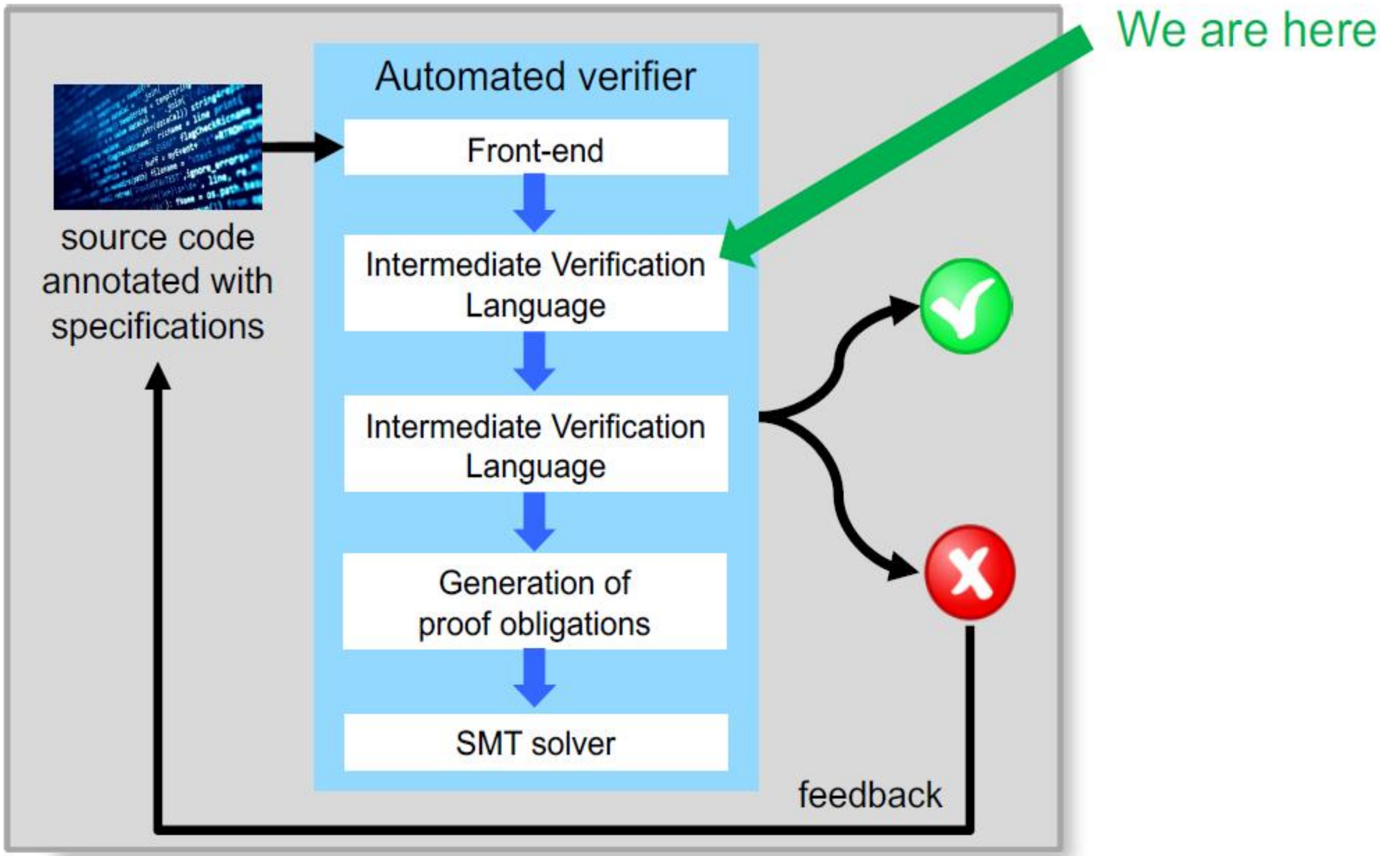
Program Analysis for Software Security

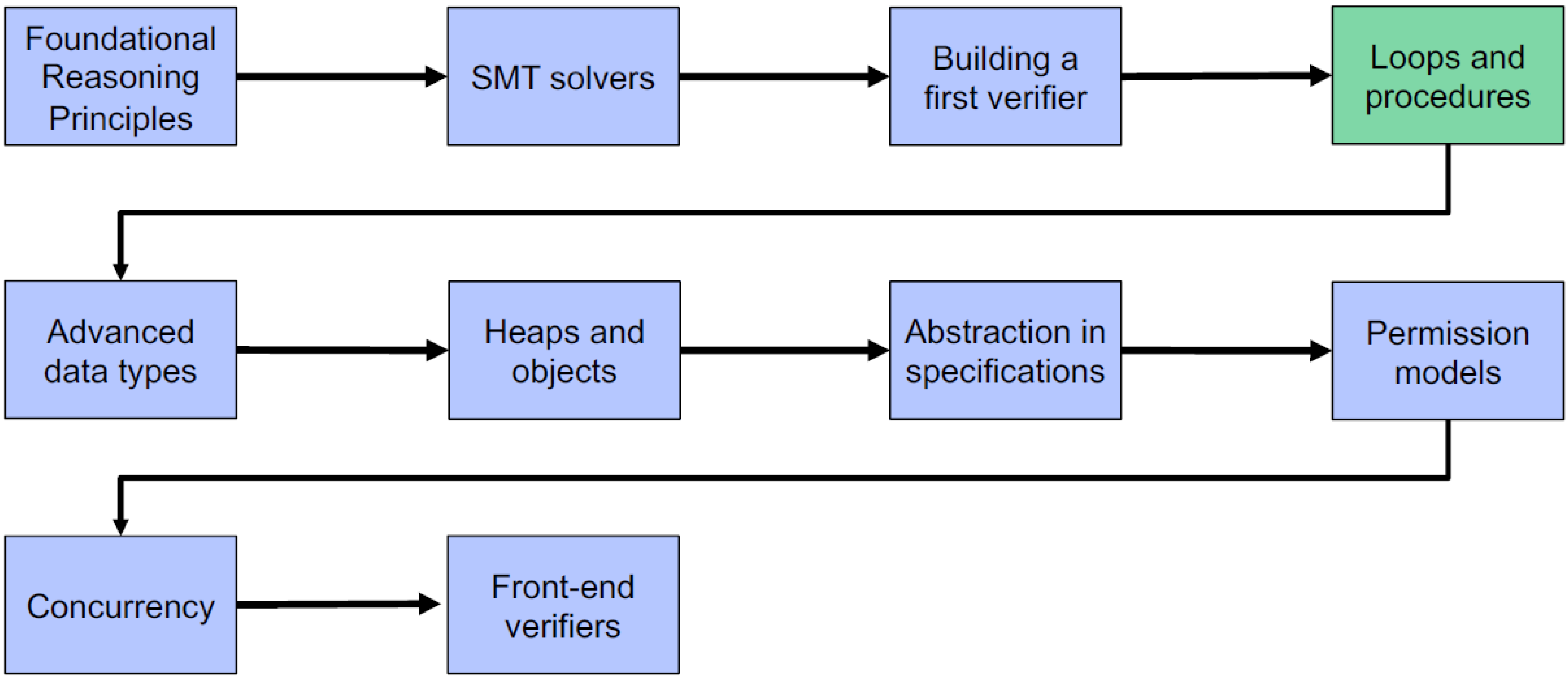
Lecture 5-6

Roadmap



We are here
Automated verifier
Front-end
Intermediate Verification Language
Intermediate Verification Language
Generation of proof obligations
SMT solver
source code annotated with specifications
feedback

Tentative course outline



LOOPS

Loops – operationally

Statements

```
S  ::=  ...  |  while (b) { S }
```

- If guard b holds, execute S and run loop again
- If b does not hold, terminate without an effect

Semantics

$$\frac{}{\texttt{while (b) \{ S \}, } \sigma \Rightarrow \texttt{if (b) \{ S; while (b) \{ S \} \} else \{ skip \}, } \sigma}$$

`assert true`

Loops – by example

Statements

```
S  ::=  ...  |  while (b) { S }
```

- If guard b holds, execute loop body S and repeat
- If guard b does not hold, terminate

```
assume n >= 0

var i: Int := 1
var r: Int := 0

while (i <= n) {
  r := r + i
  i := i + 1
}

assert ???
```

n = 5 (before guard)	
i	r
1	0
2	1
3	3
4	6
5	10

What should hold after the loop?

