Design and Analysis of Algorithms CS218M

Correctness of Algorithms

Paritosh Pandya

Indian Institute of Technology, Bombay

Autumn, 2022

Programs and Assertions

Programs (e expressions, b boolean expr.)

x:=e S1;S2 if b then S1 else S2 fi while b do S od

A State assigns a value to each variable.

A program starts in an initial state. It ends in a final state or does not terminate.

Assertions

Assertions

Conditions on state. They specify a subset of states. E.g. x > y. Formally, assertions are formulae of first-order logic.

Assertions use logical connectives.

 $P \wedge Q$ P and Q

 $P \lor Q$ P or Q

 $\neg P$ not P

 $P \Rightarrow Q$ whenever P is true so is Q

Reasoning

 $AXIOMS \models P \Rightarrow Q$

Problem

```
\begin{array}{l} \text{r:=x; q:=0;}\\ \text{while r} > \text{ y do}\\ \text{r:=r-y; q:=q+1}\\ \text{od} \end{array}
```

Problem

$$\begin{array}{l} \text{r:=x; q:=0;}\\ \text{while r} > \text{ y do}\\ \text{r:=r-y; q:=q+1}\\ \text{od} \end{array}$$

Х	у	q	r
8	3	2	2
8	0		
-8	3	0	-8
6	3	1	3

Problem

$$\begin{cases} 0 < y & \land & 0 \leq x \} \\ r{:=}x; \ q{:=}0; \\ \text{while } r > y \ do \\ r{:=}r{-}y; \ q{:=}q{+}1 \\ \text{od}$$

Х	у	q	r
8	3	2	2
8	0		
-8	3	0	-8
6	3	1	3

Problem

Program specification

 $\{P\} S \{Q\}$

- Hoave Triple
- S Program (fragment)
- P Precondition

Assumed to be true when S starts.

• Q Postcondition

Required to be true when S terminates.

Advantages

- Clear and Unambiguous articulation of what program must do.
- Separation of concern: User versus developer. interface specification.
- Can be formally verified.

Annotated Program

Assertions

Pre-condition and post-condition.

Location Invariants

- Control location: a position before a program statement
- Location Invariant: Condition which is true every time control reaches the location.

Assertions

Pre-condition and post-condition.

Location Invariants

- Control location: a position before a program statement
- Location Invariant: Condition which is true every time control reaches the location.

Loop Invariant

Consider while b do S od.

• A condition which holds every time condition b is tested.

Hoare Logic

Given predicate Q Q[e/x] denotes Q with x substituted by e E.g. x < 0[x+1/x] gives x+1 < 0.

Assignment

$${Q[e/x]} \quad x := e \quad {Q}$$

RI

Example: $\{x + 1 < 0\} = x + 1 \{x < 0\}$

Sequential Composition

$$\frac{\{P\} \ S_1 \ \{Q_1\}, \quad Q_1 \Rightarrow Q_2, \quad \{Q_2\} \ S_2 \ \{R\}}{\{P\} \ S_1; S_2 \ \{R\}}$$

22

Hoare Logic (2)

Consequence

$$\frac{P \Rightarrow P_1, \quad \{P_1\} \ S \ \{Q_1\}, \quad Q_1 \Rightarrow Q}{\{P\} \ S \ \{Q\}}$$



Conditional Statement

$$\frac{\{P \wedge b\} \ S_1 \ \{Q\}, \quad \{P \wedge \neg b\} \ S_2 \ \{Q\}}{\{P\} \ \text{if } b \ \text{then} \ S_1 \ \text{else} \ S_2 \ \text{fi} \ \{Q\}}$$



```
Claim: \{0 \le r \land 0 < y \land x = y*q+r \land y \le r\}
r:=r-y; q:=q+1
\{0 \le r \land 0 < y \land x = y*q+r\}
```

Claim:
$$\{0 \le r \land 0 < y \land x = y * q + r \land y \le r\}$$

r:=r-y; q:=q+1
 $\{0 \le r \land 0 < y \land x = y * q + r\}$

Proof

(1)

(2)

(3)

(4)

(5)

(6)

Claim:
$$\{0 \le r \land 0 < y \land x = y * q + r \land y \le r\}$$

r:=r-y; q:=q+1
 $\{0 \le r \land 0 < y \land x = y * q + r\}$

Proof (1)(2)(3)(4)(5)q := q+1 $\{0 \le r \land 0 < y \land x = y * q + r\}$ (6)

