Introduction

Operational semantics
Models precisely program execution as low-level transitions between internal states (transition systems, execution traces, big-step semantics)

Denotational semantics
Maps programs into objects in a mathematical domain (higher level, compositional, domain oriented)

Aximoatic semantics (today)
Prove properties about programs
- programs are annotated with logical assertions
- a rule-system defines the validity of assertions (logical proofs)
- clearly separates programs from specifications (specification $\simeq$ user-provided abstraction of the behavior, it is not unique)
- enables the use of logic tools (partial automation)
Overview

- Specifications (informal examples)

- Floyd–Hoare logic

- Dijkstra’s predicate transformers
  (weakest precondition, strongest postcondition)

- Verification conditions
  (partially automated program verification)

- Advanced topics
  - Auxiliary variables
  - Non-determinism
  - Total correctness (termination)
  - Arrays
Specifications
int mod(int A, int B) {
    int Q = 0;
    int R = A;
    while (R >= B) {
        R = R - B;
        Q = Q + 1;
    }
    return R;
}
Example: function specification

```c
//@ ensures \result == A mod B;
int mod(int A, int B) {
    int Q = 0;
    int R = A;
    while (R >= B) {
        R = R - B;
        Q = Q + 1;
    }
    return R;
}
```

- express the intended behavior of the function (returned value)
Example: function specification

```c
//@ requires A>=0 && B>=0;
//@ ensures \result == A mod B;
int mod(int A, int B) {
    int Q = 0;
    int R = A;
    while (R >= B) {
        R = R - B;
        Q = Q + 1;
    }
    return R;
}
```

- express the intended behavior of the function (returned value)
- add requirements for the function to actually behave as intended (a requires/ensures pair is a function contract)
Example: function specification

Example in C + ACSL

```c
//@ requires A>=0 && B>0;
//@ ensures \result == A mod B;
int mod(int A, int B) {
    int Q = 0;
    int R = A;
    while (R >= B) {
        R = R - B;
        Q = Q + 1;
    }
    return R;
}
```

- Express the intended behavior of the function (returned value)
- Add requirements for the function to actually behave as intended (a requires/ensures pair is a function contract)
- Strengthen the requirements to ensure termination
Example: program annotations

```c
//@ requires A>=0 && B>0;
//@ ensures \result == A mod B;
int mod(int A, int B) {
    int Q = 0;
    int R = A;
    //@ assert A>=0 && B>0 && Q=0 && R==A;
    while (R >= B) {
        //@ assert A>=0 && B>0 && R>=B && A==Q*B+R;
        R = R - B;
        Q = Q + 1;
    }
    //@ assert A>=0 && B>0 && R>=0 && R<B && A==Q*B+R;
    return R;
}
```

Assertions give detail about the internal computations why and how contracts are fulfilled

(Note: $r = a \mod b$ means $\exists q: a = qb + r \land 0 \leq r < b$)
Example: ghost variables

```c
//@ requires A>=0 && B>0;
//@ ensures \result == A mod B;
int mod(int A, int B) {
    int R = A;
    while (R >= B) {
        R = R - B;
    }
    // ∃Q: A = QB + R and 0 ≤ R < B
    return R;
}
```

Program annotations can be more complex than the program.
 Specifications

Example: ghost variables

def mod(A, B):
    A >= 0 and B > 0
    result == A mod B
    R = A
    assert A >= 0 and B > 0 and q = 0 and R == A
    while R >= B:
        assert A >= 0 and B > 0 and R >= B and A == q * B + R
        R = R - B
        q = q + 1
    assert A >= 0 and B > 0 and R >= 0 and R < B and A == q * B + R
    return R

Program annotations can be more complex than the program and require reasoning on enriched states (ghost variables)
Example: class invariants

class invariant: property of the fields true outside all methods
it can be temporarily broken within a method
but it must be restored before exiting the method
Contracts (and class invariants):

- built in few languages  
  (Eiffel)
- available as a library / external tool  
  (C, Java, C#, etc.)

Contracts can be:

- checked dynamically
- checked statically  
  (Frama-C, Why, ESC/Java)
- inferred statically  
  (CodeContracts)

**In this course:**

deductive methods (logic) to check (prove) statically (at compile-time)
partially automatically (with user help) that contracts hold
Floyd–Hoare logic
**Hoare triples**

A **Hoare triple** is defined as:

\[
\{ P \} \text{prog} \{ Q \}
\]

- \( P \) and \( Q \) are logical assertions over program variables.
- \( P \) is the precondition,
- \( Q \) is the postcondition.

A triple means:
- if \( P \) holds before \( \text{prog} \) is executed,
- then \( Q \) holds after the execution of \( \text{prog} \),
- unless \( \text{prog} \) does not terminate or encounters an error.

\( \{ P \} \text{prog} \{ Q \} \) expresses partial correctness

\( \{ P \} \text{prog} \{ Q \} \) expresses partial correctness

(\( \{ P \} \text{prog} \{ Q \} \) expresses partial correctness.

