Programming Languages and Compilers (CS 421)

Talia Ringer (they/them) 4218 SC, UIUC



https://courses.grainger.illinois.edu/cs421/fa2023/

Based heavily on slides by Elsa Gunter, which were based in part on slides by Mattox Beckman, as updated by Vikram Adve and Gul Agha

Midterm Study Guide

Three Main Topics of the Course









Questions before we start?



Expresses the **meaning** of syntax

- Static semantics:
 - Meaning based only on the form of the expression without executing it
 - Usually restricted to type checking / type inference
 - **Dynamic** semantics:
 - Describes meaning of executing a program
 - Kinds: operational, axiomatic, denotational

- Expresses the **meaning** of syntax
- **Static** semantics:
 - Meaning based only on the form of the expression without executing it
 - Usually restricted to type checking / type inference
 - **Dynamic** semantics:
 - Describes meaning of executing a program
 - Kinds: operational, axiomatic, denotational

- Expresses the **meaning** of syntax
- **Static** semantics:
 - Meaning based only on the form of the expression without executing it
 - Usually restricted to type checking / type inference
 - **Dynamic** semantics:
 - Describes meaning of **executing** a program
 - Kinds: operational, axiomatic, denotational

Expresses the **meaning** of syntax

- Static semantics:
 - Meaning based only on the form of the expression without executing it
 - Usually restricted to type checking / type inference
- **Dynamic** semantics:
 - Describes meaning of **executing** a program

Kinds: operational, axiomatic, denotational

Dynamic Semantics

Why so many kinds of dynamic semantics?
 Different languages better suited to different kinds of semantics
 Different kinds serve different purposes

- Common to have **multiple** kinds and show how they **relate** to each other
- **Dynamic** semantics:
 - Describes meaning of **executing** a program
 - Kinds: operational, axiomatic, denotational

What it is:

- Describe how to execute (implement) programs of language on a virtual machine, by describing how to execute each program statement (i.e., following the structure of the program)
 Meaning of program is how its execution changes the state of the machine
 Tradeoffs:
 - Easy to implement
 - Hard to reason about abstractly (without thinking about implementation details)

What it is:

- Describe how to execute (implement) programs of language on a virtual machine, by describing how to execute each program statement (i.e., following the structure of the program)
 Meaning of program is how its execution changes the state of the machine Tradeoffs:
- Easy to implement
- Hard to reason about abstractly (without thinking about implementation details)

What it is:

Also called a **Program Logic**

- Commonly Floyd-Hoare logic
- These days, also separation logic
- Logical system built from *axioms* and *inference rules*
- Often written as pre-conditions and post-conditions on programs

- Mainly suited to imperative languages
- Good for external reasoning

What it is:

Also called a **Program Logic**

- Commonly Floyd-Hoare logic
- These days, also separation logic
- Logical system built from *axioms* and *inference rules*
- Often written as pre-conditions and post-conditions on programs

- Mainly suited to imperative languages
- Good for external reasoning

What it is:

Also called a **Program Logic**

- Commonly Floyd-Hoare logic
- These days, also separation logic
- Logical system built from *axioms* and *inference rules*
- Often written as pre-conditions and post-conditions on programs

- Mainly suited to imperative languages
- Good for external reasoning

 Used to formally prove a **post-condition** (property) of the **state** (the values of the program variables) after the execution of program, assuming a **pre-condition** (another property) holds before execution
 Written :

{Precondition} Program {Postcondition}Source of idea of **loop invariant**

Used to formally prove a **post-condition** (property) of the **state** (the values of the program variables) after the execution of program, assuming a **pre-condition** (another property) holds before execution
 Written :

{Precondition} Program {Postcondition}
 Source of idea of loop invariant

Used to formally prove a **post-condition** (property) of the **state** (the values of the program variables) after the execution of program, assuming a **pre-condition** (another property) holds before execution
 Written :

