CS 599: Formal Methods in Security and Privacy: An imperative programming language and Hoare Triples

Marco Gaboardi
gaboardi@bu.edu

Alley Stoughton
stough@bu.edu

From the previous class

Does the program comply with the specification?

```
Precondition: x ≥ 0 and y ≥ 0
Function Add(x: int, y: int) : int
{
  r = 0;
  n = y;
  while n != 0
  {
    r = r + 1;
    n = n - 1;
  }
  return r
}
Postcondition: r == x + y
```

Does the program comply with the specification?

```
Precondition: x ≥ 0 and y ≥ 0
Function Add(x: int, y: int) : int
{
  r = 0;
  n = y;
  while n != 0
  {
    r = r + 1;
    n = n - 1;
  }
  return r
}
Postcondition: r == x + y
```

Fail to meet the specification

How about this one?

```
Precondition: x ≥ 0 and y ≥ 0
Function Add(x: int, y: int) : int
{
  r = x;
  n = y;
  while n != 0
  {
    r = r + 1;
    n = n - 1;
  }
  return r
}
Postcondition: r == x + y
```

How about this one?

```
Precondition: x ≥ 0 and y ≥ 0
Function Add(x: int, y: int) : int
{
  r = x;
  n = y;
  while n != 0
  {
    r = r + 1;
    n = n - 1;
  }
  return r
}
Postcondition: r == x + y
```

It meets
the specification

How can we make this reasoning mathematically precise?

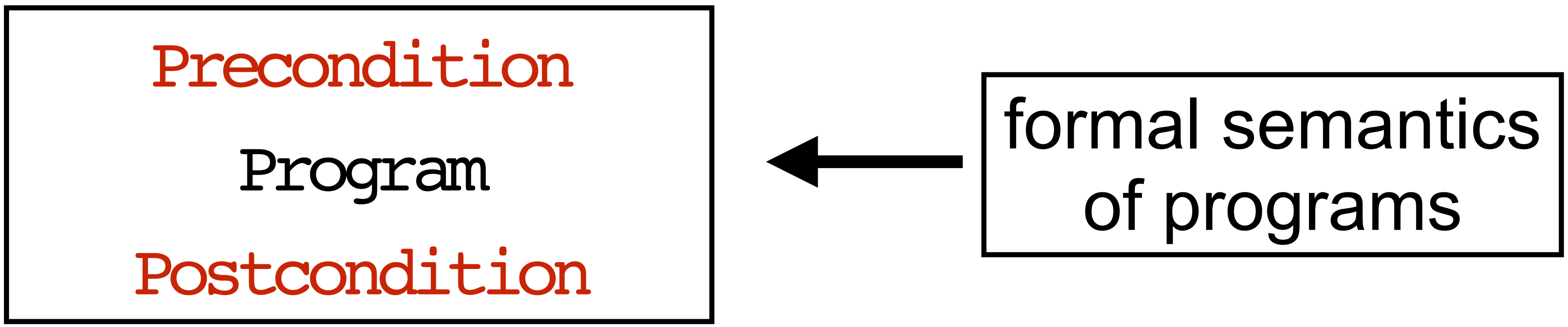
Formal Semantics

We need to assign a formal meaning to the different components:

Precondition
Program
Postcondition

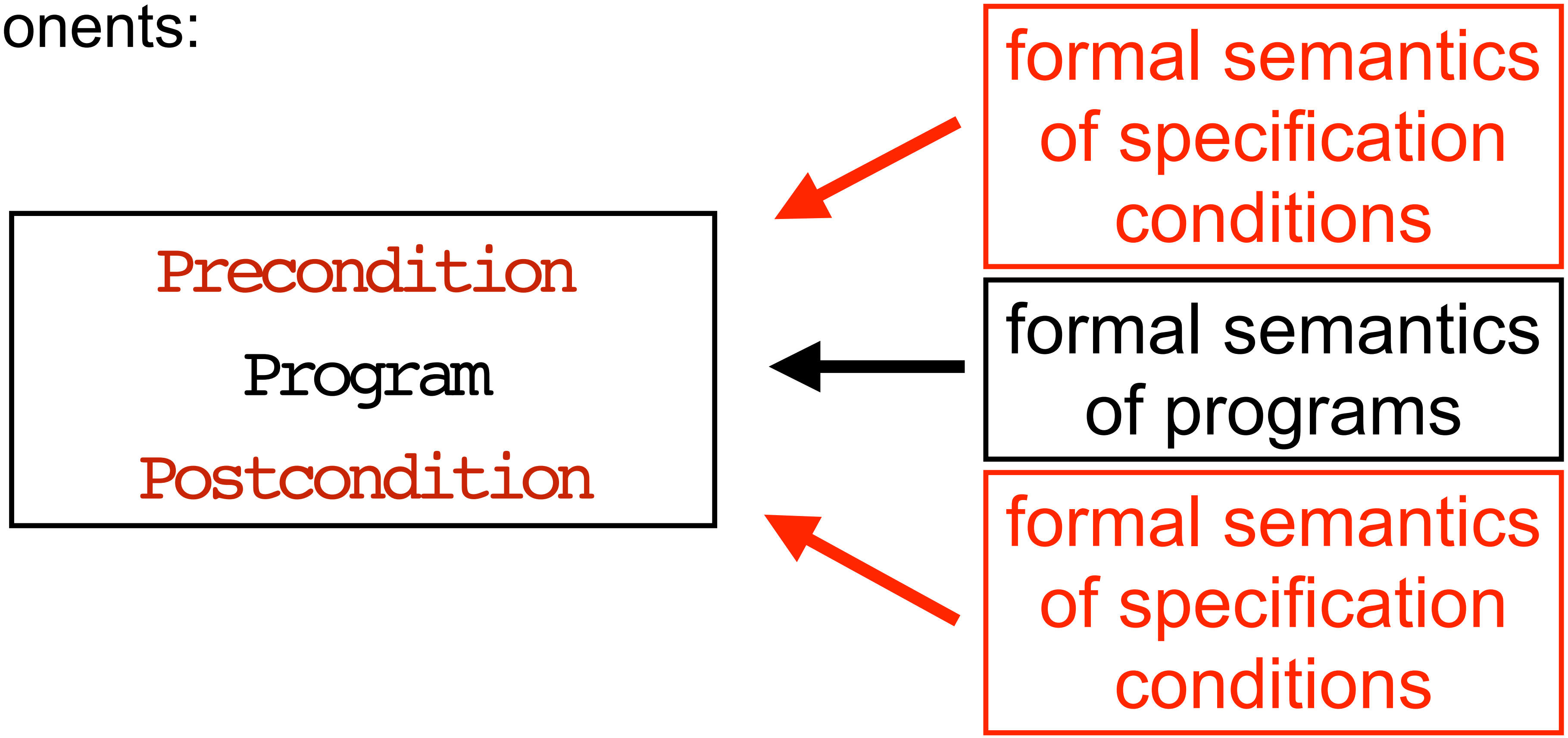
Formal Semantics

We need to assign a formal meaning to the different components:



Formal Semantics

We need to assign a formal meaning to the different components:



Formal Semantics

We need to assign a formal meaning to the different components:

Precondition

Program

Postcondition

formal semantics of specification conditions

formal semantics of programs

formal semantics of specification conditions

We also need to describe the rules which combine program and specifications.

Goal for today

- Formalize the semantics of a simple imperative programming language.

A first example

```
FastExponentiation(n, k : Nat) : Nat
  n′:= n; k′:= k; r := 1;
  if k′ > 0 then
    while k′ > 1 do
       if even(k′) then
         n′ := n′ ∗ n′;
         k′ := k′/2;
       else
         r := n′ ∗ r;
         n′ := n′ ∗ n′;
         k′ := (k′ − 1)/2;
    r := n′ ∗ r;
  (* result is r *)
```

Programming Language

```
c::= abort
   | skip
   | x:=e
   | c;c
   | if e then c else c
   | while e do c
```

x, y, z, … program variables

$e_1, e_2, \ldots$ expressions

$c_1, c_2, \ldots$ commands

Expressions

We want to be able to write complex programs with our language.

```
e::= x
   | f(e₁,…,eₙ)
```

Where `f` can be any arbitrary operator.

Some expression examples

`x+5`   `x mod k`   `x[i]`   `(x[i+1] mod 4)+5`

Types

In expressions we want to be able to use “arbitrary” data types.

```
t::= b
   | T(t₁,…,tₙ)
```

Types

In expressions we want to be able to use "arbitrary" data types.

```
t::= b
   | T(t1,…,tn)
```

We assume a collection of base types `b` including

`Bool`   `Int`   `Nat`   `String`

We also assume a set of type constructors `T` that we can use to build more complex types, such as:

`Bool list`   `Int*Bool`   `Int*String -> Bool`

Types

We also use types to guarantee that commands are well-formed.