Hoare triples serve as judgements in a proof system

(introduced in [Hoare69])
Language

\[
\text{stat} \quad ::= \quad X \leftarrow \text{expr} \quad \text{(assignment)} \\
| \quad \text{skip} \quad \text{(do nothing)} \\
| \quad \text{fail} \quad \text{(error)} \\
| \quad \text{stat; stat} \quad \text{(sequence)} \\
| \quad \text{if expr then stat else stat} \quad \text{(conditional)} \\
| \quad \text{while expr do stat} \quad \text{(loop)}
\]

- \( X \in \mathbb{V} \): integer-valued variables
- \( \text{expr} \): integer arithmetic expressions

we assume that:
  - expressions are deterministic (for now)
  - expression evaluation do not cause error

for instance, to avoid division by zero, we can:
  - either define \( 1/0 \) to be a valid value, such as 0
  - or systematically guard divisions
    (e.g.: \( \text{if } X = 0 \text{ then fail else } \cdots /X \cdots \) )
**Hoare rules: axioms**

**Axioms:**

\[
\begin{align*}
\{P\} \text{ skip } \{P\} & \quad \{P\} \text{ fail } \{Q\}
\end{align*}
\]

- any property true before \textbf{skip} is true afterwards
- any property is true after \textbf{fail}
Assignment axiom:

\[
\{P[e/X]\} \; X \leftarrow e \; \{P\}
\]

for \(P\) over \(X\) to be true after \(X \leftarrow e\)

\(P\) must be true over \(e\) before the assignment

- \(P[e/X]\) is \(P\) where free occurrences of \(X\) are replaced with \(e\)
- \(e\) must be deterministic
- the rule is “backwards” \((P\) appears as a postcondition)

examples:

\[
\{true\} \; X \leftarrow 5 \; \{X = 5\}
\]

\[
\{Y = 5\} \; X \leftarrow Y \; \{X = 5\}
\]

\[
\{X + 1 \geq 0\} \; X \leftarrow X + 1 \; \{X \geq 0\}
\]

\[
\{false\} \; X \leftarrow Y + 3 \; \{Y = 0 \land X = 12\}
\]

\[
\{Y \in [0, 10]\} \; X \leftarrow Y + 3 \; \{X = Y + 3 \land Y \in [0, 10]\}
\]
Floyd–Hoare logic

Hoare rules: consequence

**Rule of consequence:**

\[
\begin{align*}
P \Rightarrow P' & \quad Q' \Rightarrow Q \\
\{P'\} \ c \ \{Q'\} & \quad \{P\} \ c \ \{Q\}
\end{align*}
\]

we can weaken a Hoare triple by:

- weakening its postcondition \( Q \leftarrow Q' \)
- strengthening its precondition \( P \Rightarrow P' \)

we assume a logic system to be available to prove formulas on assertions, such as \( P \Rightarrow P' \) (e.g., arithmetic, set theory, etc.)

examples:

- the axiom for `fail` can be replaced with \( \{\text{true}\} \ \text{fail} \ \{\text{false}\} \) (as \( P \Rightarrow \text{true} \) and \( \text{false} \Rightarrow Q \) always hold)
- \( \{X = 99 \land Y \in [1, 10]\} \ \ X \leftarrow Y + 10 \ \{X = Y + 10 \land Y \in [1, 10]\} \)
  (as \( \{Y \in [1, 10]\} \ \ X \leftarrow Y + 10 \ \{X = Y + 10 \land Y \in [1, 10]\} \))
Hoare rules: tests

Tests:

\[
\begin{align*}
\{P \land e\} & \quad s \quad \{Q\} \quad \quad \{P \land \neg e\} & \quad t \quad \{Q\} \\
\{P\} & \quad \text{if } e \text{ then } s \text{ else } t \quad \{Q\}
\end{align*}
\]

to prove that \( Q \) holds after the test

we prove that it holds after each branch \((s, t)\)

under the assumption that it is executed \((e, \neg e)\)

example:

\[
\begin{align*}
\{X < 0\} & \quad X \leftarrow -X \quad \{X > 0\} \\
\{(X \neq 0) \land (X < 0)\} & \quad X \leftarrow -X \quad \{X > 0\}
\end{align*}
\]

\[
\begin{align*}
\{X > 0\} & \quad \text{skip} \quad \{X > 0\} \\
\{(X \neq 0) \land (X \geq 0)\} & \quad \text{skip} \quad \{X > 0\}
\end{align*}
\]

\[
\begin{align*}
\{X \neq 0\} & \quad \text{if } X < 0 \text{ then } X \leftarrow -X \text{ else } \text{skip} \quad \{X > 0\}
\end{align*}
\]
Hoare rules: sequences

Sequences:

\[
\begin{array}{c}
\{P\} \ s \ \{R\} \ \{R\} \ t \ \{Q\} \\
\{P\} \ s; \ t \ \{Q\}
\end{array}
\]

to prove a sequence \(s; t\)

we must invent an intermediate assertion \(R\)

implied by \(P\) after \(s\), and implying \(Q\) after \(t\)

(often denoted \(\{P\} \ s \ \{R\} \ t \ \{Q\}\))

eample:

\[
\{X = 1 \land Y = 1\} \ X \gets X + 1 \ \{X = 2 \land Y = 1\} \ Y \gets Y - 1 \ \{X = 2 \land Y = 0\}
\]
Floyd–Hoare logic

**Hoare rules: loops**

Loops:

\[
\begin{align*}
\{P \land e\} \ s \ \{P\} \\
\{P\} \textbf{ while } e \textbf{ do } s \ \{P \land \neg e\}
\end{align*}
\]

\(P\) is a **loop invariant**

\(P\) holds before each loop iteration, before even testing \(e\)

**Practical use:**

actually, we would rather prove the triple: \(\{P\} \textbf{ while } e \textbf{ do } s \ \{Q\}\)

it is sufficient to **invent an assertion** \(I\) that:

- holds when the loop start: \(P \Rightarrow I\)
- is invariant by the body \(s\): \(\{I \land e\} \ s \ \{I\}\)
- implies the assertion when the loop stops: \((I \land \neg e) \Rightarrow Q\)

we can derive the rule:

\[
\begin{align*}
P \Rightarrow I & \quad I \land \neg e \Rightarrow Q \\
\{I\} \textbf{ while } e \textbf{ do } s \ \{I \land \neg e\}
\end{align*}
\]

\(\{P\} \textbf{ while } e \textbf{ do } s \ \{Q\}\)
Hoare rules: logical part

Hoare logic is parameterized by the choice of logical theory of assertions. The logical theory is used to:

- **prove** properties of the form $P \Rightarrow Q$ (rule of consequence)
- **simplify** formulas (replace a formula with a simpler one, equivalent in a logical sense: $\Leftrightarrow$)

**Examples:** (generally first order theories)
- booleans ($B, \neg, \land, \lor$)
- bit-vectors ($B^n, \neg, \land, \lor$)
- Presburger arithmetic ($\mathbb{N}, +$)
- Peano arithmetic ($\mathbb{N}, +, \times$)
- linear arithmetic on $\mathbb{R}$
- Zermelo-Fraenkel set theory ($\in, \{\}$)
- theory of arrays (lookup, update)

Theories have different expressiveness, decidability and complexity results. This is an important factor when trying to automate program verification.
Floyd–Hoare logic

Hoare rules: summary

\[ \{P\} \textbf{skip} \{P\} \]
\[ \{\text{true}\} \textbf{fail} \{\text{false}\} \]
\[ \{P[e/X]\} X \leftarrow e \{P\} \]

\[ \{P\} \ s \ \{R\} \quad \{R\} \ t \ \{Q\} \]
\[ \{P\} \ s; \ t \ \{Q\} \]

\[ \{P \land e\} \ s \ \{Q\} \quad \{P \land \neg e\} \ t \ \{Q\} \]
\[ \{P\} \ \textbf{if} \ e \ \textbf{then} \ s \ \textbf{else} \ t \ \{Q\} \]

\[ \{P \land e\} \ s \ \{P\} \]
\[ \{P\} \ \textbf{while} \ e \ \textbf{do} \ s \ \{P \land \neg e\} \]

\[ P \Rightarrow P' \quad Q' \Rightarrow Q \quad \{P'\} \ c \ \{Q'\} \]
\[ \{P\} \ c \ \{Q\} \]
Proof tree example

\[ s \overset{\text{def}}{=} \textbf{while } I < N \textbf{ do } (X \leftarrow 2X; \ I \leftarrow I + 1) \]

\[
\begin{array}{c}
\text{C} \quad \{P_3\} \ X \leftarrow 2X \quad \{P_2\} \\
\{P_1 \land I < N\} \ X \leftarrow 2X; \ I \leftarrow I + 1 \quad \{P_1\}
\end{array}
\]

\[
\begin{array}{c}
A \quad B \\
\{P_1\} \ s \{P_1 \land I \geq N\} \\
\{X = 1 \land I = 0 \land N \geq 0\} \ s \{X = 2^N \land N = I \land N \geq 0\}
\end{array}
\]

\[
\begin{align*}
P_1 & \overset{\text{def}}{=} X = 2^I \land I \leq N \land N \geq 0 \\
P_2 & \overset{\text{def}}{=} X = 2^{I+1} \land I+1 \leq N \land N \geq 0 \\
P_3 & \overset{\text{def}}{=} 2X = 2^{I+1} \land I+1 \leq N \land N \geq 0 \\ & \overset{\text{equiv}}{=} X = 2^I \land I < N \land N \geq 0
\end{align*}
\]

\[
\begin{align*}
A &: \ (X = 1 \land I = 0 \land N \geq 0) \Rightarrow P_1 \\
B &: \ (P_1 \land I \geq N) \Rightarrow (X = 2^N \land N = I \land N \geq 0) \\
C &: \ P_3 \iff (P_1 \land I < N)
\end{align*}
\]
Proof tree example

\[ s \overset{\text{def}}{=} \text{while } l \neq 0 \text{ do } l \leftarrow l - 1 \]

\[
\begin{array}{c}
\{ \text{true} \} \ l \leftarrow l - 1 \ \{ \text{true} \} \\
\{ l \neq 0 \} \ l \leftarrow l - 1 \ \{ \text{true} \} \\
\{ \text{true} \} \ \text{while } l \neq 0 \text{ do } l \leftarrow l - 1 \ \{ \text{true} \land \neg l \neq 0 \} \\
\{ \text{true} \} \ \text{while } l \neq 0 \text{ do } l \leftarrow l - 1 \ \{ l = 0 \}
\end{array}
\]

- in some cases, the program does not terminate
  (if the program starts with \( l < 0 \))
- the same proof holds for: \( \{ \text{true} \} \ \text{while } l \neq 0 \text{ do } J \leftarrow J - 1 \ \{ l = 0 \} \)
- anything can be proven of a program that never terminates: \( \{ l = 1 \land l \neq 0 \} \ J \leftarrow J - 1 \ \{ l = 1 \} \)
  \( \{ l = 1 \} \ \text{while } l \neq 0 \text{ do } J \leftarrow J - 1 \ \{ l = 1 \land l = 0 \} \)
  \( \{ l = 1 \} \ \text{while } l \neq 0 \text{ do } J \leftarrow J - 1 \ \{ \text{false} \} \)
Example: we wish to prove:

\[ \{ X = Y = 0 \} \textbf{while} X < 10 \textbf{ do } (X \leftarrow X + 1; \; Y \leftarrow Y + 1) \{ X = Y = 10 \} \]

we need to find an invariant assertion \( P \) for the \textbf{while} rule

**Incorrect invariant:** \( P \overset{\text{def}}{=} X, Y \in [0, 10] \)

- \( P \) indeed holds at each loop iteration \((P \text{ is an invariant})\)
- but \( \{ P \land (X < 10) \} \; X \leftarrow X + 1; \; Y \leftarrow Y + 1 \{ P \} \)
  does not hold

\( P \land X < 10 \) does not prevent \( Y = 10 \) after \( Y \leftarrow Y + 1 \), \( P \) does not hold anymore
Invariants and inductive invariants

Example: we wish to prove:

$\{X = Y = 0\} \textbf{while } X < 10 \textbf{ do } (X \leftarrow X + 1; \ Y \leftarrow Y + 1) \{X = Y = 10\}$

we need to find an invariant assertion $P$ for the \textbf{while} rule

\textbf{Correct invariant: } $P' \overset{\text{def}}{=} X \in [0, 10] \land X = Y$

- $P'$ also holds at each loop iteration \hspace{1cm} \textit{(P' is an invariant)}
- $\{P' \land (X < 10)\} X \leftarrow X + 1; \ Y \leftarrow Y + 1 \{P'\}$ can be proven
- $P'$ is an \textit{inductive invariant} \hspace{1cm} \textit{(passes to the induction, stable by a loop iteration)}

$\implies$ to prove a loop invariant
- it is often necessary to find a \textit{stronger} loop invariant
Soundness and completeness

Validity:
\[ \{ P \} c \{ Q \} \text{ is valid } \iff \text{executions starting in a state satisfying } P \]
\[ \text{and terminating} \]
\[ \text{end in a state satisfying } Q \]

(it is an operational notion)

- **soundness**
  a proof tree exists for \( \{ P \} c \{ Q \} \Rightarrow \{ P \} c \{ Q \} \) is valid

- **completeness**
  \( \{ P \} c \{ Q \} \) is valid \( \Rightarrow \) a proof tree exists for \( \{ P \} c \{ Q \} \)

(technically, by Gödel’s incompleteness theorem, \( P \Rightarrow Q \) is not always provable for strong theories; hence, Hoare logic is incomplete; we consider relative completeness by adding all valid properties \( P \Rightarrow Q \) on assertions as axioms)

**Theorem (Cook 1974)**
Hoare logic is sound (and relatively complete)

Completeness no longer holds for more complex languages (Clarke 1976)
Link with denotational semantics

Reminder: \( S[\text{ stat }] : \mathcal{P}(\mathcal{E}) \rightarrow \mathcal{P}(\mathcal{E}) \) where \( \mathcal{E} \overset{\text{def}}{=} \forall \mapsto \bot \)

\[ S[\text{ skip }] R \overset{\text{def}}{=} R \]

\[ S[\text{ fail }] R \overset{\text{def}}{=} \emptyset \]

\[ S[ s_1 ; s_2 ] \overset{\text{def}}{=} S[ s_2 ] \circ S[ s_1 ] \]

\[ S[ X \gets e ] R \overset{\text{def}}{=} \{ \rho[X \mapsto v] \mid \rho \in R, \ v \in E[ e ] \rho \} \]

\[ S[ \text{ if } e \text{ then } s_1 \text{ else } s_2 ] R \overset{\text{def}}{=} S[ s_1 ] \{ \rho \in R \mid \text{true} \in E[ e ] \rho \} \cup S[ s_2 ] \{ \rho \in R \mid \text{false} \in E[ e ] \rho \} \]

\[ S[ \text{ while } e \text{ do } s ] R \overset{\text{def}}{=} \{ \rho \in \text{Ifp } F \mid \text{false} \in E[ e ] \rho \} \]

where \( F(X) \overset{\text{def}}{=} R \cup S[ s ] \{ \rho \in X \mid \text{true} \in E[ e ] \rho \} \)

Theorem

\[ \{ P \} c \{ Q \} \overset{\text{def}}{\iff} \forall R \subseteq \mathcal{E} : R \models P \implies S[ c ] R \models Q \]

\((A \models P \text{ means } \forall \rho \in A, \text{ the formula } P \text{ is true on the variable assignment } \rho)\)
Floyd–Hoare logic

Link with denotational semantics

- Hoare logic reasons on formulas
- Denotational semantics reasons on state sets

We can assimilate assertion formulas and state sets
(logical abuse: we assimilate formulas and models)

Let \([R]\) be any formula representing the set \(R\), then:

- \([R]\) \(\subseteq\) \([S[:c:]R]\) is always valid
- \([R]\) \(\subseteq\) \([R']\) \(\Rightarrow\) \(S[:c:]R \subseteq R'\)

\(\implies\) \([S[:c:]R]\) provides the best valid postcondition
Floyd–Hoare logic

Link with denotational semantics

**Loop invariants**

- **Hoare:**
  to prove \(\{P\} \textbf{while} e \textbf{do} s \{P \land \neg e\}\) we must prove \(\{P \land e\} s \{P\}\)
i.e., \(P\) is an inductive invariant

- **Denotational semantics:**
  we must find \(\text{lfp } F\) where \(F(X) \overset{\text{def}}{=} R \cup S[ s ] \{ \rho \in X \mid \rho \models e \}\)
  - \(\text{lfp } F = \cap \{ X \mid F(X) \subseteq X \}\) (Tarski’s theorem)
  - \(F(X) \subseteq X \iff ([R] \Rightarrow [X]) \land \{[X \land e]\} s \{[X]\}\)
    - \(R \subseteq X\) means \([R] \Rightarrow [X]\),
    - \(S[ s ] \{ \rho \in X \mid \rho \models e \}\) \(\subseteq X\) means \([X \land e]\) \(s \{[X]\}\)

As a consequence:

- any \(X\) such that \(F(X) \subseteq X\) gives an inductive invariant \([X]\)
- \(\text{lfp } F\) gives the best inductive invariant
- any \(X\) such that \(\text{lfp } F \subseteq X\) gives an invariant
  (not necessarily inductive)

(see [Cousot02])
Predicate transformers
Dijkstra’s weakest liberal preconditions

**Principle:**
- calculus to derive preconditions from postconditions
- order and mechanize the search for intermediate assertions
  (easier to go backwards, mainly due to assignments)

**Weakest liberal precondition** \( \text{wlp} : (\text{prog} \times \text{Prop}) \to \text{Prop} \)

\( \text{wlp}(c, P) \) is the weakest, i.e. most general, precondition ensuring that \( \{ \text{wlp}(c, P) \} \ c \ \{ P \} \) is a Hoare triple