{Precondition} Program {Postcondition}
 Source of idea of loop invariant

Denotational Semantics

What it is:

- Construct function *M* assigning mathematical meaning to each program construct
 - via category theory, algebra, probability theory, topology, lambda calculus, ...
- Meaning function is compositional: meaning of construct built from meaning of parts

Tradeoffs:

- Useful for proving properties of programs
- Doesn't help much with implementation

Denotational Semantics

What it is:

- Construct function *M* assigning mathematical meaning to each program construct
 - via category theory, algebra, probability theory, topology, lambda calculus, ...
- Meaning function is compositional: meaning of construct built from meaning of parts

- Useful for **proving** properties of programs
- Doesn't help much with implementation

What it is:

- Describe how to execute (implement) programs of language on a virtual machine, by describing how to execute each program statement (i.e., following the structure of the program)
 Meaning of program is how its execution changes the state of the machine Tradeoffs:
- Easy to implement
- Hard to reason about abstractly (without thinking about implementation details)

- Can be small step or big step
 - Small step: define meaning of one step of execution of a program statement at a time
 - Big step: define meaning in terms of value of execution of whole program statement
 - Common to have both and relate them

- Can be small step or big step
 - Small step: define meaning of one step of execution of a program statement at a time
 - **Big step:** define meaning in terms of value of execution of whole program statement
 - Common to have both and relate them

Natural (Big Step) Semantics

Natural Semantics

- Also known as Structural Operational Semantics or Big Step Semantics
- Provide value for a program by rules and derivations, similar to type derivations
- Rule conclusions look like:

(C, m) ↓ m' or (E, m) ↓ v

Natural Semantics

Natural Semantics

- Also known as Structural Operational Semantics or Big Step Semantics
- Provide value for a program by rules and derivations, similar to type derivations
- Rule conclusions look like:

(C, m) ↓ m' or (E, m) ↓ v

Natural Semantics

- $I \in Identifiers$
- $N \in Numerals$
- B ::= true | false | B & B | B or B | not B | E < E | E = E</p>
- E ::= N | I | E + E | E * E | E E | E | (E)
 C ::= skip | C; C | I := E |
 if B then C else C fi | while B do C od



- $I \in Identifiers$
- $N \in Numerals$
- B ::= true | false | B & B | B or B | not B | E < E | E = E</p>
- E ::= N | I | E + E | E * E | E E | E | (E)
 C ::= skip | C; C | I := E |
 if B then C else C fi | while B do C od

Natural Semantics

- $I \in Identifiers$
- $N \in Numerals$
- B ::= true | false | B & B | B or B | not B | E < E | E = E</p>
- E ::= N | I | E + E | E * E | E E | E | (E)
- C ::= skip | C; C | I := E | if B then C else C fi | while B do C od



- $I \in Identifiers$
- $N \in Numerals$
- B ::= true | false | B & B | B or B | not B | E < E | E = E
- E ::= N | I | E + E | E * E | E E | E | (E)
- C ::= skip | C; C | I := E | if B then C else C fi | while B do C od

Simple Imperative Language Semantics



True

(true, m) ↓ true

(false, m) ↓ false

Natural Semantics

Simple Imperative Language Semantics





(false, m) ↓ false

Natural Semantics

Simple Imperative Language Semantics






Simple Imperative Language Semantics (B, m) ↓ v

$\begin{array}{c} (B, m) \Downarrow false \\ (B \& B', m) \Downarrow false \end{array} \begin{array}{c} (B, m) \Downarrow true (B', m) \Downarrow b \\ (B \& B', m) \Downarrow b \end{array}$

 $\frac{(B, m) \Downarrow true}{(B \text{ or } B', m) \Downarrow true} \xrightarrow[(B, m) \lor false (B', m) \Downarrow b]{}_{Or-F}$

(B, m) ↓ true _{№t-т} (not B, m) ↓ false

Boolean combinators have the standard meaning $(B, m) \Downarrow false_{Not-F}$ (not B, m) \Downarrow true