For example, in the commands

```
while e do c          if e then c1 else c2
```

We require that `e` is of type `Bool`.

Types

We also use types to guarantee that commands are well-formed.

For example, in the commands

```
while e do c          if e then c1 else c2
```

We require that `e` is of type `Bool`.

We omit the details of the type system here but you can find them in the notes by Gilles Barthe

Values

Values are atomic expressions whose semantics is self-evident and which do not need a further analysis.

For example, we have the following values

```
true      5        [1,2,3,4]       "Hello"
```

The following are not values:

```
not true      x+5       [x,x+1]        x[1]
```

Values

Values are atomic expressions whose semantics is self-evident and which do not need a further analysis.

For example, we have the following values

`true`   `5`   `[1,2,3,4]`   `“Hello”`

The following are not values:

`not true`   `x+5`   `[x,x+1]`   `x[1]`

We could define a grammar for values, but we prefer to leave this at the intuitive level for now.

How can we give semantics to expressions and commands?

Memories

We can formalize a memory as a total map $\mathtt{m}$ from variables to values. $\mathtt{m} = [\mathtt{x}_1 \longmapsto \mathtt{v}_1, \ldots, \mathtt{x}_n \longmapsto \mathtt{v}_n]$

We consider only maps that respect types.

Memories

We can formalize a memory as a total map $\mathtt{m}$ from variables to values. $\mathtt{m}=[\mathtt{x}_1 \longmapsto \mathtt{v}_1, \ldots, \mathtt{x}_n \longmapsto \mathtt{v}_n]$

We consider only maps that respect types.

We want to read the value associated to a particular variable:

$$\mathtt{m}(\mathtt{x})$$

We want to update the value associated to a particular variable:

$$\mathtt{m}[\mathtt{x} \leftarrow \mathtt{v}]$$

This is defined as

$$\mathtt{m}[\mathtt{x} \leftarrow \mathtt{v}](\mathtt{y}) = \begin{cases} \mathtt{v} & \text{If } \mathtt{x}=\mathtt{y} \\ \mathtt{m}(\mathtt{y}) & \text{Otherwise} \end{cases}$$

Semantics of Expressions

What is the meaning of the following expressions?

`x+5`   `x mod k`   `x[i]`   `(x[i+1] mod 4)+5`

Semantics of Expressions

What is the meaning of the following expressions?

`x+5`   `x mod k`   `x[i]`   `(x[i+1] mod 4)+5`

We can give the semantics as a relation between expressions, memories and values.

```
Exp * Mem -> Val
```

We will denote this relation as:

$$\{e\}_m=v$$

Semantics of Expressions

What is the meaning of the following expressions?

`x+5`   `x mod k`   `x[i]`   `(x[i+1] mod 4)+5`

We can give the semantics as a relation between expressions, memories and values.

```
Exp * Mem -> Val
```

We will denote this relation as:

$\{e\}_m = v$

This is commonly typeset as:

$$[\![e]\!]_m = v$$

Semantics of Expressions

This is defined on the structure of expressions:

$$\{x\}_m = m(x)$$

$$\{f(e_1, \ldots, e_n)\}_m = \{f\}(\{e_1\}_m, \ldots, \{e_n\}_m)$$

where $\{f\}$ is the semantics associated with the basic operation we are considering.

Semantics of Expressions

Suppose we have a memory

$$\texttt{m}=[\texttt{i}\longmapsto 1, \texttt{x}\longmapsto [1,2,3], \texttt{y}\longmapsto 2]$$

That $\{\texttt{mod}\}$ is the modulo operation and $\{+\}$ is addition, we can derive the meaning of the following expression:

Semantics of Expressions

Suppose we have a memory

$$\texttt{m=[i}\longmapsto\texttt{1,x}\longmapsto\texttt{[1,2,3],y}\longmapsto\texttt{2]}$$