Claim:
$$\{0 \le r \land 0 < y \land x = y * q + r \land y \le r\}$$

r:=r-y; q:=q+1
 $\{0 \le r \land 0 < y \land x = y * q + r\}$

Proof

(1)

(2)

(3)

$$\{0 \le r \land 0 < y \land x = y * (q+1) + r\}$$

(4)

$$q := q+1$$

(5)

$$\{0 \le r \land 0 < y \land x = y * q + r\}$$

(6)

Claim:
$$\{0 \le r \land 0 < y \land x = y * q + r \land y \le r\}$$

r:=r-y; q:=q+1
 $\{0 \le r \land 0 < y \land x = y * q + r\}$

Proof

(1)

(2)

$$r:=r-y;$$
 (3)

$$\{0 \le r \land 0 < y \land x = y * (q+1) + r\} \tag{4}$$

$$q := q+1 \tag{5}$$

$$\{0 \le r \land 0 < y \land x = y * q + r\} \tag{6}$$

Claim:
$$\{0 \le r \land 0 < y \land x = y * q + r \land y \le r\}$$

r:=r-y; q:=q+1
 $\{0 \le r \land 0 < y \land x = y * q + r\}$

Proof

(1)

$$\{0 \le r - y \land 0 < y \land x = y * (q+1) + (r-y)\}$$
 (2)

$$r:=r-y;$$
 (3)

$$\{0 \le r \land 0 < y \land x = y * (q+1) + r\} \tag{4}$$

$$\mathbf{q} := \mathbf{q} + \mathbf{1} \tag{5}$$

$$\{0 \le r \land 0 < y \land x = y * q + r\} \tag{6}$$

Claim:
$$\{0 \le r \land 0 < y \land x = y * q + r \land y \le r\}$$

 $r:=r-y; q:=q+1$
 $\{0 \le r \land 0 < y \land x = y * q + r\}$



Proof

$$\{0 \le r \land 0 < y \le r \land x = y * q + r\}$$

 $^{1)}$

$$\{0 \le r - y \land 0 < y \land$$

$$x = y * (q + 1) + (r - y)$$

$$r:=r-y;$$

(3)

(2)

$$\{0 \le r \land 0 < y \land x = y * (q+1) + r\}$$

(4)

$$q := q+1$$

(5)

$$\{0 \le r \land 0 < y \land x = y * q + r\}$$

(6)

While Statement

Let P be loop invariant. It holds every time the loop condition is bound Finel + tested.

$$\frac{\{P \land b\} \ S \ \{P\}}{\{P\} \ \text{while} \ b \ \text{do} \ S \ \text{od} \ \ \{P \land \neg b\}}$$

Proving Termination

Let t be bound function. Bound function is integer valued total function.

While Rule



$$Q \Rightarrow P$$

 $\{P \land b\} S \{P\}$
 $P \land \neg b \Rightarrow R$
 $P \land b \Rightarrow t > 0$
 $\{P \land b \land t = k\} S \{t < k\}$
 $\{Q\}$ while $b \text{ do } S \text{ od } \{R\}$

While Rule Intuition

Premises:

Initially the invariant holds. (PR1)

Each loop iteration preserves loop invariant. (PR2)

Each loop iteration decrements bound function from a positive value. (PR4)

Loop terminates before making bound function non-positive. (PR5)

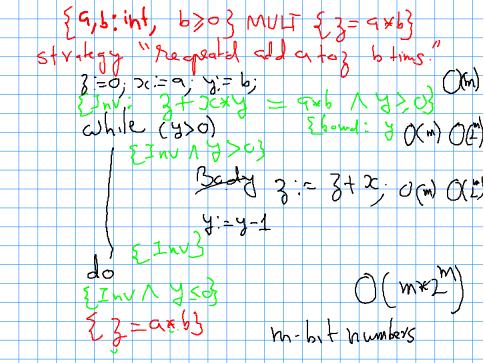
Conclusion:

On termination invariant must hold and also loop condition must be false. These together imply post condition by PR3.

The loop must terminate as bound function cannot decrement indefinitely from a positive value.

