terms:

\[ m_1\ m_2\ \ldots\ m_k \rightarrow m'_1\ m'_2\ \ldots\ m'_l \]

- Rewriting takes place only within solutions

- Three (metaphoric) kinds of rules
  - Heating rules using \( \rightarrow \): structurally rearrange solution
  - Cooling rules using \( \rightarrow \): clean up solution after reactions
  - Reaction rules using \( \rightarrow \): change solution irreversibly
CHAM Airlock

- Allows to extract molecules from encapsulated solutions
- Governed by two rules coming in a heating/cooling pair:

\[ \{m_1, m_2, \ldots, m_k\} \Leftrightarrow \{m_1 \triangleleft \{m_2, \ldots, m_k\}\} \]
A top-level solution containing two subsolution molecules
- One for holding the syntax
- Another for holding the state

```
{ { Syntax } { State } }
```

Example:
```
{ { x := (3 / (x + 2)) } { x ⟷ 1  y ⟷ 0 } }
```
Airlock can be Problematic

- Airlock cannot be used to encode evaluation strategies; heating/cooling rules of the form

\[
\begin{align*}
x & : = a \iff a \triangleleft \{ x : = \Box \} \\
a_1 + a_2 & \iff a_1 \triangleleft \{ \Box + a_2 \} \\
a_1 + a_2 & \iff a_2 \triangleleft \{ a_1 + \Box \}
\end{align*}
\]

are problematic, because they yield ambiguity, e.g.,

\[
\{ x : = (3 / (x + 2)) \} \iff x : = ((3 / x) + 2)
\]
Correct Representation of Syntax

- Other attempts fail, too (see the lecture notes)
- We need some mechanism which is not based on airlocks
- We borrow the representation approach of K
  - Term $x := (3 / (x + 2))$ represented as
    $x \leadsto (\square + 2) \leadsto (3 / \square) \leadsto (x := \square) \leadsto \square$
  
- Can be achieved using heating/cooling rules of the form
  $$(x := a) \leadsto c \quad \Rightarrow \quad a \leadsto (x := \square) \leadsto c$$
  $$\left(a_1 / a_2\right) \leadsto c \quad \Rightarrow \quad a_2 \leadsto \left(a_1 / \square\right) \leadsto c$$
  $$\left(a_1 + a_2\right) \leadsto c \quad \Rightarrow \quad a_1 \leadsto \left(\square + a_2\right) \leadsto c$$
  $$s \quad \Rightarrow \quad s \leadsto \square$$
CHAM Heating-Cooling Rules for IMP

\[
\begin{align*}
    a_1 + a_2 &\sim c \quad \Rightarrow \quad a_1 \sim \wedge + a_2 \sim c \\
    a_1 + a_2 &\sim c \quad \Rightarrow \quad a_2 \sim a_1 + \wedge \sim c \\
    a_1 / a_2 &\sim c \quad \Rightarrow \quad a_1 \sim \wedge / a_2 \sim c \\
    a_1 / a_2 &\sim c \quad \Rightarrow \quad a_2 \sim a_1 / \wedge \sim c \\
    a_1 \leq a_2 &\sim c \quad \Rightarrow \quad a_1 \sim \wedge \leq a_2 \sim c \\
    i_1 \leq a_2 &\sim c \quad \Rightarrow \quad a_2 \sim i_1 \leq \wedge \sim c \\
    \text{not } b &\sim c \quad \Rightarrow \quad b \sim \text{not } \wedge \sim c \\
    b_1 \text{ and } b_2 &\sim c \quad \Rightarrow \quad b_1 \sim \wedge \text{ and } b_2 \sim c \\
    x := a &\sim c \quad \Rightarrow \quad a \sim x := \wedge \sim c \\
    s_1 ; s_2 &\sim c \quad \Rightarrow \quad s_1 \sim \wedge ; s_2 \sim c \\
    s &\Rightarrow \quad s \sim \wedge \\
    \text{if } b \text{ then } s_1 \text{ else } s_2 &\sim c \quad \Rightarrow \quad b \sim \text{if } \wedge \text{ then } s_1 \text{ else } s_2 \sim c
\end{align*}
\]
Examples of Syntax Heating/Cooling

- The following is correct heating/cooling of syntax:

\[
\{ x := 1 ; x := (3 / (x + 2)) \} \iff^*
\]
\[
\{ x := 1 \sim (\square ; x := (3 / (x + 2))) \sim \square \}
\]