Loops – by example

Statements

```
S  ::=  ...  |  while (b) { S }
```

- If guard b holds, execute loop body S and repeat
- If guard b does not hold, terminate

```
assume n >= 0

var i: Int := 1
var r: Int := 0

while (i <= n) {
  r := r + i
  i := i + 1
}

assert n >= 0
assert r == n * (n+1) / 2
```

n = 5 (before guard)	
i	r
1	0
2	1
3	3
4	6
5	10

$$\sum_{i=1}^{n+1} i = \frac{n \cdot (n+1)}{2}$$

- Incompleteness: cannot reason *fully* automatically about all executions of unbounded loops
  ➔ need human interaction
- Model checking: unbounded loops yield infinite-state systems
- Static program analysis: infinite-height domains

Reminder

`{ P } S { Q }` is valid for total correctness iff

1. Safety:

   executing `S` on any state in `P` never fails an assertion

2. Partial correctness:

   every terminating execution of `S` on a state in `P` ends in a state in `Q`

3. Termination:

   every execution of `S` on a state from `P` stops after finitely many steps

iff verification condition `P ==> WP(S, Q)` is valid

Loops – by example

- Safety: loop execution does not fail

- Partial correctness: postcondition is satisfied if the loop terminates

- Termination of the loop

```
assume n >= 0

var i: Int := 1
var r: Int := 0

while (i <= n) {
  r := r + i
  i := i + 1
}

assert r == n * (n+1) / 2
assert n >= 0
```

Loops – by example with proof arguments

- Safety: loop execution does not fail
  - No assertion (failure) in the loop
- Partial correctness: postcondition is satisfied if the loop terminates
  - Before every loop iteration: `r == (i - 1) * i / 2`
  - Upon termination we also know `i == n + 1`
- Termination of the loop
  - `n – i + 1 >= 0`, always
  - `n – i + 1` decreases in every loop iteration

```
assume n >= 0

var i: Int := 1
var r: Int := 0

while (i <= n)
  invariant ...
{
  z := variant

  r := r + i
  i := i + 1
  assert variant < z

}

assert n >= 0
assert r == n * (n+1) / 2
```

➔ How do these annotations work?

Outline

- Weakest preconditions of loops
- Partial correctness reasoning
- Termination
- Encoding to PL0

Loops – operationally (reminder)

Statements

```
S  ::=  ...  |  while (b) { S }
```

- If guard b holds, execute S and run loop again
- If b does not hold, terminate without an effect

Semantics

```
―――――――――――――――――――――――――――――――――――――――――――――――――――――――――――――――
while (b) { S }, σ  ⇒  if (b) { S; while (b) { S } } else { skip }, σ
```

`assert true`

Loops – via unrolling

```
WP(while (b) { S }, Q)

=

WP(if (b) { S; while (b) { S } } else { skip }, Q)

=

(b ==> WP(S, WP(while (b) { S }, Q))) && (!b ==> Q)

::=  Φ(WP(while (b) { S }, Q))
```

➔ Solution is a fixed point of $X = \Phi(X)$

Running example

```
Φ(X) ::= (b ==> WP(S, X)) && (!b ==> Q)
```

```
Φ(X) ::=
  (i <= n ==> X[i / i+1][r / r+i]) &&
  (!(i <= n) ==> n >= 0 &&
                 r == n * (n+1) / 2)
```

```
while (i <= n) {
  { X[i / i+1][r / r+i] }
  r := r + i
  { X[i / i+1] }
  i := i + 1
  { X }
}

assert n >= 0
assert r == n * (n+1) / 2
```

Loops – as fixed points

`WP(while (b) { S }, Q)` must be a fixed point of

$\Phi(X)$ `::= b ==> WP(S, X) && !b ==> Q`

- `(Pred, ==>)` is a complete lattice
- `WP(S,_), b ==> _, &&` are monotone and continuous
- $\Phi(X)$ is monotone and continuous
- Tarski-Knaster Theorem: $\Phi(X)$ has at least one fixed point
- Which fixed point do we choose if there is more than one?

Exercise

1. Determine *all* fixed points of $\Phi(X)$ for the loop on the right and an arbitrary `Q`.

```
while (true) {
  skip // assert true
}
```

```
Φ(X) ::= b ==> WP(S, X)
         && !b ==> Q
```

2. Which fixed point corresponds to the weakest precondition of the loop, that is, what is

    `WP(while(true) { skip }, Q) ?`

Hint: recall that `WP(S, Q)` is the largest predicate P such that `{ P } S { Q }` is valid for total correctness.