Annotated Program

```
\{0 < y \land 0 \le x\}
                                                   (1)
                                                   (2)
  r:=x; q:=0;
\{inv: 0 \le r \land 0 \le v \land x = v * q + r\}
  \{bound: r\}
                                                   (5)
  while \uparrow r\geqy do
                                                   (6)
        \{0 \le r \land 0 < y \land x = y * q + r \land y \le r\}
       r := r - y; q := q + 1
       \{0 \le r \land 0 < y \land x = y * q + r\}
       inv N
                                                   (10)
(11)
  \{x = v * a + r \land 0 < r < v\}
                                                   (12)
```



Efficient Multiplication

Efficient Multipliation (2)

```
Cowl
                                                     (1)
\{0 \le b\}
                                                     (2)
x:=a; y:=b; z:=0;
                                                     (3)
{inv: 0 \le y \land z + x * y = a * b}
                                                     (4)
{bound: y}
                                                            O(m)
                                                     (5)
while v > 0 do
                                                     (6)
       \{inv \land y > 0\}
       if even(y) then
                                                     (7)
                                                     (8)
               \{inv \land y > 0 \land even(x)\}
                                                     (9)
               x := x + x; y := y/2
               {inv}
                                                     (10)
                                                     (11)
       else
               \{inv \land y > 0 \land \neg even(x)\}
                                                     (12)
                                                     (13)
               y:=y-1; z:=z+x
               {inv}
                                                      (14)
       fi
                                                     (15)
       {inv}
                                                     (16)
od
                                                     (17)
\{inv \land y \leq 0\}
                                                     (18)
\{z = a * b\}
                                                     (19)
```

Founders of Formal Verification

First Order Logic for Assertions

Alan Turing



Tony Hoare



Edsgar Dijkstra



Other Contributors

- O.J. Dahl (Data structuring)
- S. Cook (Relative Completeness)

David Gries, The Science of Programming, Springer-Verlag.

Essentials of First-Order Predicate Logic

A language for describing mathematical structures.

```
A structure \mathcal{U} = (S, F, G)

S - set of values.

called Domain, written as |\mathcal{U}|

F - set of functions over S

G - set of relations over S
```

Pair (F, G) is called the signature.

Examples

```
\omega Natural Numbers Real Numbers Bool (\{0,1\},\ \{\land,\lnot\},\ \{=\})
```

Formalizing Properties of Structure

Stak

an colly

Some valid properties of ω

$$\forall y. \ (0 < y \lor 0 = y)$$

 $\forall x. \ x < x + 1$
 $\forall x, y, z. \ (x * (y + z) = x * y + x * z)$

$$div(x, y)$$
 means x "divides" y $div(x, y) \stackrel{\text{def}}{=} \exists z. \ x * z = y$

J.v (3,4)

prime(x) means x is a prime.

33, 3×3=4

$$prime(x) \stackrel{\text{def}}{=} \forall y.$$

$$(div(y,x) \Rightarrow y = 1 \lor y = x)$$

Domain $\{0,1,2,\ldots\}$ Functions 0, 1, +, * Relations <, =

What do f.o.l. formulas over ω look like?

Terms

• Examples:

```
Domain \{0,1,2,\ldots\}
Functions 0, 1, +, *
Relations <, =
```

What do f.o.l. formulas over ω look like?

- Examples: x+0*y 1*z
- Syntax: $t ::= x \mid f(t_1, \ldots, t_n)$

```
Domain \{0,1,2,\ldots\}
Functions 0, 1, +, *
Relations <, =
```

What do f.o.l. formulas over ω look like?

- Examples: x+0*y 1*z
- Syntax: $t ::= x \mid f(t_1, \ldots, t_n)$
- State (valuation) $\sigma: Var \rightarrow |\mathcal{U}|$. E.g. $\sigma(x) = 3, \sigma(y) = 4, \sigma(z) = 2$.

```
\begin{array}{lll} \text{Domain } \{0,1,2,\ldots\} \\ \text{Functions } 0,\ 1,\ +,\ * \\ \text{Relations } <,\ = \end{array}
```

What do f.o.l. formulas over ω look like?

- Examples: x+0*y 1*z
- Syntax: $t ::= x \mid f(t_1, \ldots, t_n)$
- State (valuation) $\sigma: Var \rightarrow |\mathcal{U}|$. E.g. $\sigma(x) = 3, \sigma(y) = 4, \sigma(z) = 2$.
- Value of term t in structure $\mathcal U$ and state σ is $\hat{\sigma}(t) \in |\mathcal U|$

```
Domain \{0,1,2,\ldots\}
Functions 0, 1, +, *
Relations <, =
```

What do f.o.l. formulas over ω look like?