- The following is incorrect heating/cooling of syntax:

\[
\{ x := 1 ; x := (3 / (x + 2)) \} \iff^*
\]
\[
\{ x := 1 ; (x \sim (\square + 2) \sim (3 / \square) \sim (x := \square) \sim \square) \}
\]
CHAM Reaction Rules for IMP

\[
\begin{align*}
\{x \sim c\} \& \{x \mapsto i \triangleright \sigma\} & \rightarrow & \{i \sim c\} \& \{x \mapsto i \triangleright \sigma\} \\
i_1 + i_2 \sim c & \rightarrow & i_1 +_{\text{Int}} i_2 \sim c \\
i_1 / i_2 \sim c & \rightarrow & i_1 /_{\text{Int}} i_2 \sim c \quad \text{when } i_2 \neq 0 \\
i_1 \leq i_2 \sim c & \rightarrow & i_1 \leq_{\text{Int}} i_2 \sim c \\
\text{not true} \sim c & \rightarrow & \text{false} \sim c \\
\text{not false} \sim c & \rightarrow & \text{true} \sim c \\
\text{true and } b_2 \sim c & \rightarrow & b_2 \sim c \\
\text{false and } b_2 \sim c & \rightarrow & \text{false} \sim c \\
\{x := i \sim c\} \& \{x \mapsto j \triangleright \sigma\} & \rightarrow & \{\text{skip} \sim c\} \& \{x \mapsto i \triangleright \sigma\} \\
\text{skip } ; s_2 \sim c & \rightarrow & s_2 \sim c \\
\text{if true then } s_1 \text{ else } s_2 \sim c & \rightarrow & s_1 \sim c \\
\text{if false then } s_1 \text{ else } s_2 \sim c & \rightarrow & s_2 \sim c \\
\text{while } b \text{ do } s \sim c & \rightarrow & \text{if } b \text{ then } (s ; \text{while } b \text{ do } s) \text{ else skip} \sim c \\
\text{var } x_l ; s & \rightarrow & \{s\} \& \{x_l \mapsto 0\} \\
(x, x_l) \mapsto i & \rightarrow & x \mapsto i \triangleright \{x_l \mapsto i\}
\end{align*}
\]
Sample CHAM Rewriting

\[
\begin{align*}
\{ \text{var } x, y ; x := 1 ; x := (3 / (x + 2)) \} & \rightarrow \\
\{ \{ x := 1 ; x := (3 / (x + 2)) \} \} & \rightarrow \{ x, y \mapsto 0 \} \\
\{ \{ x := 1 ; x := (3 / (x + 2)) \rightarrow \Box \} \} & \rightarrow \{ x \mapsto 0 \ \ y \mapsto 0 \} \\
\{ \{ x := 1 \rightarrow \Box ; x := (3 / (x + 2)) \rightarrow \Box \} \} & \rightarrow \{ x \mapsto 0 \ \ y \mapsto 0 \} \\
\{ \{ \text{skip} \rightarrow \Box ; x := (3 / (x + 2)) \rightarrow \Box \} \} & \rightarrow \{ x \mapsto 1 \ \ y \mapsto 0 \} \\
\{ \{ \text{skip} ; x := (3 / (x + 2)) \rightarrow \Box \} \} & \rightarrow \{ x \mapsto 1 \ \ y \mapsto 0 \} \\
\{ \{ x := (3 / (x + 2)) \rightarrow \Box \} \} & \rightarrow \{ x \mapsto 1 \ \ y \mapsto 0 \} \\
\{ \{ x \rightarrow \Box + 2 \rightarrow 3 / \Box \rightarrow x := \Box \rightarrow \Box \} \} & \rightarrow \{ x \mapsto 1 \ \ y \mapsto 0 \} \\
\{ \{ 1 \rightarrow \Box + 2 \rightarrow 3 / \Box \rightarrow x := \Box \rightarrow \Box \} \} & \rightarrow \{ x \mapsto 1 \ \ y \mapsto 0 \} \\
\{ \{ 1 + 2 \rightarrow 3 / \Box \rightarrow x := \Box \rightarrow \Box \} \} & \rightarrow \{ x \mapsto 1 \ \ y \mapsto 0 \} \\
\{ \{ 3 \rightarrow 3 / \Box \rightarrow x := \Box \rightarrow \Box \} \} & \rightarrow \{ x \mapsto 1 \ \ y \mapsto 0 \} \\
\{ \{ x := 1 \rightarrow \Box \} \} & \rightarrow \{ x \mapsto 1 \ \ y \mapsto 0 \} \\
\{ \{ \text{skip} \rightarrow \Box \} \} & \rightarrow \{ x \mapsto 1 \ \ y \mapsto 0 \} \\
\{ \{ \text{skip} \} \} & \rightarrow \{ x \mapsto 1 \ \ y \mapsto 0 \}
\end{align*}
\]
COMPARING
CONVENTIONAL
EXECUTABLE
SEMANTICS