That $\{\texttt{mod}\}$ is the modulo operation and $\{\texttt{+}\}$ is addition, we can derive the meaning of the following expression:

$$\{\texttt{(x[i+1] mod y)+5}\}_{\texttt{m}}$$

Semantics of Expressions

Suppose we have a memory

$$\mathtt{m=[i\longmapsto 1,x\longmapsto [1,2,3],y\longmapsto 2]}$$

That $\{\mathtt{mod}\}$ is the modulo operation and $\{\mathtt{+}\}$ is addition, we can derive the meaning of the following expression:

$$\begin{aligned}&\{\mathtt{(x[i+1]\ mod\ y)+5}\}_{\mathtt{m}}\\ &\qquad = \{\mathtt{(x[i+1]\ mod\ y)}\}_{\mathtt{m}}\{\mathtt{+}\}\{\mathtt{5}\}_{\mathtt{m}}\end{aligned}$$

Semantics of Expressions

Suppose we have a memory

$$\mathtt{m=[i\mapsto 1,x\mapsto [1,2,3],y\mapsto 2]}$$

That $\{\mathtt{mod}\}$ is the modulo operation and $\{+\}$ is addition, we can derive the meaning of the following expression:

$$
\begin{aligned}
&\{\mathtt{(x[i+1]\ mod\ y)+5}\}_{\mathtt{m}} \\
&\qquad = \{\mathtt{(x[i+1]\ mod\ y)}\}_{\mathtt{m}}\{+\}\{\mathtt{5}\}_{\mathtt{m}} \\
&\qquad = (\{\mathtt{x[i+1]}\}_{\mathtt{m}}\ \{\mathtt{mod}\}\ \{\mathtt{y}\}_{\mathtt{m}})\{+\}\{\mathtt{5}\}_{\mathtt{m}}
\end{aligned}
$$

Semantics of Expressions

Suppose we have a memory

$$\mathtt{m=[i \longmapsto 1, x \longmapsto [1,2,3], y \longmapsto 2]}$$

That $\{\mathtt{mod}\}$ is the modulo operation and $\{\mathtt{+}\}$ is addition, we can derive the meaning of the following expression:

$$
\begin{aligned}
&\{\mathtt{(x[i+1]\ mod\ y)+5}\}_{\mathtt{m}} \\
&\quad = \{\mathtt{(x[i+1]\ mod\ y)}\}_{\mathtt{m}}\{\mathtt{+}\}\{\mathtt{5}\}_{\mathtt{m}} \\
&\quad = (\{\mathtt{x[i+1]}\}_{\mathtt{m}}\ \{\mathtt{mod}\}\ \{\mathtt{y}\}_{\mathtt{m}})\{\mathtt{+}\}\{\mathtt{5}\}_{\mathtt{m}} \\
&\quad = (\{\mathtt{x}\}_{\mathtt{m}}[\{\mathtt{i}\}_{\mathtt{m}}\{\mathtt{+}\}\{\mathtt{1}\}_{\mathtt{m}}]\ \{\mathtt{mod}\}\ \{\mathtt{y}\}_{\mathtt{m}})\{\mathtt{+}\}\{\mathtt{5}\}_{\mathtt{m}}
\end{aligned}
$$

Semantics of Expressions

Suppose we have a memory

$$m = [i \mapsto 1, x \mapsto [1,2,3], y \mapsto 2]$$

That $\{mod\}$ is the modulo operation and $\{+\}$ is addition, we can derive the meaning of the following expression:

$$
\begin{aligned}
&\{(x[i+1] \text{ mod } y)+5\}_m \\
&\quad = \{(x[i+1] \text{ mod } y)\}_m\{+\}\{5\}_m \\
&\quad = (\{x[i+1]\}_m \ \{mod\} \ \{y\}_m)\{+\}\{5\}_m \\
&\quad = (\{x\}_m[\{i\}_m\{+\}\{1\}_m] \ \{mod\} \ \{y\}_m)\{+\}\{5\}_m \\
&\quad = (\{x\}_m[1\{+\}1] \ \{mod\} \ 2)\{+\}5
\end{aligned}
$$

Semantics of Expressions

Suppose we have a memory

$$\mathtt{m} = [\mathtt{i} \longmapsto 1, \mathtt{x} \longmapsto [1, 2, 3], \mathtt{y} \longmapsto 2]$$