Simple Imperative Language Semantics (B, m) ↓ v

 $\begin{array}{c} (B, m) \Downarrow \text{false}_{And-F} \\ (B \& B', m) \Downarrow \text{false} \end{array} \begin{array}{c} (B, m) \Downarrow \text{true} (B', m) \Downarrow b \\ (B \& B', m) \Downarrow b \end{array}$

 $\frac{(B, m) \Downarrow true}{(B \text{ or } B', m) \Downarrow true} \xrightarrow[(B, m) \lor false (B', m) \Downarrow b]{(B \text{ or } B', m) \lor b}$

(B, m) ↓ true _{№t-т} (not B, m) ↓ false

Boolean combinators have the **standard** meaning

 $(B, m) \Downarrow false_{Not-F}$ (not B, m) \Downarrow true

Simple Imperative Language Semantics (B, m) ↓ v

 $\begin{array}{c} (B, m) \Downarrow false \\ (B \& B', m) \Downarrow false \end{array} \begin{array}{c} (B, m) \Downarrow true (B', m) \Downarrow b \\ (B \& B', m) \Downarrow b \end{array}$

 $\frac{(B, m) \Downarrow true}{(B \text{ or } B', m) \Downarrow true} \quad \frac{(B, m) \Downarrow false (B', m) \Downarrow b}{(B \text{ or } B', m) \Downarrow b}$

 $(B, m) \Downarrow true_{Not-T}$ (not B, m) \Downarrow false

Boolean combinators have the **standard** meaning

 $\frac{(B, m) \Downarrow false}{(not B, m) \Downarrow true}$

Simple Imperative Language Semantics (E, m) ↓ v

$\begin{array}{ccc} (\mathsf{E},\,\mathsf{m}) \Downarrow \mathsf{U} & (\mathsf{E}',\,\mathsf{m}) \Downarrow \mathsf{V} & \mathsf{U} \sim \mathsf{V} = \mathsf{b} \\ & & (\mathsf{E} \sim \mathsf{E}',\,\mathsf{m}) \Downarrow \mathsf{b} \end{array}$

- By U ~ V = b, we mean: does (the meaning of) the relation ~ hold on the meaning of U and V?
- May be specified by a mathematical expression/equation or rules matching U and V

Simple Imperative Language Semantics ((E, m) ↓ v

$$\begin{array}{ccc} (\mathsf{E},\,\mathsf{m}) \Downarrow \mathsf{U} & (\mathsf{E}',\,\mathsf{m}) \Downarrow \mathsf{V} & \mathbf{U} \sim \mathbf{V} = \mathbf{b} \\ & (\mathsf{E} \sim \mathsf{E}',\,\mathsf{m}) \Downarrow \mathsf{b} \end{array}$$

- By U ~ V = b, we mean: does (the meaning of) the relation ~ hold on the meaning of U and V?
- May be specified by a mathematical expression/equation or rules matching U and V

Simple Imperative Language Semantics ((E, m) ↓ v

$$(\mathbf{E}, \mathbf{m}) \Downarrow \mathbf{U} \quad (\mathbf{E}', \mathbf{m}) \Downarrow \mathbf{V} \quad \mathbf{U} \sim \mathbf{V} = \mathbf{b}_{\mathsf{Rel}}$$
$$(\mathbf{E} \sim \mathbf{E}', \mathbf{m}) \Downarrow \mathbf{b}$$

- By U ~ V = b, we mean: does (the meaning of) the relation ~ hold on the meaning of U and V?
- May be specified by a mathematical expression/equation or rules matching U and V

Simple Imperative Language Semantics (E, m) ↓ v

$(E, m) \Downarrow U \quad (E', m) \Downarrow V \quad U \sim V = b_{Rel}$ $(E \sim E', m) \Downarrow b$