3. Does your answer change if we reason about *partial* instead of *total* correctness? Why (not)?

Solution: multiple fixed points

$\Phi(X)$

=

`true ==>` ***WP***`(`**`skip`**`, `X`) && !true ==> Q`

=

`true ==>` X

=

X

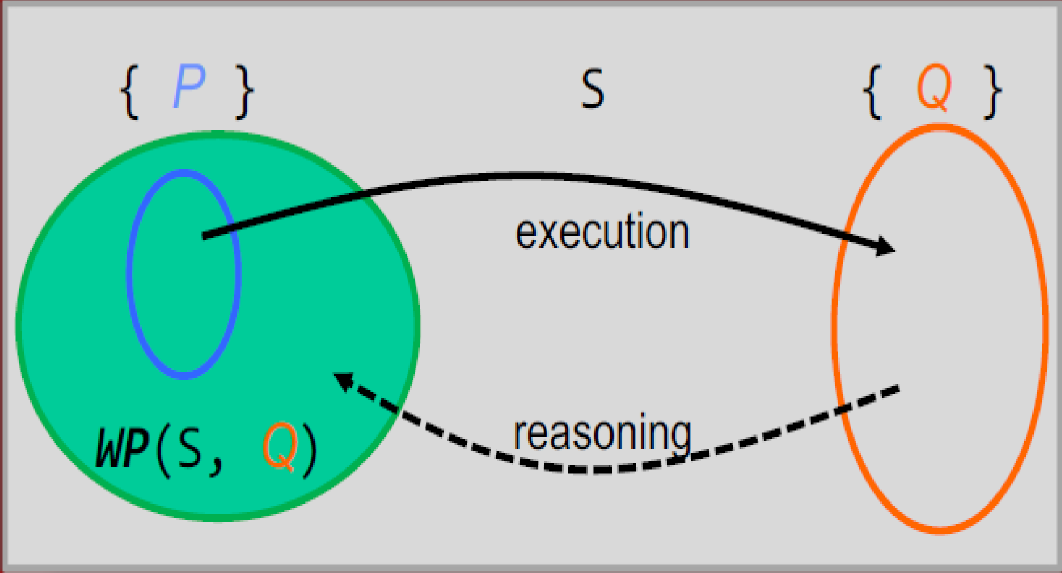
➔ every predicate is a fixed point

```
while (true) {
  skip // assert true
}
```

$\Phi(X)$ `::= b ==>` ***WP***`(S, `X`)`
`&& !b ==> Q`

Solution: multiple fixed points

Backward VC: P ==> WP(S, Q)
(are all initial states from which we must terminate in Q *included in* P?)




```
while (true) {
  skip
} // unreachable, regardless of the initial state
```

WP(while(true) { skip }, Q)
=
false
=
fix(Φ)

$\Phi(X) = X$

➔ pick *least* fixed point fix(Φ)

Loops – via weakest precondition

continuous predicate transformer

Weakest precondition of loops

```
WP(while (b) { S }, Q) ::= fix(Φ)

Φ (X) ::= b ==> WP(S, X) && !b ==> Q
```

Relative Completeness Theorem (Cook, 1974).

For PL0 programs and predicates, there exists a predicate that is logically equivalent to `fix(Φ)`.

Loops – via weakest precondition

continuous predicate transformer that depends on `b, S, Q`

Weakest precondition of loops

$$\mathit{WP}(\texttt{while (b) \{ S \}, Q}) \;::=\; \texttt{fix}(\Phi)$$

$$\Phi(X) \;::=\; \texttt{b} \implies \mathit{WP}(\texttt{S}, X) \;\&\&\; \texttt{!b} \implies \texttt{Q}$$

Kleene's fixed point theorem (applied to loops)

$$\texttt{fix}(\Phi) = \sup \{\Phi^{n}(\texttt{false}) \mid n \in \mathbb{N}\}$$

least fixed point may only be reached *in the limit*

$$\begin{array}{c}
\Phi^{\infty}(\texttt{false}) \\
\uparrow \\
\vdots \\
\uparrow \\
\Phi^{3}(\texttt{false}) \\
\uparrow \\
\Phi(\Phi(\texttt{false})) \\
\uparrow \\
\Phi(\texttt{false}) \\
\uparrow \\
\texttt{false}
\end{array}$$

Loops – a proof rule using Kleene's theorem

If we can find a parameterized predicate `I(k)` such that

1. `I(0) ==> Φ(false)`

2. `I(k+1) ==> Φ(I(k))`

3. `P ==>` $\left(\lim_{k\to\infty} \mathtt{I(k)}\right)$,

then `P ==>` $\underbrace{\texttt{wp(}\textbf{while}\texttt{ (b) \{ S \}}}_{=\ \texttt{fix}(\Phi)}\texttt{, Q).}$

$$
\begin{array}{c}
\lim_{k\to\infty} \mathtt{I(k)} \\
\uparrow \\
\Phi^{\infty}\mathtt{(false)} \\
\uparrow \\
\vdots \\
\mathtt{I(1)} \\
\uparrow \\
\Phi^{2}\mathtt{(false)} \\
\uparrow \\
\mathtt{I(0)} \\
\uparrow \\
\Phi\mathtt{(false)} \\
\uparrow \\
\mathtt{false}
\end{array}
$$