That $\{\mathtt{mod}\}$ is the modulo operation and $\{+\}$ is addition, we can derive the meaning of the following expression:

$$
\begin{aligned}
&\{(\mathtt{x[i+1]\ mod\ y)+5}\}_{\mathtt{m}} \\
&\quad = \{(\mathtt{x[i+1]\ mod\ y})\}_{\mathtt{m}}\{+\}\{5\}_{\mathtt{m}} \\
&\quad = (\{\mathtt{x[i+1]}\}_{\mathtt{m}}\ \{\mathtt{mod}\}\ \{\mathtt{y}\}_{\mathtt{m}})\{+\}\{5\}_{\mathtt{m}} \\
&\quad = (\{\mathtt{x}\}_{\mathtt{m}}[\{\mathtt{i}\}_{\mathtt{m}}\{+\}\{1\}_{\mathtt{m}}]\ \{\mathtt{mod}\}\ \{\mathtt{y}\}_{\mathtt{m}})\{+\}\{5\}_{\mathtt{m}} \\
&\quad = (\{\mathtt{x}\}_{\mathtt{m}}[1\{+\}1]\ \{\mathtt{mod}\}\ 2)\{+\}5 \\
&\quad = (\{\mathtt{x}\}_{\mathtt{m}}[2]\ \{\mathtt{mod}\}\ 2)\{+\}5
\end{aligned}
$$

Semantics of Expressions

Suppose we have a memory

$$m=[i\mapsto 1, x\mapsto [1,2,3], y\mapsto 2]$$

That $\{mod\}$ is the modulo operation and $\{+\}$ is addition, we can derive the meaning of the following expression:

$$
\begin{aligned}
&\{(x[i+1]\ mod\ y)+5\}_m \\
&\quad = \{(x[i+1]\ mod\ y)\}_m\{+\}\{5\}_m \\
&\quad = (\{x[i+1]\}_m\ \{mod\}\ \{y\}_m)\{+\}\{5\}_m \\
&\quad = (\{x\}_m[\{i\}_m\{+\}\{1\}_m]\ \{mod\}\ \{y\}_m)\{+\}\{5\}_m \\
&\quad = (\{x\}_m[1\{+\}1]\ \{mod\}\ 2)\{+\}5 \\
&\quad = (\{x\}_m[2]\ \{mod\}\ 2)\{+\}5 \\
&\quad = (2\ \{mod\}\ 2)\{+\}5 = 0\ \{+\}\ 5 = 5
\end{aligned}
$$

Operational vs Denotational Semantics

The style of semantics we are using is denotational, in the sense that we describe the meaning of an expression by means of the value it denotes.

A different approach, more operational in nature, would be to describe the meaning of an expression by means of the value that the expression evaluates to in an abstract machine.

Semantics of Commands

What is the meaning of the following command?

```
k:=2; z:=x mod k; if z=0 then r:=1 else r:=2
```

Semantics of Commands

What is the meaning of the following command?

```
k:=2; z:=x mod k; if z=0 then r:=1 else r:=2
```

We can give the semantics as a relation between command, memories and memories or failure.

```
Exp * Mem -> Mem
```

Semantics of Commands

What is the meaning of the following command?

```
k:=2; z:=x mod k; if z=0 then r:=1 else r:=2
```

We can give the semantics as a relation between command, memories and memories or failure.

```
Exp * Mem -> Mem
```

Would this work?

Semantics of Commands

What is the meaning of the following command?

```
k:=2; z:=x mod k; if z=0 then r:=1 else r:=2
```

Semantics of Commands

What is the meaning of the following command?

```
k:=2; z:=x mod k; if z=0 then r:=1 else r:=2
```

We can give the semantics as a relation between command, memories and memories or failure.

$$\text{Exp} * \text{Mem} \rightarrow (\text{Mem} \cup \{\bot\})$$

We will denote this relation as:

$\{c\}_m = m'$ Or $\{c\}_m = \bot$

Semantics of Commands

What is the meaning of the following command?

```
k:=2; z:=x mod k; if z=0 then r:=1 else r:=2
```

We can give the semantics as a relation between command, memories and memories or failure.

$$\texttt{Exp * Mem -> (Mem } \mathbf{U} \texttt{ \{}\bot\texttt{\})}$$

We will denote this relation as:

$\{c\}_m = m'$ Or $\{c\}_m = \bot$

This is commonly typeset as:

$$[\![c]\!]_m = m'$$

Semantics of Commands

This is defined on the structure of commands:

Semantics of Commands

This is defined on the structure of commands:

$\{\texttt{abort}\}_m = \bot$

Semantics of Commands

This is defined on the structure of commands:

$$\{\texttt{abort}\}_{\texttt{m}} = \bot$$

$$\{\texttt{skip}\}_{\texttt{m}} = \texttt{m}$$

Semantics of Commands

This is defined on the structure of commands:

$$\{\texttt{abort}\}_{\texttt{m}} = \bot$$

$$\{\texttt{skip}\}_{\texttt{m}} = \texttt{m}$$

$$\{\texttt{x:=e}\}_{\texttt{m}} = \texttt{m}[\texttt{x} \leftarrow \{\texttt{e}\}_{\texttt{m}}]$$

Semantics of Commands

This is defined on the structure of commands:

$$\{\texttt{abort}\}_{m} = \bot$$

$$\{\texttt{skip}\}_{m} = m$$

$$\{\texttt{x:=e}\}_{m} = m[x \leftarrow \{e\}_{m}]$$

$$\{\texttt{c;c'}\}_{m} = \{c'\}_{m'} \quad \text{If} \quad \{c\}_{m} = m'$$

Semantics of Commands

This is defined on the structure of commands:

$$\{\texttt{abort}\}_m = \bot$$

$$\{\texttt{skip}\}_m = m$$

$$\{\texttt{x:=e}\}_m = m[x \leftarrow \{e\}_m]$$

$$\{\texttt{c;c'}\}_m = \{\texttt{c'}\}_{m'} \quad \text{If} \quad \{\texttt{c}\}_m = m'$$

$$\{\texttt{c;c'}\}_m = \bot \quad \text{If} \quad \{\texttt{c}\}_m = \bot$$

Semantics of Commands

This is defined on the structure of commands:

$\{\texttt{abort}\}_m = \bot$

$\{\texttt{skip}\}_m = m$

$\{\texttt{x:=e}\}_m = m[x \leftarrow \{e\}_m]$

$\{\texttt{c;c'}\}_m = \{c'\}_{m'}$ If $\{c\}_m = m'$

$\{\texttt{c;c'}\}_m = \bot$ If $\{c\}_m = \bot$

$\{\texttt{if e then } c_t \texttt{ else } c_f\}_m = \{c_t\}_m$ If $\{e\}_m\texttt{=true}$

Semantics of Commands

This is defined on the structure of commands:

$\{\texttt{abort}\}_m = \bot$

$\{\texttt{skip}\}_m = m$

$\{\texttt{x:=e}\}_m = m[x \leftarrow \{e\}_m]$

$\{\texttt{c;c'}\}_m = \{c'\}_{m'}$ If $\{c\}_m = m'$

$\{\texttt{c;c'}\}_m = \bot$ If $\{c\}_m = \bot$

$\{\texttt{if e then } c_t \texttt{ else } c_f\}_m = \{c_t\}_m$ If $\{e\}_m = \texttt{true}$

$\{\texttt{if e then } c_t \texttt{ else } c_f\}_m = \{c_f\}_m$ If $\{e\}_m = \texttt{false}$

Semantics of While

What about while

Semantics of While

What about while

$$\{\texttt{while e do c}\}_{\texttt{m}} = \texttt{???}$$

Semantics of While

What about while

$$\{\texttt{while e do c}\}_m = \texttt{???}$$

We omit the semantics of while, you can find it in the notes by Gilles Barthe.
Alternatively, you can look at these notes:
https://groups.seas.harvard.edu/courses/cs152/2016sp/lectures/lec06-denotational.pdf
https://web.cecs.pdx.edu/~apt/cs578_2022/imp.pdf

Hoare Triples

Hoare triple

Precondition
Program
Postcondition

Precondition
(a logical formula)

$$c : P \Rightarrow Q$$

Program

Postcondition
(a logical formula)

Some examples

Precondition

$$x := z + 1 : \{z = n\} \Rightarrow \{x = n + 1\}$$

Postcondition

Is it a good specification?

Some examples

Precondition

$$x := z + 1 : \{z = n\} \Rightarrow \{x = n + 1\}$$

Postcondition

Is it a good specification?

✓

How do we determine the validity of an Hoare triple?

Specification can also be imprecise.

Some examples

Precondition

$$x := z + 1 : \{z > 0\} \Rightarrow \{x > 0\}$$

Postcondition

Is it a good specification?

Some examples

Precondition

$$x := z + 1 : \{z > 0\} \Rightarrow \{x > 0\}$$

Postcondition

Is it a good specification?

✓

Some examples

Precondition

$$x := z + 1 : \{z + 1 > 0\} \Rightarrow \{x > 0\}$$

Postcondition

Is it a good specification?