- By U ~ V = b, we mean: does (the meaning of) the relation ~ hold on the meaning of U and V?
- May be specified by a mathematical expression/equation or rules matching U and V

Simple Imperative Language Semantics ((E, m) ↓ v

$$\begin{array}{ccc} (\mathsf{E},\,\mathsf{m}) \Downarrow \mathsf{U} & (\mathsf{E}',\,\mathsf{m}) \Downarrow \mathsf{V} & \mathbf{U} \sim \mathbf{V} = \mathbf{b} \\ & (\mathsf{E} \sim \mathsf{E}',\,\mathsf{m}) \Downarrow \mathsf{b} \end{array}$$

- By U ~ V = b, we mean: does (the meaning of) the relation ~ hold on the meaning of U and V?
- May be specified by a mathematical expression/equation or rules matching U and V

Simple Imperative Language Semantics (E, m) ↓ v

$\begin{array}{ccc} (\mathsf{E},\,\mathsf{m}) \Downarrow \mathsf{U} & (\mathsf{E}',\,\mathsf{m}) \Downarrow \mathsf{V} & \mathsf{U} \text{ op } \mathsf{V} = \mathsf{N} \\ & & \mathsf{(E} \text{ op } \mathsf{E}',\,\mathsf{m}) \Downarrow \mathsf{N} \end{array}$

where **N** is the specified value for **U op V**

Arithmetic expressions are defined similarly

$\begin{array}{ccc} (\mathsf{E},\,\mathsf{m}) \Downarrow \mathsf{U} & (\mathsf{E}',\,\mathsf{m}) \Downarrow \mathsf{V} & \mathbf{U} \, \mathbf{op} \, \mathbf{V} = \mathbf{N} \\ & & (\mathsf{E} \, \mathsf{op} \, \mathsf{E}',\,\mathsf{m}) \Downarrow \mathsf{N} \end{array}$

where **N** is the specified value for **U op V**

Arithmetic expressions are defined similarly







 $\frac{(\mathsf{C},\mathsf{m})\Downarrow\mathsf{m}'\ (\mathsf{C}',\mathsf{m}')\Downarrow\mathsf{m}''}{(\mathsf{C};\mathsf{C}',\mathsf{m})\Downarrow\mathsf{m}''}$

Natural Semantics



$$(C; C', m) \Downarrow m'$$













If then else is split into two cases, one for true and one for false



(B, m) \Downarrow true (C, m) \Downarrow m' (if B then C else C' fi, m) \Downarrow m'

(B, m) \Downarrow false (C', m) \Downarrow m' (if B then C else C' fi, m) \Downarrow m'

If then else is split into two cases, one for true and one for false



(B, m) \Downarrow **true** (**C**, m) \Downarrow m' (if B then **C** else C' fi, m) \Downarrow m'

(B, m) \Downarrow false (C', m) \Downarrow m' (if B then C else C' fi, m) \Downarrow m'

If then else is split into two cases, one for true and one for false

(B, m) \Downarrow **true** (**C**, m) \Downarrow m' (if B then **C** else C' fi, m) \Downarrow m'

(B, m)
$$\Downarrow$$
 false (C', m) \Downarrow m'
(if B then C else C' fi, m) \Downarrow m'

Simple Imperative Language Semantics (C, m) ↓ m′ (B, m) \Downarrow false While-F (while B do C od, m) \downarrow m (B, m) ↓ true (C, m) ↓ m′ (while B do C od, m') ↓ m" while-T (while B do C od, m) \downarrow m"

While is likewise split into two cases, one for **true** and one for **false**

Simple Imperative Language Semantics (C, m) ↓ m′ (B, m) ↓ false While-F (while **B** do C od, m) \downarrow m (B, m) ↓ true (C, m) ↓ m′ (while B do C od, m') ↓ m" while-T (while B do C od, m) \downarrow m"