Example – via Kleene's theorem

If we can find a parameterized predicate `I(k)` such that

1. `I(0) ==> Φ(false)`
2. `I(k+1) ==> Φ(I(k))`
3. `P ==>` $\left(\lim_{k\to\infty} \texttt{I(k)}\right)$,

then `P ==> wp(while (b) { S }, Q)`.

```
I(k) ::= n >= 0 &&
         (i > n ==> r == n * (n+1) / 2) &&
         forall j:Int ::
             1 <= j && j <= k ==>
                 i == n - j + 1 ==>
                     r == (n-j) * (n-j+1) / 2
```

```
assume n >= 0

var i: Int := 1
var r: Int := 0

while (i <= n) {
  r := r + i
  i := i + 1
}

assert r == n * (n-1)/2
```

Example – via Kleene's theorem

If we can find a parameterized predicate `I(k)` such that

1. `I(0) ==> Φ(false)`
2. `I(k+1) ==> Φ(I(k))`
3. `P ==>` $\left(\lim_{k\to\infty} \texttt{I(k)}\right)$,

then `P ==> wp(while (b) { S }, Q)`.

$\lim_{k\to\infty}$ `I(k) = n >= 0 &&`
```
        (i > n ==> r == n * (n+1) / 2) &&
        forall j:Int ::
            1 <= j ~~&& j <= k~~ ==>
                i == n - j + 1 ==>
                    r == (n-j) * (n-j+1) / 2
```

```
assume n >= 0

var i: Int := 1
var r: Int := 0

while (i <= n) {
  r := r + i
  i := i + 1
}

assert r == n * (n-1)/2
```

→ Proves total correctness

→ Finding I(k) is challenging

→ Step 3 is hard to automate

Outline

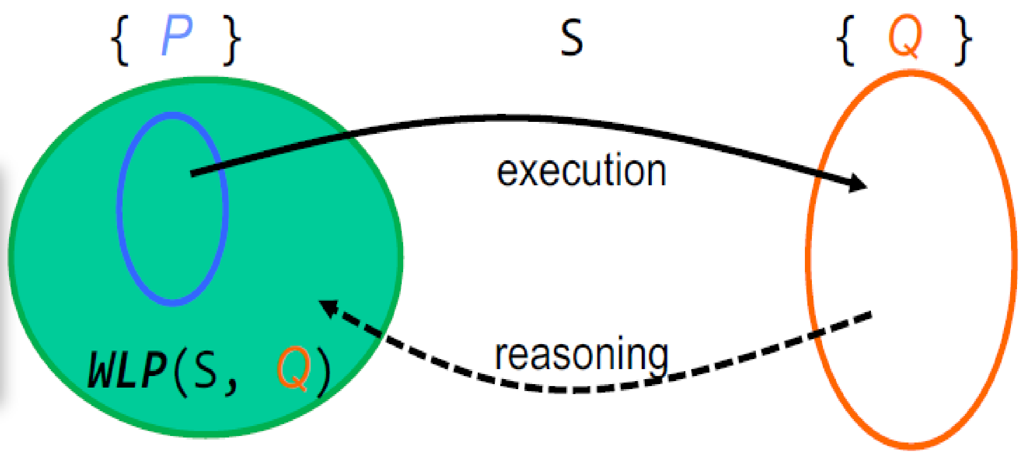
- Weakest preconditions of loops
- Partial correctness reasoning
- Termination
- Encoding to PL0

Loops – by example with proof arguments

- Safety: loop execution does not fail
  - No assertion (failure) in the loop
- Partial correctness: postcondition is satisfied if the loop terminates
  - Before every loop iteration: `r == (i - 1) * i / 2`
  - Upon termination we also know `i == n + 1`

```
assume n >= 0

var i: Int := 1
var r: Int := 0

while (i <= n)
  invariant ...
{


  r := r + i
  i := i + 1


}

assert n >= 0
assert r == n * (n+1) / 2
```

Loops – fixed points for partial correctness



Backward VC: P ==> *WLP*(S, Q)
(are all initial states from which every terminating execution of S ends in Q)

```
while (true) {
  skip
}
```