(greatest state set that ensures that the computation ends up in \( P \))

formally: \[ \{ P \} \ c \ \{ Q \} \iff (P \Rightarrow \text{wlp}(c, Q)) \]

“liberal” means that we do not care about termination and errors

**Examples:**

\[
\begin{align*}
\text{wlp}(X \leftarrow X + 1, X = 1) &= \\
\text{wlp}(&\text{while } X < 0 X \leftarrow X + 1, X \geq 0) &= \\
\text{wlp}(&\text{while } X \neq 0 X \leftarrow X + 1, X \geq 0) &=
\end{align*}
\]

(introduced in [Dijkstra75])
**Predicate transformers**

**Dijkstra’s weakest liberal preconditions**

**Principle:**
- **calculus** to derive preconditions from postconditions
- order and mechanize the search for intermediate assertions
  (easier to go backwards, mainly due to assignments)

**Weakest liberal precondition** \( \text{wlp} : (\text{prog} \times \text{Prop}) \rightarrow \text{Prop} \)

\( \text{wlp}(c, P) \) is the weakest, i.e. **most general**, precondition ensuring that \( \{ \text{wlp}(c, P) \} \ c \ \{ P \} \) is a Hoare triple

(greatest state set that ensures that the computation ends up in \( P \))

formally:

\[ \{ P \} \ c \ \{ Q \} \iff (P \Rightarrow \text{wlp}(c, Q)) \]

“liberal” means that we do not care about termination and errors

**Examples:**

\[
\begin{align*}
\text{wlp}(X \leftarrow X + 1, \ X = 1) &= (X = 0) \\
\text{wlp}(\text{while } X < 0 \ X \leftarrow X + 1, \ X \geq 0) &= \text{true} \\
\text{wlp}(\text{while } X \neq 0 \ X \leftarrow X + 1, \ X \geq 0) &= \text{true}
\end{align*}
\]

(introduced in [Dijkstra75])
A calculus for \textit{wlp}

\textit{wlp} is defined by induction on the syntax of programs:

\begin{align*}
\text{\texttt{wlp}}(\texttt{skip}, P) & \overset{\text{def}}{=} P \\
\text{\texttt{wlp}}(\texttt{fail}, P) & \overset{\text{def}}{=} \text{true} \\
\text{\texttt{wlp}}(X \leftarrow e, P) & \overset{\text{def}}{=} P[e/X] \\
\text{\texttt{wlp}}(s; t, P) & \overset{\text{def}}{=} \text{\texttt{wlp}}(s, \text{\texttt{wlp}}(t, P)) \\
\text{\texttt{wlp}}(\texttt{if } e\texttt{ then } s \texttt{ else } t, P) & \overset{\text{def}}{=} (e \Rightarrow \text{\texttt{wlp}}(s, P)) \land (\neg e \Rightarrow \text{\texttt{wlp}}(t, P)) \\
\text{\texttt{wlp}}(\texttt{while } e\texttt{ do } s, P) & \overset{\text{def}}{=} I \land ((e \land I) \Rightarrow \text{\texttt{wlp}}(s, I)) \land ((\neg e \land I) \Rightarrow P)
\end{align*}

- \(e \Rightarrow Q\) is equivalent to \(Q \lor \neg e\)
  - weakest property that matches \(Q\) when \(e\) holds
  - but says nothing when \(e\) does not hold

- \textbf{while} loops require \textbf{providing} an \textbf{invariant predicate} \(I\)
  - intuitively, \(\text{\texttt{wlp}}\) checks that \(I\) is an inductive invariant implying \(P\)
  - if so, it returns \(I\); otherwise, it returns false

\(\text{\texttt{wlp}}\) is the weakest precondition only if \(I\) is well-chosen...
Alternate form for loops

Unrolling of the loop while \( e \) do \( s \):

- \( L_0 \overset{\text{def}}{=} \text{fail} \)
- \( L_{i+1} \overset{\text{def}}{=} \text{if } e \text{ then } (s; L_i) \text{ else skip} \)
- \( L_i \) runs the loop and fails after \( i \) iterations

we have:

\[
\begin{align*}
\text{wlp}(L_0, P) &= \text{true} \\
\text{wlp}(L_{i+1}, P) &= (e \Rightarrow \text{wlp}(s, \text{wlp}(L_i, P))) \land (\neg e \Rightarrow P)
\end{align*}
\]

Alternate \( \text{wlp} \) for loops:

\[\text{wlp(while } e \text{ do } s, P) \overset{\text{def}}{=} \forall i: X_i\]

where \( X_0 \overset{\text{def}}{=} \text{true} \)

\[X_{i+1} \overset{\text{def}}{=} (e \Rightarrow \text{wlp}(s, X_i)) \land (\neg e \Rightarrow P)\]

\( X_i \Rightarrow X_{i+1} \): sequence of assertions of increasing strength

\((\forall i: X_i)\) is the limit, with an arbitrary number of iterations

\((\forall i: X_i)\) is a closed form guaranteed to be the weakest precondition

\((\forall i: X_i)\) is the fixpoint of a second-order formula

\(\implies\) very difficult to handle
wlp(if \( X < 0 \) then \( Y \leftarrow -X \) else \( Y \leftarrow X, \ Y \geq 10 \) =

\((X < 0 \Rightarrow wlp(Y \leftarrow -X, Y \geq 10)) \land (X \geq 0 \Rightarrow wlp(Y \leftarrow X, Y \geq 10))\)

\((X < 0 \Rightarrow -X \geq 10) \land (X \geq 0 \Rightarrow X \geq 10) =\)

\((X \geq 0 \lor -X \geq 10) \land (X < 0 \lor X \geq 10) =\)

\(X \geq 10 \lor X \leq -10\)