- Examples: x+0*y 1*z
- Syntax: $t ::= x \mid f(t_1, \ldots, t_n)$
- State (valuation) $\sigma: Var \rightarrow |\mathcal{U}|$. E.g. $\sigma(x) = 3, \sigma(y) = 4, \sigma(z) = 2$.
- Value of term t in structure $\mathcal U$ and state σ is $\hat{\sigma}(t) \in |\mathcal U|$
- $\hat{\sigma}(x+0*y) = 3+0*4 = 3.$

```
Domain \{0,1,2,\ldots\}
Functions 0, 1, +, *
Relations <, =
```

What do f.o.l. formulas over ω look like?

- Examples: x+0*y 1*z
- Syntax: $t ::= x \mid f(t_1, \ldots, t_n)$
- State (valuation) $\sigma: Var \rightarrow |\mathcal{U}|$. E.g. $\sigma(x) = 3, \sigma(y) = 4, \sigma(z) = 2$.
- Value of term t in structure $\mathcal U$ and state σ is $\hat{\sigma}(t) \in |\mathcal U|$
- $\hat{\sigma}(x+0*y) = 3+0*4 = 3.$
- Semantics: $\hat{\sigma}(x) = \sigma(x)$ $\hat{\sigma}(f(t_1, ..., t_n)) = f(\hat{\sigma}(t_1), ..., \hat{\sigma}(t_n))$

First order logic (cont)

Atomic Formulae

• Example: x + 0 * y < 1 * z

- Example: x + 0 * y < 1 * z
- Syntax: ψ ::= $t_1 = t_2 \mid R(t_1, \ldots, t_n)$

- Example: x + 0 * y < 1 * z
- Syntax: ψ ::= $t_1 = t_2 \mid R(t_1, ..., t_n)$
- $\mathcal{U}, \sigma \models \psi$ denotes that ψ evaluates to true in \mathcal{U}, σ . $\mathcal{U}, \sigma \not\models \psi$ denotes that ψ evaluates to false in \mathcal{U}, σ .

- Example: x + 0 * y < 1 * z
- Syntax: ψ ::= $t_1 = t_2 \mid R(t_1, ..., t_n)$
- $\mathcal{U}, \sigma \models \psi$ denotes that ψ evaluates to true in \mathcal{U}, σ . $\mathcal{U}, \sigma \not\models \psi$ denotes that ψ evaluates to false in \mathcal{U}, σ .
- Let $\sigma(x) = 3$, $\sigma(y) = 4$, $\sigma(z) = 2$. Then, ω , $\sigma \not\models (x + 0 * y < 1 * z)$. (why?)

- Example: x + 0 * y < 1 * z
- Syntax: ψ ::= $t_1 = t_2 \mid R(t_1, ..., t_n)$
- $\mathcal{U}, \sigma \models \psi$ denotes that ψ evaluates to true in \mathcal{U}, σ . $\mathcal{U}, \sigma \not\models \psi$ denotes that ψ evaluates to false in \mathcal{U}, σ .
- Let $\sigma(x) = 3$, $\sigma(y) = 4$, $\sigma(z) = 2$. Then, $\omega, \sigma \not\models (x + 0 * y < 1 * z)$. (why?)
- Semantics:

$$\mathcal{U}, \sigma \models t_1 = t_2 \quad \mathbf{iff} \quad \hat{\sigma}(t_1) = \hat{\sigma}(t_2)$$

$$\mathcal{U}, \sigma \models R(t_1, \dots, t_n) \quad \mathbf{iff}$$

$$R(\hat{\sigma}(t_1), \dots, \hat{\sigma}(t_n))$$

• Formula ϕ is made of atomic formulas using boolean connectives $\land, \lor, \neg, \Rightarrow$ as well as quantifiers $\exists x. \phi$ and $\forall x. \phi$.

- Formula ϕ is made of atomic formulas using boolean connectives $\land, \lor, \neg, \Rightarrow$ as well as quantifiers $\exists x. \phi$ and $\forall x. \phi$.
- E.g. $(\forall y. (x < y \lor x = y)).$

- Formula ϕ is made of atomic formulas using boolean connectives $\land, \lor, \neg, \Rightarrow$ as well as quantifiers $\exists x. \phi$ and $\forall x. \phi$.
- E.g. $(\forall y. (x < y \lor x = y)).$
- Syntax: ϕ ::= $\psi \mid \phi_1 \land \phi_2 \mid \neg \phi \mid \exists x.\phi$.