WLP(while(true) { skip }, Q)
=
true
=
FIX(Φ)

Φ(*X*) = *X*

➔ Pick *greatest* fixed point FIX(Φ)

Loops – weakest liberal preconditions

Backward VC: P ==> ***WLP***(S, Q)
(are all initial states from which every terminating execution of S ends in Q)

S	*WLP*(S, Q)
var x	**forall** x :: Q
x := a	Q[x / a]
assert R	R && Q
assume R	R ==> Q
S1; S2	***WLP***(S1, ***WLP***(S2, Q))
S1 [] S2	***WLP***(S1, Q) && ***WLP***(S2, Q)

Weakest *liberal* precondition of loops

WLP(**while** (b) { S }, Q) ::= FIX(Φ)

Φ (X) ::= b ==> ***WLP***(S, X) && !b ==> Q

Loops – inductive invariants

Weakest *liberal* precondition of loops

greatest fixed point

```
WLP(while (b) { S }, Q) ::= FIX(Φ)
Φ (X) ::= b ==> WLP(S, X) && !b ==> Q
```

Tarski-Knaster fixed point theorem

pre-fixed point

```
FIX(Φ) = sup { I | I ==> Φ(I) }
```

Inductive invariant rule

loop invariant

```
          I ==> Φ(I)
------------------------------
I ==> WLP(while (b) { S }, Q)
```

Loop invariants

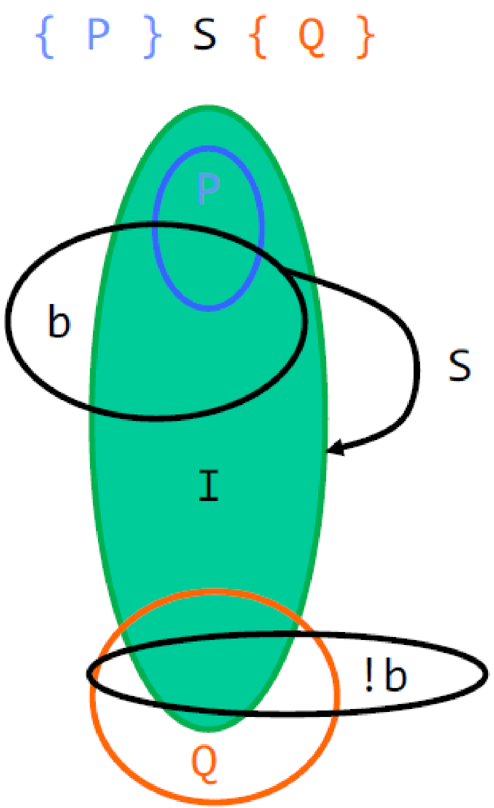
- Predicate that holds before every iteration

invariant I is preserved by one iteration

$$\frac{\{\ I\ \&\&\ b\ \}\ S\ \{\ I\ \}}{\{\ I\ \}\ \texttt{while (b) \{ S \}}\ \{\ I\ \&\&\ !b\ \}}$$

How can we derive this rule?

- Can be viewed as an induction proof
  - **Base:** invariant holds before the loop
  - **Hypothesis:** invariant holds before a fixed number of loop iterations
  - **Step:** invariant is preserved after performing one more iteration



Soundness

```
Assume { I && b } S { I }.

Equivalently:
I && b ==> WLP(S, I)
implies
I ==> b ==> WLP(S, I)
implies
I ==> (b ==> WLP(S, I)) && (!b ==> I)
implies (for Q ::= I && !b)
I ==> Φ(I)

By the inductive invariant rule:
I ==> WLP(while (b) { S }, I && !b)
equivalent to
{ I } while (b) { S } { I && !b }.
```

```
        { I && b } S { I }
-------------------------------------
{ I } while (b) { S } { I && !b }
```

Inductive invariant rule

```
          I ==> Φ(I)
-------------------------------------
I ==> WLP(while (b) { S }, Q)
```

```
Φ(X) ::= b ==> WLP(S, X)
         && !b ==> Q
```

Loop invariants

- Predicate that holds before every iteration

loop invariant I is preserved by one iteration

$$\frac{\{\ I\ \&\&\ b\ \}\ S\ \{\ I\ \}}{\{\ I\ \}\ \texttt{while}\ (b)\ \{\ S\ \}\ \{\ I\ \&\&\ !b\ \}}$$