Some examples

Precondition

$$x := z + 1 : \{z + 1 > 0\} \Rightarrow \{x > 0\}$$

Postcondition

Is it a good specification?

✓

Some examples

Precondition

$$x := z + 1 : \{z < 0\} \Rightarrow \{x < 0\}$$

Postcondition

Is it a good specification?

Some examples

Precondition

$$x := z + 1 : \{z < 0\} \Rightarrow \{x < 0\}$$

Postcondition

Is it a good specification?

✗

Some examples

Precondition

$$x := z + 1 : \{z < 0\} \Rightarrow \{x < 0\}$$

Postcondition

Is it a good specification?

✗

$$m_{in} = [z = -1, x = 2] \qquad m_{out} = [z = -1, x = 0]$$

Some examples

```
i:=0;
r:=1;
while(i≤k)do
  r:=r * n;
  i:=i + 1
```

Precondition

$$: \{0 \leq k\} \Rightarrow \{r = n^k\}$$

Postcondition

Is it a good specification?

Some examples

```
i:=0;
r:=1;
while(i≤k)do
  r:=r * n;
  i:=i + 1
```

Precondition

$$: \{0 \leq k\} \Rightarrow \{r = n^k\}$$

Postcondition

Is it a good specification?

✗

Some examples

```
i:=0;
r:=1;
while(i≤k)do
  r:=r * n;
  i:=i + 1
```

Precondition

$$: \{0 \leq k\} \Rightarrow \{r = n^k\}$$

Postcondition

Is it a good specification? ✗

$$m_{in} = [k = 0, n = 2, i = 0, r = 0]$$

$$m_{out} = [k = 0, n = 2, i = 1, r = 2]$$

Some examples

```
i:=0;
r:=1;
while(i≤k)do
  r:=r * n;
  i:=i + 1
```

Precondition

$$: \{0 < k\} \Rightarrow \{r = n^k\}$$

Postcondition

Is it a good specification?

Some examples

```
i:=0;
r:=1;
while(i≤k)do
  r:=r * n;
  i:=i + 1
```

Precondition

$$: \{0 < k\} \Rightarrow \{r = n^k\}$$

Postcondition

Is it a good specification?

✗

Some examples

```
i:=0;
r:=1;
while(i≤k)do
  r:=r * n;
  i:=i + 1
```

Precondition

$$: \{0 < k\} \Rightarrow \{r = n^k\}$$

Postcondition

Is it a good specification?

✗

$$m_{in} = [k = 1, n = 2, i = 0, r = 0]$$

$$m_{out} = [k = 1, n = 2, i = 2, r = 4]$$

Some examples

```
i:=0;
r:=1;
while(i<k)do
  r:=r * n;
  i:=i + 1
```

Precondition

$$: \{0 \leq k\} \Rightarrow \{r = n^k\}$$

Postcondition

Is it a good specification?

Some examples

```
i:=0;
r:=1;
while(i<k)do
  r:=r * n;
  i:=i + 1
```

Precondition

$$: \{0 \le k\} \Rightarrow \{r = n^k\}$$

Postcondition

Is it a good specification?

✓

Some examples

```
i:=0;
r:=1;
while(i≤k)do
  r:=r * n;
  i:=i + 1
```

Precondition

$$: \{0 \le k\} \Rightarrow \{r = n^i\}$$

Postcondition

Is it a good specification?

Some examples

```
i:=0;
r:=1;
while(i≤k)do
  r:=r * n;
  i:=i + 1
```

Precondition

$$: \{0 \leq k\} \Rightarrow \{r = n^i\}$$

Postcondition

Is it a good specification? ✓

Some examples

```
i:=0;
r:=1;
while(i≤k)do
  r:=r * n;
  i:=i + 1
```

Precondition

$$: \{0 < k \wedge k < 0\} \Rightarrow \{r = n^k\}$$

Postcondition

Is it a good specification?

Some examples

```
i:=0;
r:=1;
while(i≤k)do
  r:=r * n;
  i:=i + 1
```

Precondition

$$: \{0 < k \land k < 0\} \Rightarrow \{r = n^k\}$$

Postcondition

Is it a good specification? ✓

Some examples

```
i:=0;
r:=1;
while(i≤k)do
  r:=r * n;
  i:=i + 1
```

Precondition

$$: \{0 < k \wedge k < 0\} \Rightarrow \{r = n^k\}$$

Postcondition

Is it a good specification?

✓

This is good because there is no memory that satisfies the precondition.