- Formula ϕ is made of atomic formulas using boolean connectives $\land, \lor, \neg, \Rightarrow$ as well as quantifiers $\exists x. \phi$ and $\forall x. \phi$.
- E.g. $(\forall y. (x < y \lor x = y)).$
- Syntax: ϕ ::= $\psi \mid \phi_1 \land \phi_2 \mid \neg \phi \mid \exists x. \phi$.
- $\mathcal{U}, \sigma \models \phi$ denotes that ϕ evaluates to true in \mathcal{U}, σ .

- Formula ϕ is made of atomic formulas using boolean connectives $\land, \lor, \neg, \Rightarrow$ as well as quantifiers $\exists x. \phi$ and $\forall x. \phi$.
- E.g. $(\forall y. (x < y \lor x = y)).$
- Syntax: ϕ ::= $\psi \mid \phi_1 \land \phi_2 \mid \neg \phi \mid \exists x. \phi$.
- $\mathcal{U}, \sigma \models \phi$ denotes that ϕ evaluates to true in \mathcal{U}, σ .
- Formula $\exists x.\phi$ states that there exists a choice of value of x (ignoring the value given by $\sigma(x)$) which makes ϕ true. Formula $\forall x.\phi$ states that all choice of value of x (ignoring the value given by $\sigma(x)$) make ϕ true.

- Formula ϕ is made of atomic formulas using boolean connectives $\land, \lor, \neg, \Rightarrow$ as well as quantifiers $\exists x. \phi$ and $\forall x. \phi$.
- E.g. $(\forall y. (x < y \lor x = y)).$
- Syntax: ϕ ::= $\psi \mid \phi_1 \land \phi_2 \mid \neg \phi \mid \exists x. \phi$.
- $\mathcal{U}, \sigma \models \phi$ denotes that ϕ evaluates to true in \mathcal{U}, σ .
- Formula $\exists x. \phi$ states that there exists a choice of value of x (ignoring the value given by $\sigma(x)$) which makes ϕ true. Formula $\forall x. \phi$ states that all choice of value of x (ignoring the value given by $\sigma(x)$) make ϕ true.
- Let $\sigma(x) = 0$. Then, $\omega, \sigma \not\models (\forall y. (x < y \lor x = y))$. (why?)

- Formula ϕ is made of atomic formulas using boolean connectives $\land, \lor, \neg, \Rightarrow$ as well as quantifiers $\exists x. \phi$ and $\forall x. \phi$.
- E.g. $(\forall y. (x < y \lor x = y)).$
- Syntax: ϕ ::= $\psi \mid \phi_1 \land \phi_2 \mid \neg \phi \mid \exists x. \phi$.
- $\mathcal{U}, \sigma \models \phi$ denotes that ϕ evaluates to true in \mathcal{U}, σ .
- Formula $\exists x. \phi$ states that there exists a choice of value of x (ignoring the value given by $\sigma(x)$) which makes ϕ true. Formula $\forall x. \phi$ states that all choice of value of x (ignoring the value given by $\sigma(x)$) make ϕ true.
- Let $\sigma(x) = 0$. Then, $\omega, \sigma \not\models (\forall y. (x < y \lor x = y))$. (why?)
- Semantics: σ' is x-variant of σ if $\sigma(y) = \sigma'(y)$ for all $y \neq x$.

$$\mathcal{U}, \sigma \models \exists x. \phi \quad \text{iff}$$

$$\mathcal{U}, \sigma' \models \phi \quad \text{for some x-variant σ' of σ}$$

Quicksort

Sorted
$$(A, i, j) \stackrel{\text{def}}{=}$$

$$1 \le i \le j \le n \quad \Rightarrow \quad \forall i'. i \le i' < j \Rightarrow A[i'] \le A[i'+1]$$
Partition $(A, i, j, k) \stackrel{\text{def}}{=}$

$$1 \le i \le j \le k \le n \quad \land \quad (\forall i'. (i \le i' < j \Rightarrow A[i'] \le A[j]) \quad \land$$

 $(\forall k'.(j < k' < k \Rightarrow A[i] < A[k'])$

Then,

$$\models \left(egin{array}{c} \textit{Partition}(A,i,j,k) \\ \land & \textit{Sorted}(A,i,j-1) \\ \land & \textit{Sorted}(A,j+1,k) \end{array} \right) \Rightarrow \textit{Sorted}(A,i,k)$$