- Can be viewed as an induction proof
  - **Base:** invariant holds before the loop
  - **Hypothesis:** invariant holds before a fixed number of loop iterations
  - **Step:** invariant is preserved after performing one more iteration

```
i := 1
r := 0

{ 0 <= r && 1 <= i }
while (i <= n) {
{ 0 <= r && 1 <= i && i <= n }
==>
{ 0 <= r + i && 1 <= i + 1 }
  r := r + i
{ 0 <= r && 1 <= i + 1 }
  i := i + 1
{ 0 <= r && 1 <= i }
}
{ 0 <= r && 1 <= i && !(i <= n) }
==>
{ 0 <= r }
```

Inductive loop invariants

```
{ I && b } S { I }
─────────────────────────────────
{ I } while (b) { S } { I && !b }
```

- Some predicates hold before every iteration but are not loop invariants
- We must be able to prove that the invariant is preserved
- Often requires strengthening the proposed invariant

```
i := 1
r := 0


while (i <= n) {
{ 0 <= r && i <= n }
==>  // proof fails
{ 0 <= r + i }
  r := r + i
{ 0 <= r }
  i := i + 1
{ 0 <= r }
}
{ 0 <= r && !(i <= n) }
==>
{ 0 <= r }
```

PL1: PL0 + loops with invariants

PL1 Statements

```
S ::= PL0... | while (b) invariant I { S }
```

Approximation of WLP with invariants

```
WLP(while (b) invariant I { S }, Q) ::= I
```

if predicate `I` is a loop invariant

```
i := 1; r := 0

while (i <= n)
  invariant 0 <= r && 1 <= i
{
  r := r + i
  i := i + 1
}
```

- We require loop invariants to be provided by the programmer
- Writing loop invariants is one of the main challenges for program verification
- Preservation of invariants needs to be checked as a side condition
  - invariant wrong → failure

Loops – in Viper

- Viper supports multiple invariants
  - all invariants are conjoined

```
while (0 < x)
  invariant 0 < x
  invariant x < 10
{ … }
```

- Error messages indicate why an invariant does not hold

```
var x: Int

while (0 < x)
  invariant 0 < x
{ … }
```

"Loop invariant might not hold on entry"

```
var x: Int
x := 5

while (0 < x)
  invariant 0 < x
{
  x := x - 1
}
```

"Loop invariant might not be preserved"

```
method main() {
  var n: Int
  var i: Int
  var r: Int

  assume n >= 0

  i := 1
  r := 0

  while (i <= n)
    invariant ??
  {
    r := r + i
    i := i + 1
  }

  assert r == n * (n+1) / 2
}
```

```
method main() {
  var n: Int
  var i: Int
  var r: Int

  assume n >= 0

  i := 1
  r := 0

  while (i <= n)
    invariant i <= n + 1
    invariant r == (i - 1) * i / 2
{
    r := r + i
    i := i + 1
  }

  assert r == n * (n+1) / 2
}
```

Outline

- Weakest preconditions of loops
- Partial correctness reasoning
- Termination
- Encoding to PL0

Proving termination

A loop **variant** is an an expression V that decreases in every loop iteration (for some well-founded ordering <).

< has no infinite descending chains

Well-founded	Not-well-founded
< over Nat	< over Int
⊂ over finite sets	< over positive reals

A loop terminates iff there exists a loop variant.

\# loop iterations

$V_1 > V_2 > V_3 > V_4 > \ \ldots > \ V_k$

Loop must stop after some finite number k of iterations because < has no infinite descending chains

old value of variant (>= 0)

V decreases

```
{ I && b && V == z } S { I && V < z }
--------------------------------------
   { I } while (b) { S } { I && !b }
```

Example – loops with variants

old value of variant (>= 0)

V decreases

```
{ I && b && V == z } S { I && V < z }
-------------------------------------
     { I } while (b) { S } { I && !b }
```

- Termination is experimental in Viper
- We can model variants with ghost code
  - code that does not affect execution
  - can be safely removed again
  - example: variables that keep track of old values

```
assume n >= 0

var i: Int := 1
var r: Int := 0

while (i <= n)
{
  var z: Int := n - i + 1
  assert z >= 0
  r := r + i
  i := i + 1
  assert n - i + 1 >= 0
  assert n - i + 1 < z
}

assert n >= 0
assert r == n * (n+1) / 2
```

V = n - i + 1

Example – loops with variants

old value of variant (>= 0)

V decreases

```
{ I && b && V == z } S { I && V < z }
─────────────────────────────────────
{ I } while (b) { S } { I && !b }
```