How good are the various semantic approaches?
We next discuss the conventional executable semantics approaches in depth, aiming at understanding their pros and cons.

Our approach is to extend each semantics of IMP with various common features (we call the resulting language IMP++)

- **Variable increment** – this will add side effects to expressions
- **Input/Output** – this will require changes in the configuration
- **Abrupt termination** – this requires explicit handling of control
- **Dynamic threads** – this requires handling concurrency and sharing
- **Local variables** – this requires handling environments

We will first treat each extension of IMP independently, i.e., we do not pro-actively take semantic decisions when defining a feature that will help the definition of other features later on. Then, we will put all features together into our IMP++ final language.
IMP++ Variable Increment

- **Syntax:**

  \[ AExp ::= \ldots \mid ++ \text{Id} \]

- Variable increment is very common (C, C++, Java, etc.)
  - We only consider pre-increment (first increment, then return value)
  - The problem with increment in some semantic approaches is that it adds side effects to expressions. Therefore, if one did not pro-actively account for that then one needs to change many existing and unrelated semantics rules, if not all.
IMP++ Variable Increment

Big-Step SOS

- Previous big-step SOS rules had the form:

\[
\frac{\langle a_1, \sigma \rangle \Downarrow \langle i_1 \rangle \quad \langle a_2, \sigma \rangle \Downarrow \langle i_2 \rangle}{\langle a_1 / a_2, \sigma \rangle \Downarrow \langle i_1 / \text{Int} \ i_2 \rangle}, \text{ where } i_2 \neq 0
\]

- Big-step SOS is the most affected by side effects
  - Needs to change its sequents from \( \langle a, \sigma \rangle \Downarrow \langle i \rangle \) to \( \langle a, \sigma \rangle \Downarrow \langle i, \sigma' \rangle \)
  - And all the existing rules accordingly, e.g.:

\[
\frac{\langle a_1, \sigma \rangle \Downarrow \langle i_1, \sigma_1 \rangle, \quad \langle a_2, \sigma_1 \rangle \Downarrow \langle i_2, \sigma_2 \rangle}{\langle a_1 / a_2, \sigma \rangle \Downarrow \langle i_1 / \text{Int} \ i_2, \sigma_2 \rangle}, \text{ where } i_2 \neq 0
\]
Recall IMP operators like / were non-deterministically strict. Here is an attempt to achieve that with big-step SOS

\[
\begin{align*}
\langle a_1, \sigma \rangle & \downarrow \langle i_1, \sigma_1 \rangle, \quad \langle a_2, \sigma_1 \rangle \downarrow \langle i_2, \sigma_2 \rangle \\
\langle a_1 / a_2, \sigma \rangle & \downarrow \langle i_1 / \text{Int}i_2, \sigma_2 \rangle
\end{align*}
\]

where \( i_2 \neq 0 \)

\[
\begin{align*}
\langle a_1, \sigma_2 \rangle & \downarrow \langle i_1, \sigma_1 \rangle, \quad \langle a_2, \sigma \rangle \downarrow \langle i_2, \sigma_2 \rangle \\
\langle a_1 / a_2, \sigma \rangle & \downarrow \langle i_1 / \text{Int}i_2, \sigma_1 \rangle
\end{align*}
\]

where \( i_2 \neq 0 \)

All we got is “non-deterministic choice” strictness: choose an order, then evaluate the arguments in that order.