While is likewise split into two cases, one for **true** and one for **false**

Simple Imperative Language Semantics (C, m) ↓ m′ (B, m) ↓ false While-F (while **B** do C od, **m**) \Downarrow **m** (B, m) ↓ true (C, m) ↓ m′ (while B do C od, m') ↓ m" while-T (while B do C od, m) \downarrow m"

While is likewise split into two cases, one for **true** and one for **false**

Simple Imperative Language Semantics (C, m) ↓ m′ (B, m) \Downarrow false While-F (while B do C od, m) \Downarrow m (B, m) ↓ true (C, m) ↓ m' (while B do C od, m') \Downarrow m" _{while-T} (while **B** do C od, **m**) \Downarrow m"

While is likewise split into two cases, one for **true** and one for **false**

Simple Imperative Language Semantics (C, m) ↓ m' (B, m) \Downarrow false While-F (while B do C od, m) \downarrow m (B, m) ↓ true (C, m) **↓** m′ (while B do C od, m') \Downarrow m'' _{while-T} (while B do C od, m) \Downarrow m" While is likewise split into two cases,

one for **true** and one for **false**

Simple Imperative Language Semantics (C, m) ↓ m' (B, m) \Downarrow false While-F (while B do C od, m) \Downarrow m (B, m) ↓ true (C, m) ↓ **m'** (while B do C od, m') U m"_{While-T} (while B do C od, m) \Downarrow m" While is likewise split into two cases, one for true and one for false

Simple Imperative Language Semantics (C, m) ↓ m′ (B, m) \Downarrow false While-F (while B do C od, m) \Downarrow m (B, m) ↓ true (C, m) ↓ m' (while B do C od, m') ↓ m''_{while-T} (while B do C od, m) \Downarrow **m**" While is likewise split into two cases,

one for **true** and one for **false**





Want to determine the **semantics** of this command, using the **natural semantics** for the language that we just defined.

(if x > 5 then y := 2 + 3 else y := 3 + 4 fi, $\{x \rightarrow 7\}$) \Downarrow ??



First, **if-then-else rule**, but we don't know if the guard is **true** or **false** yet.

(if x > 5 then y := 2 + 3 else y := 3 + 4 fi,
$$\{x \rightarrow 7\}$$
) \Downarrow ??



First, **if-then-else rule**, but we don't know if the guard is **true** or **false** yet.

$$\frac{(x > 5, \{x -> 7\}) \Downarrow ??}{(if x > 5 then y := 2 + 3 else y := 3 + 4 fi, \{x -> 7\}) \Downarrow ??}$$



The guard is a **relation**.

$$\frac{(x > 5, \{x -> 7\}) \Downarrow ??}{(if x > 5 then y := 2 + 3 else y := 3 + 4 fi, \{x -> 7\}) \Downarrow ??}$$


The guard is a **relation**.

$$\begin{array}{cccc} (x, \{x - > 7\}) \Downarrow \ref{eq: x->7}) \ \ref{eq: x->7}$$



So we determine the meaning of **each side** of the **relation** ...

$$(\mathbf{x}, \{x ->7\}) \Downarrow \ref{x} (\mathbf{5}, \{x ->7\}) \Downarrow \ref{x} ?? > \ref{x} = \ref{x} Rel$$

$$(\mathbf{x} > \mathbf{5}, \{x ->7\}) \Downarrow \ref{x} ??$$

$$(if x > 5 then y := 2 + 3 else y := 3 + 4 fi, \{x ->7\}) \Downarrow \ref{x}$$

















If-T



Rel

 (x > 5, {x -> 7})
$$\Downarrow$$
 true

 (if x > 5 then y := 2 + 3 else y := 3 + 4 fi, {x -> 7}) \Downarrow ??