Ensures we use a well-founded ordering
(< over Nat)

Check that the variant V descreases after execution of the loop body, that is, V < z

```
assume n >= 0

var i: Int := 1
var r: Int := 0

while (i <= n)
{
  var z: Int := n - i + 1
  assert z >= 0
  r := r + i
  i := i + 1
  assert n - i + 1 >= 0
  assert n - i + 1 < z
}

assert n >= 0
assert r == n * (n+1) / 2
```

Chosen variant:
V = n - i + 1

Store value of V before loop body in ghost variable

Outline

- Weakest preconditions of loops
- Partial correctness reasoning
- Termination
- Encoding to PL0

Encoding of loops: naive attempt

$$\frac{\{\ I\ \&\&\ b\ \}\ S\ \{\ I\ \}}{\{\ I\ \}\ \text{while (b) \{ S \}}\ \{\ I\ \&\&\ !b\ \}}$$

- Check that loop invariant is preserved via a separate proof obligation

```
assume I
assume b

// encoding of S

assert I
```

- Verify the surrounding code by replacing the loop with statements that check and use the loop invariant

```
assert I

// havoc (reset) the state
var x; var y; // ...

assume I
assume !b
```

Loop framing

```
    { I && b } S { I }
----------------------------------
{ I } while (b) { S } { I && !b }
```

```
assert I

// havoc (reset) the state
var x; var y; // ...

assume I
assume !b
```

```
x := 0

while (false)
  invariant true
{ skip }

assert x == 0
```

- We often need to prove that a property is not affected by a loop
- Proving the preservation of a property across operations is called framing
- Our rule and our preliminary encoding require all framed properties to be conjoined to the loop invariant

Improved encoding for surrounding code

- It is sufficient to havoc those variables that get assigned to in the loop body
  - all other variables will not change
  - we do not forget their values

Frame rule

$$\frac{\{\ P\ \}\ S\ \{\ Q\ \} \qquad S \text{ modifies no var. in } R}{\{\ P\ \&\&\ R\ \}\ S\ \{\ Q\ \&\&\ R\ \}}$$

- We call the assigned variables loop targets

```
assert I
// havoc all loop targets
assume I
assume !b
```

```
x := 0
while (false)
  invariant true
{ skip }
assert x == 0
```

Improved encoding of invariant preservation

- If we check the invariant in a separate proof, we also check it for states we can never reach given the remaining code

```
assume I
assume b

// encoding of S

assert I
```

```
x := 0

while (true)
  invariant true
{ assert x == 0 }
```

invariant is checked for x == -1

- Solution check loop preservation after prior code

```
// prior code
// reset all loop targets
assume I
assume b

// encoding of S

assert I
```

```
x := 0

while (true)
  invariant true
{ assert x == 0 }
```

Final loop encoding

```
// prior code
// havoc all loop targets
assume I
assume b

// encoding of S

assert I
```

```
// prior code
assert I

// havoc all loop targets
assume I

assume !b
// subsequent code
```

```
// prior code

assert I

// havoc all loop targets

assume I

{
  assume b

  // encoding of S

  assert I
  assume false
} [] {
  assume !b
}

// subsequent code
```

Final loop encoding

```
// prior code
havoc all loop targets
assume I
assume b

// encoding of S

assert I
```

```
// prior code
assert I

havoc all loop targets
assume I

assume !b
// subsequent code
```

```
// prior code
{ I && forall ... :: (I && b ==> WP(S,I)) && (I && !b ==> Q) }
assert I
{ forall ... :: (I && b ==> WP(S,I)) && (I && !b ==> Q) }
// havoc all loop targets
{ (I && b ==> WP(S,I)) && (I && !b ==> Q) }
assume I
{ (b ==> WP(S,I)) && (!b ==> Q) }
{ { b ==> WP(S,I) }
  assume b { WP(S,I) }
  // encoding of S
  { I }
  assert I { true }
  assume false
} [] { { !b ==> Q }
  assume !b
} { Q }
// subsequent code
```

Loops: wrap-up

- Loop semantics is characterized by fixed points
  - total correctness: least fixed point
  - partial correctness: greatest fixed point
- We use loop invariants for proving partial correctness
  - Strong enough to prove correctness of loop body
  - Strong enough to establish postcondition
  - Preserved by loop body
- We use variants for proving termination
  - descreases in every loop iteration
  - well-founded: cannot decrease infinitely often
- Finding invariants and variants is one of the main sources of manual overhead in deductive verification