\(wlp\) generates complex formulas

it is important to simplify them from time to time
Properties of \( wlp \)

- \( wlp(c, \text{false}) \equiv \text{false} \)  
  
  (excluded miracle)

- \( wlp(c, P) \land wlp(d, Q) \equiv wlp(c, P \land Q) \)  
  
  (distributivity)

- \( wlp(c, P) \lor wlp(d, Q) \equiv wlp(c, P \lor Q) \)  
  
  (distributivity)

  (\( \Rightarrow \) always true, \( \Leftarrow \) only true for deterministic, error-free programs)

- if \( P \Rightarrow Q \), then \( wlp(c, P) \Rightarrow wlp(c, Q) \)  
  
  (monotonicity)

\( A \equiv B \) means that the formulas \( A \) and \( B \) are equivalent

i.e., \( \forall \rho: \rho \models A \iff \rho \models B \)

(stronger than syntactic equality)
we can define $slp: (\text{Prop} \times \text{prog}) \to \text{Prop}$

- $\{ P \}\ c\ \{ slp(P, c) \}$
  (postcondition)

- $\{ P \}\ c\ \{ Q \} \iff (slp(P, c) \Rightarrow Q)$
  (strongest postcondition)

  (corresponds to the smallest state set)

- $slp(P, c)$ does not care about non-termination
  (liberal)

- allows forward reasoning

we have a duality:

$\ (P \Rightarrow wlp(c, Q)) \iff (slp(P, c) \Rightarrow Q)$

proof: $\ (P \Rightarrow wlp(c, Q)) \iff \{ P \}\ c\ \{ Q \} \iff (slp(P, c) \Rightarrow Q)$
Calculus for slp

\[
\text{slp}(P, \text{skip}) \overset{\text{def}}{=} P
\]

\[
\text{slp}(P, \text{fail}) \overset{\text{def}}{=} \text{false}
\]

\[
\text{slp}(P, X \leftarrow e) \overset{\text{def}}{=} \exists v : P[v/X] \land X = e[v/X]
\]

\[
\text{slp}(P, s; t) \overset{\text{def}}{=} \text{slp}(\text{slp}(P, s), t)
\]

\[
\text{slp}(P, \text{if } e \text{ then } s \text{ else } t) \overset{\text{def}}{=} \text{slp}(P \land e, s) \lor \text{slp}(P \lor \neg e, t)
\]

\[
\text{slp}(P, \text{while } e \text{ do } s) \overset{\text{def}}{=} (P \Rightarrow l) \land (\text{slp}(l \land e, s) \Rightarrow l) \land (\neg e \land l)
\]

(the rule for \( X \leftarrow e \) makes \( \text{slp} \) much less attractive than \( \text{wlp} \))
Verification conditions
How can we automate program verification using logic?

- Hoare logic: deductive system
  can only automate the checking of proofs

- predicate transformers: \( wlp, slp \) calculus
  construct (big) formulas mechanically
  invention is still needed for loops

- verification condition generation
  take as input a program with annotations
  (at least contracts and loop invariants)
  generate mechanically logic formulas ensuring the correctness
  (reduction to a mathematical problem, no longer any reference to a program)
  use an automatic SAT/SMT solver to prove (discharge) the formulas
  or an interactive theorem prover

(the idea of logic-based automated verification appears as early as [King69])
**Language**

\[
stat ::= \begin{array}{c}
X \leftarrow expr \\
\mid \text{skip} \\
\mid stat; \; stat \\
\mid \text{if} \; expr \; \text{then} \; stat \; \text{else} \; stat \\
\mid \text{while} \; \{Prop\} \; expr \; \text{do} \; stat \\
\mid \text{assert} \; expr
\end{array}
\]

\[
prog ::= \{Prop\} \; stat \; \{Prop\}
\]

- loops are annotated with loop invariants
- optional assertions at any point
- programs are annotated with a contract
  (precondition and postcondition)
Verification condition generation algorithm

by induction on the syntax of statements

\[
\text{vcg}_p : \text{prog} \rightarrow \mathcal{P}(\text{Prop})
\]

\[
\text{vcg}_p(\{P\} \ c \ \{Q\}) \defeq \\
\text{let } (R, C) = \text{vcg}_s(c, Q) \text{ in } C \cup \{P \Rightarrow R\}
\]

\[
\text{vcg}_s : (\text{stat} \times \text{Prop}) \rightarrow (\text{Prop} \times \mathcal{P}(\text{Prop}))
\]

\[
\text{vcg}_s(\text{skip}, Q) \defeq (Q, \emptyset)
\]

\[
\text{vcg}_s(X \leftarrow e, Q) \defeq (Q[e/X], \emptyset)
\]

\[
\text{vcg}_s(s; t, Q) \defeq \\
\text{let } (R, C) = \text{vcg}_s(t, Q) \text{ in let } (P, D) = \text{vcg}_s(s, R) \text{ in } (P, C \cup D)
\]

\[
\text{vcg}_s(\text{if } e \text{ then } s \text{ else } t, Q) \defeq \\
\text{let } (S, C) = \text{vcg}_s(s, Q) \text{ in let } (T, D) = \text{vcg}_s(t, Q) \text{ in } ((e \Rightarrow S) \land (\neg e \Rightarrow T), C \cup D)
\]

\[
\text{vcg}_s(\text{while } \{I\} e \text{ do } s, Q) \defeq \\
\text{let } (R, C) = \text{vcg}_s(s, I) \text{ in } (I, C \cup \{(I \land e) \Rightarrow R, (I \land \neg e) \Rightarrow Q\})
\]

\[
\text{vcg}_s(\text{assert } e, Q) \defeq (e \Rightarrow Q, \emptyset)
\]