Some behaviors are thus lost, but this is relatively acceptable in practice since programmers should not rely on those behaviors in their programs anyway (the loss of behaviors when we add threads is going to be much worse)
We are now ready to add the big-step SOS for variable increment (this is easy now, the hard part was to get here):

\[
\langle \text{++ } x, \sigma \rangle \Downarrow \langle \sigma(x) + \text{Int } 1, \sigma[((\sigma(x) + \text{Int } 1)/x]\rangle
\]

Example:
- How many values can the following expression possibly evaluate to under the big-step SOS of IMP++ above (assume \( x \) is initially 1)?

\[
\text{++ } x / (\text{++ } x / x)
\]
- Can it evaluate to 0 or even be undefined under a fully non-deterministic evaluation strategy?
Previous small-step SOS rules had the form:

\[
\langle a_1, \sigma \rangle \rightarrow \langle a'_1, \sigma \rangle \\
\langle a_1 / a_2, \sigma \rangle \rightarrow \langle a'_1 / a_2, \sigma \rangle
\]

Small-step SOS less affected than big-step SOS, but still requires many rule changes to account for the side effects:

\[
\langle a_1, \sigma \rangle \rightarrow \langle a'_1, \sigma_1 \rangle \\
\langle a_1 / a_2, \sigma \rangle \rightarrow \langle a'_1 / a_2, \sigma_1 \rangle
\]
IMP++ Variable Increment
Small-Step SOS

- Since small-step SOS “gets back to the top” at each step, it actually does not lose any non-deterministic behaviors
  - We get fully non-deterministic evaluation strategies for all the IMP constructs instead of “non-deterministic choice” ones
- The semantics of variable increment almost the same as in big-step SOS (indeed, variable increment is an atomic operation):

\[
\langle \mathtt{++x}, \sigma \rangle \rightarrow \langle \sigma(x) +_{\text{int}} 1, \sigma[(\sigma(x) +_{\text{int}} 1)/x] \rangle
\]
IMP++ Variable Increment

**MSOS**

- Previous MSOS rules had the form:

\[
\frac{a_1 \rightarrow a_1'}{a_1 / a_2 \rightarrow a_1' / a_2}
\]

- All semantic changes are hidden within labels, which are implicitly propagated through the general MSOS mechanism.

- Consequently, the MSOS of IMP only needs the following rule to accommodate variable updates; nothing else changes!

\[
++ x \quad \{\text{state}=\sigma, \text{state'}=\sigma[(\sigma(x)+\text{Int } 1)/x], \ldots\} \quad \sigma(x) +_{\text{Int}} 1
\]
IMP++ Variable Increment
Reduction Semantics with Eval. Contexts

- Previous RSEC evaluation contexts and rules had the form:

\[
\text{Context} ::= \square \mid \text{Context} / \text{AExp} \mid \text{AExp} / \text{Context}
\]

\[
i_1 / i_2 \rightarrow i_1 /_{\text{Int}}i_2, \quad \text{when } i_2 \neq 0
\]

\[
\langle c, \sigma \rangle[x] \rightarrow \langle c, \sigma \rangle[\sigma(x)]
\]

- Evaluation contexts, together with the characteristic rule of RSEC, allows for compact unconditional rules, mentioning only what is needed from the entire configuration.

- Consequently, the RSED of IMP only needs the following rule to accommodate variable updates; nothing else changes!

\[
\langle c, \sigma \rangle[+x] \rightarrow \langle c, \sigma[\left(\sigma(x) +_{\text{Int}} 1\right)/x]\rangle[\sigma(x) +_{\text{Int}} 1]
\]
Previous CHAM heating/cooling/reaction rules had the form:

\[
\begin{align*}
a_1 / a_2 \sim c &\Rightarrow a_1 \sim \square / a_2 \sim c \\
a_1 / a_2 \sim c &\Rightarrow a_2 \sim a_1 / \square \sim c \\
i_1 / i_2 \sim c &\Rightarrow i_1 /_{Int} i_2 \sim c \quad \text{when } i_2 \neq 0
\end{align*}
\]

\[
\{x \sim c\} \ {\{x \leftrightarrow i \triangleright \sigma\}} \rightarrow \{i \sim c\} \ {\{x \leftrightarrow i \triangleright \sigma\}}
\]

Since the heating/cooling rules achieve the role of the evaluation contexts and since one can only mention the necessary molecules in each rule, one does not need to change anything either!

All one needs to do is to add the following rule:

\[
\{++x \sim c\} \ {\{x \leftrightarrow i \triangleright \sigma\}} \rightarrow \{i +_{Int} 1 \sim c\} \ {\{x \leftrightarrow i +_{Int} 1 \triangleright \sigma\}}
\]
Where is the rest?

- We discussed the remaining features in class, using the whiteboard and colors.
- The lecture notes contain the complete information, even more than we discussed in class.