We need the meaning of the **if** branch, not the **else** branch

$$\begin{array}{cccc}
 & & & & & & & & & & \\ \hline (x > 5, \{x -> 7\}) \Downarrow & & & & & & & \\ \hline (\text{if } x > 5 & & & & & \\ \hline (\text{if } x > 5 & & & & & \\ \hline \{x -> 7\}) \Downarrow & & & \\ \hline \end{array}$$



















Awkward Example



Where m" (y) = m' (y) for $y \neq I$ and m" (I) = m (I) if m(I) is defined, and m" (I) is undefined otherwise

Let in Command

$$\begin{array}{c} (x,\{x->5\}) \Downarrow 5 \quad (3,\{x->5\}) \Downarrow 3 \\ (x+3,\{x->5\}) \Downarrow 8 \\ (5,\{x->17\}) \Downarrow 5 \quad (x:=x+3,\{x->5\}) \Downarrow \{x->8\} \\ \hline (\text{let } x = 5 \text{ in } (x:=x+3), \{x ->17\}) \Downarrow \ref{eq: constraint} \end{array}$$

Awkward Example ₉₇

Let in Command

$$\begin{array}{c} (x,\{x->5\}) \Downarrow 5 \quad (3,\{x->5\}) \Downarrow 3 \\ \hline (x+3,\{x->5\}) \Downarrow 8 \\ \hline (5,\{x->17\}) \Downarrow 5 \quad (x:=x+3,\{x->5\}) \Downarrow \{x->8\} \\ \hline (\text{let } x = 5 \text{ in } (x:=x+3), \{x ->17\}) \Downarrow \{x->17\} \end{array}$$

Awkward Example ₉₈

Comment

- Simple Imperative Programming Language introduces variables **implicitly** through assignment
- The let-in command introduces scoped variables explicitly
- Clash of constructs apparent in awkward semantics





Interpretation Versus Compilation

- A compiler from language L1 to language L2 is a program that takes an L1 program and for each piece of code in L1 generates a piece of code in L2 of same meaning
- An interpreter of L1 in L2 is an L2 program that executes the meaning of a given L1 program
- Compiler would examine the body of a loop once; an interpreter would examine it every time the loop was executed

Interpretation Versus Compilation

- A compiler from language L1 to language L2 is a program that takes an L1 program and for each piece of code in L1 generates a piece of code in L2 of same meaning
- An interpreter of L1 in L2 is an L2 program that executes the meaning of a given L1 program
- Compiler would examine the body of a loop once; an interpreter would examine it every time the loop was executed

Interpreter

- An Interpreter represents the operational semantics of a language L1 (source language) in the language of implementation L2 (target language)
- Built incrementally
 - Start with literals
 - Variables
 - Primitive operations
 - Evaluation of expressions
 - Evaluation of commands/declarations

Interpreter

- Takes abstract syntax trees as input
 - In simple cases could be just strings
- One procedure for each syntactic category (nonterminal)
 - e.g., one for expressions, another for commands
- From semantics to implementation:
 - If Natural Semantics used, tells how to compute final value from code
 - If Transition Semantics used, tells how to compute next "state"
 - To get final value, put in a loop

Interpreter

- Takes abstract syntax trees as input
 - In simple cases could be just strings
- One procedure for each syntactic category (nonterminal)
 - e.g., one for expressions, another for commands
- From semantics to implementation:
 - If Natural Semantics used, tells how to compute final value from code
 - If Transition Semantics used, tells how to compute next "state"
 - To get final value, put in a loop

Natural Semantics Example

- compute_exp (Var(v), m) = look_up v m
 compute_exp (Int(n), _) = Num (n)
- compute_com(IfExp(b,c1,c2),m) =
 if compute_exp (b,m) = Bool(true)
 then compute_com (c1,m)
 else compute_com (c2,m)

Natural Semantics Example

compute_com(While(b,c), m) =
 if compute_exp (b,m) = Bool(false)
 then m
 else compute_com
 (While(b,c), compute_com(c,m))

May fail to terminate - exceed stack limitsReturns no useful information then


No Class Thursday for Midterm!