- we use \textit{wlp} to infer assertions automatically when possible
- \text{vcg}_s(c, P) = (P', C) propagates postconditions backwards (P into P') and accumulates into C verification conditions (from loops)
- we could do the same using \textit{slp} instead of \textit{wlp} (symbolic execution)
Consider the program:
\[
\begin{align*}
\{N \geq 0\} & \quad X \leftarrow 1; I \leftarrow 0; \\
\text{while} \{X = 2^I \land 0 \leq I \leq N\} & \quad I < N \text{ do} \\
& \quad (X \leftarrow 2X; I \leftarrow I + 1) \\
\{X = 2^N\}
\end{align*}
\]

we get three verification conditions:

\[
\begin{align*}
C_1 \overset{\text{def}}{=} & \quad (X = 2^I \land 0 \leq I \leq N) \land I \geq N \Rightarrow X = 2^N \\
C_2 \overset{\text{def}}{=} & \quad (X = 2^I \land 0 \leq I \leq N) \land I < N \Rightarrow 2X = 2^{I+1} \land 0 \leq I + 1 \leq N \\
& \quad \text{(from } (X = 2^I \land 0 \leq I \leq N)[I+1/I, 2X/X]) \\
C_3 \overset{\text{def}}{=} & \quad N \geq 0 \Rightarrow 1 = 2^0 \land 0 \leq 0 \leq N \\
& \quad \text{(from } (X = 2^I \land 0 \leq I \leq N)[0/I, 1/X])
\end{align*}
\]

which can be checked independently
Extensions
Auxiliary variables:

mathematical variables that do not appear in the program
they are constant during program execution

Applications:

- simplify proofs
- express more properties (contracts, input-output relations)
- achieve completeness on extended languages

Example: \( \{ X = x \land Y = y \} \text{ if } X < Y \text{ then } Y \leftarrow X \text{ else skip } \{ Y = \text{min}(x, y) \} \)

- \( x \) and \( y \) retain the values of \( X \) and \( Y \) from the program entry
- \( Y = \text{min}(X, Y) \) is much less useful as a specification of a min function

"\( \{ \text{true} \} \text{ if } X < Y \text{ then } Y \leftarrow X \text{ else skip } \{ Y = \text{min}(X, Y) \} \)" holds, but
"\( \{ \text{true} \} X \leftarrow Y + 1 \{ Y = \text{min}(X, Y) \} \)" also holds
We model non-determinism with the statement $X \leftarrow ?$ meaning: $X$ is assigned a random value

$(X \leftarrow [a, b]$ can be modeled as: $X \leftarrow ?; \text{if } X < a \lor X > b \text{ then fail;}$)

**Hoare axiom:**

\[
\{\forall X : P\} \quad X \leftarrow ? \quad \{P\}
\]

If $P$ is true after assigning $X$ to random then $P$ must hold whatever the value of $X$ before

Often, $X$ does not appear in $P$ and we get simply:

\[
\{P\} \quad X \leftarrow ? \quad \{P\}
\]

**Example:**

\[
\{X = x\} \quad Y \leftarrow X \quad \{Y = x\}
\]
\[
\{Y = x\} \quad X \leftarrow ? \quad \{Y = x\} \quad \{Y = x\} \quad X \leftarrow Y \quad \{X = x\}
\]
\[
\{X = x\} \quad Y \leftarrow X ; \quad X \leftarrow ? ; \quad X \leftarrow Y \quad \{X = x\}
\]
Non-determinism

Predicate transformers:

- \( \text{wlp}(X \leftarrow ?, P) \mathrel{\overset{\text{def}}{=}} \forall X: P \)
  
  \((P \text{ must hold whatever the value of } X \text{ before the assignment})\)

- \( \text{slp}(P, X \leftarrow ?) \mathrel{\overset{\text{def}}{=}} \exists X: P \)
  
  \((\text{if } P \text{ held for one value of } X, \ P \text{ holds for all values of } X \text{ after the assignment})\)

Link with operational semantics (as transition systems)

predicates \( P \) as sets of states \( P \subseteq \Sigma \)
commands \( c \) as transition relations \( c \subseteq \Sigma \times \Sigma \)

we define: \( \text{post}[\tau](P) \mathrel{\overset{\text{def}}{=}} \{ \sigma' \mid \exists \sigma \in P: (\sigma, \sigma') \in \tau \} \)
\( \text{\texttilde{pre}}[\tau](P) \mathrel{\overset{\text{def}}{=}} \{ \sigma \mid \forall \sigma' \in \Sigma: (\sigma, \sigma') \in \tau \implies \sigma' \in P \} \)

then: \( \text{slp}(P, c) = \text{post}[c](P) \)
\( \text{wlp}(c, P) = \text{\texttilde{pre}}[c](P) \)
Total correctness

**Hoare triple:** \([P] \text{ prog } [Q]\)

- if \(P\) holds before \(\text{ prog}\) is executed
- then \(\text{ prog}\) always terminates
- and \(Q\) holds after the execution of \(\text{ prog}\)

**Rules:** we only need to change the rule for \(\text{ while}\)

\[
\forall t \in W: [P \land e \land u = t] \text{ s } [P \land u \prec t]
\]

\([P] \text{ while } e \text{ do } s [P \land \neg e]\)

- \((W, \prec)\) well-founded \(\iff\) every \(V \subseteq W, V \neq \emptyset\) has a minimal element for \(\prec\)
  - ensures that we cannot decrease infinitely by \(\prec\) in \(W\)
  - generally, we simply use \((\mathbb{N}, <)\)
    - (also useful: lexicographic orders, ordinals)

- in addition to the loop invariant \(P\)
  - we invent an expression \(u\) that strictly decreases by \(s\)
    - \(u\) is called a “ranking function”
    - often \(\neg e \implies u = 0\): \(u\) counts the number of steps until termination
To simplify, we can decompose a proof of total correctness into:

- a proof of partial correctness $\{P\} \ c \ {Q}$
  ignoring termination

- a proof of termination $[P] \ c \ [true]$
  ignoring the specification
  (we must still include the precondition $P$
  as the program may not terminate for all inputs)

Indeed, we have:

$$
\frac{\{P\} \ c \ {Q} \ \ [P] \ c \ [true]}{[P] \ c \ [Q]}
$$
Total correctness example

We use a simpler rule for integer ranking functions $((\mathbb{W}, \prec) \equiv (\mathbb{N}, \leq))$ using an integer expression $r$ over program variables:

$$\forall n: [P \land e \land (r = n)] \rightarrow [P \land (r < n)] \land (P \land e) \Rightarrow (r \geq 0)$$

$$[P] \text{ while } e \text{ do } s [P \land \neg e]$$

Example: $p \equiv \text{ while } I < N \text{ do } I \leftarrow I + 1; \ X \leftarrow 2X \text{ done}$

we use $r \equiv N - I$ and $P \equiv \text{ true}$

$$\forall n: [I < N \land N - I = n] \rightarrow [I \leftarrow I + 1; \ X \leftarrow 2X \land [N - I = n - 1]]$$

$$I < N \Rightarrow N - I \geq 0$$

$$[\text{true}] p [I \geq N]$$
Weakest precondition

**Weakest precondition**  \( \text{wp}(\text{prog}, \text{Prop}) : \text{Prop} \)

- similar to \( \text{wp} \), but also additionally imposes termination
- \([P] \ c \ [Q] \iff (P \Rightarrow \text{wp}(c, Q))\)

As before, only the definition for **while** needs to be modified:

\[
\text{wp}(\text{while } e \text{ do } s, P) \overset{\text{def}}{=} I \land \\
(I \Rightarrow v \geq 0) \land \\
\forall n: ((e \land I \land v = n) \Rightarrow \text{wp}(s, I \land v < n)) \land \\
((\neg e \land I) \Rightarrow P)
\]

the invariant predicate \( I \) is combined with a variant expression \( v \)

- \( v \) is positive \quad (this is an invariant: \( I \Rightarrow v \geq 0 \))
- \( v \) decreases at each loop iteration

(and similarly for strongest postconditions)
Arrays

We enrich our language with:

- a set $A$ of array variables
- array access in expressions: $A(expr), A \in A$
- array assignment: $A(expr) \leftarrow expr, A \in A$

(arrays have unbounded size here, we do not care about overflow)

**Issue:**

a natural idea is to generalize the assignment axiom:

\[
\{P[f/A(e)]\} A(e) \leftarrow f \{P\}
\]

but this is not sound, due to aliasing

example:

we would derive the invalid triple: $\{A(J) = 1 \land I = J\} A(I) \leftarrow 0 \{A(J) = 1 \land I = J\}$

as $(A(J) = 1)[0/A(I)] = (A(J) = 1)$
Solution: use a specific theory of arrays (McCarthy 1962)

- enrich the assertion language with expressions $A\{e \mapsto f\}$
  (meaning: the array equal to $A$ except that index $e$ maps to value $f$)

- add the axiom
  $\{P[A\{e \mapsto f\}/A]\} \ A(e) \leftarrow f \ \{P\}$
  (intuitively, we use “functional arrays” in the logic world)

- add logical axioms to reason about our arrays in assertions

\[
A\{e \mapsto f\}(e) = f \quad (e \neq e') \Rightarrow (A\{e \mapsto f\}(e') = A(e'))
\]
Example: swap

given the program $p \overset{\text{def}}{=} T \leftarrow A(I);\ A(I) \leftarrow A(J);\ A(J) \leftarrow T$

we wish to prove: $\{ A(I) = x \land A(J) = y \} \; p \; \{ A(I) = y \land A(J) = x \}$

by propagating $A(I) = y$ backwards by the assignment rule, we get

$$A\{ J \mapsto T \}(I) = y$$
$$A\{ I \mapsto A(J) \}\{ J \mapsto T \}(I) = y$$
$$A\{ I \mapsto A(J) \}\{ J \mapsto A(I) \}(I) = y$$

we consider two cases:

if $I = J$, then $A\{ I \mapsto A(J) \}\{ J \mapsto A(I) \} = A$

so, $A\{ I \mapsto A(J) \}\{ J \mapsto A(I) \}(I) = A(I) = A(J)$

if $I \neq J$, then $A\{ I \mapsto A(J) \}\{ J \mapsto A(I) \}(I) = A\{ I \mapsto A(J) \}(I) = A(J)$

in both cases, we get $A(J) = y$ in the precondition

likewise, $A(I) = x$ in the precondition
What about real languages?

In a real language such as C, the rules are not so simple

Example: the assignment rule \( \{P[e/X]\} X \leftarrow e \{P\} \) requires that

- \( e \) has no effect
- there is no pointer aliasing
- \( e \) has no run-time error

moreover, the operators in the program and in the logic may not match:

- integers: logic models \( \mathbb{Z} \), computers use \( \mathbb{Z}/2^n\mathbb{Z} \) (wrap-around)
- continuous:
  - logic models \( \mathbb{Q} \) or \( \mathbb{R} \), programs use floating-point numbers (rounding error)
- a logic for pointers and dynamic allocation is also required (separation logic)

(see for instance the tool Why, to see how some problems can be circumvented)
Conclusion
logic allows us to reason about program correctness
verification can be reduced to proofs of simple logic statements

**Issue: automation**

- annotations are required (loop invariants, contracts)
- verification conditions must be proven

to scale up to realistic programs, we need to automate as much as possible

**Some solutions:** in the following courses

- automatic logic solvers to discharge proof obligations
  SAT / SMT solvers
- abstract interpretation to approximate the semantics
  - fully automatic
  - able to infer invariants


