TATA Institute Bombay Jan 1978

THEORY OF PROGRAMMING

(1)

and its application to the design of correct and efficient computer programs.

A COURSE OF LECTURES

given by Professor C.A.R. Hoare

Professor of Computation at Oxford University

BASED ON THE TEXT

(2)

a

discipline

of

programming

Edsger W. Dijkstra

Prentice Hall, 1976

BASIC DEFINITIONS

(3)

A number is denoted by a string of digits

e.g. 0

3

703

A variable is denoted by a string of letters

03

e.g. X Y ALPHA QUOT

A machine state is a finite mapping between variables and values.

$$m(X) = 37, m(Y) = 7, m(QUOT) = 0$$

m(REM) is undefined. m is also undefined for all other variables.

A command C is a relation between machine states, i.e. the states of (4) the machine before and after the command is obeyed.

(we define a relation as a set of ordered pairs $(m,m') \in C$).

The skip command is defined:

i.e. the identity relation.

This command is obeyed by doing nothing!

The abort command is defined

i.e. the empty relation.

A machine which attempts to obey this command will simply fail (break)!

An <u>expression</u> is constructed from variables, values, operators, and (5) brackets.

e.g.
$$X = 17 X + 17 Y*(X - 3)$$

Given machine state m and expression e, we define m*(e) - as the value taken by e, when evaluated in machine state m; i.e. when its variables are replaced by the values given by m,

e.g. if
$$m(X) = 3$$
 then $m^*(X + 17) = 20$
 $m^*(17) = 17$

if m(Y) is not defined, then nor is

$$m^{*}(Y^{*}(X + 3))$$

x := e (x becomes y)

where x is a variable

and e is an expression

e.g.
$$Y := 17$$
 $Y := X - 1$ $X := X + 1$.

$$x := e = df \{ (m,m') \mid m'(x) = m*(e).$$

&
$$\forall y \neq x m^t(y) = m(y)$$

It is obeyed by evaluating e in the initial machine state m, and then changing m(x) to have this value instead of its old one. If e is undefined in m, the machine breaks.

The composition of commands c1 and c2 is:

(7)

c1; c2 =
$$\{(m,m') \mid \exists m'' (m,m'') \in c1 \& (m'',m') \in c2 \}$$

It is obeyed by first obeying c1 and then obeying c2. m" is the final machine state of c1 and the initial machine state of c2.

Theorem. c1;(c2; c3) = (c1; c2); c3

The associativity of ; will justify omission of brackets.

Theorem. skip; c = c; skip = c

abort; c = c; abort = abort

Compare: 1*c = c*1 = c

0 * c = c * 0 = 0

A condition b is an expression which is either true \underline{T} or false \underline{F} (8) e.g. X 0 X = Y \underline{T} F

It defines a "command"

$$\{(m,m') \mid m^*(b) = \underline{T} \& m' = m \}$$

i.e. the identity relation restricted to those machine states in which b is true.

It is obeyed by evaluating b; if this is true, skip; otherwise abort.

A conditional command is defined

if
$$b \rightarrow c1$$
, $c2 \underline{fi} = b$; $c1 \cup b$; $c2$

It is obeyed by first evaluating b; if the value is true, c1 is obeyed and c2 omitted. if the value is false, c2 is obeyed and c1 omitted.

if $b \rightarrow skip$, abort $\underline{fi} = b$ Theorems if $b \rightarrow c1$, $c2 \underline{fi}$; $c3 = \underline{if} b \rightarrow (c1; c3)$, $(c2; c3) \underline{fi}$

if $b \rightarrow c1$, $c2 \underline{fi} = \underline{if} b \rightarrow c2$, $c1 \underline{fi}$

 $\underline{if} \quad \underline{T} \rightarrow c1$, $c2 \quad \underline{fi} = \underline{if} \quad \underline{F} \rightarrow c2$, $c1 \quad \underline{fi} = c1$.

The repetitive command is defined $\frac{do \ b \rightarrow c \ od}{n = 0} = \bigcup_{n=0}^{\infty} c_n \qquad (\underline{\text{while } b \ do \ c})$ where $c_0 = \overline{b}$

$$c_{n+1} = b; c; c_n \left[\begin{array}{c} \overline{b} \\ \end{array} \right]$$

It is obeyed by first evaluating b. If this is false, the task is finished. If it is true, c is next obeyed, and then the whole command is repeated.

Theorem. $do b \rightarrow c od = if b \rightarrow (c; do b \rightarrow c od)$, skip fi

 $do F \rightarrow c od = skip$

do T -> c od = abort

 $\underline{do} b \rightarrow c \underline{od} = \overline{b}, b; c; b, b; c; b; c; b, ...$

 $c_n \subseteq c_{n+1}$

 $c_n = b; c; b; c \dots b; c; b$

 \leq n times.

$$x := X; y := Y;$$

$$\underline{do} x \neq y \rightarrow \underline{if} x < y \rightarrow y := y - x, x := x - y$$

$$\underline{fi}$$

<u>od</u>

= x := X; y := Y;

$$\{x = y \ U \ (x \neq y; x < y; y := y - x; x = y)\}$$

 $U \ (x \neq y; x < y; x := x - y; x = y)$

 $0 \times \neq y$; x < y; y := y - x; $x \neq y$; x < y; x := x - y; x = y

.

EXECUTION TRACES

(11)

	EXECUTION	TRACES		•	
	X	Y	x	у	
initial m/s	111	259			
x := X	111	2 59	111		
у := Ү	11	ff	n	259	
x ≠ y ✓					
$x < \lambda$					
y := y-x				148	
x ≠ y; x <y td="" ✓<=""><td></td><td></td><td></td><td></td><td></td></y>					
у := у- х				37	
x ≠ y; x<y< del=""> ✓</y<>		,			
x := x -y			74		
$x \neq y$; $\overline{x \langle y}$				ı	
$\mathbf{x} := \mathbf{x} - \mathbf{y}$			37	•	
x = y					
final m/s.	111	259	37	37	`

A. Obey the following commands, giving their execution traces.

(12)

1.
$$y := 0$$
; $x := 3$; do $y + 1 < x \rightarrow y := y + 1$ od

2.
$$y := 2$$
; $x := 3$; do $y > 0 \rightarrow x := x + 1$; $y := y - 1$ od

3. y := 2; x := 3; s := 0;

$$do y>0 \rightarrow y := y-1; s:= s + x od$$

4. y := 3; x := 7; q := 0;

$$\underline{do} \times y \rightarrow x := x - y; q := q + 1 \underline{od}$$

B. In each example above, replace the repetitive command by a pair of assignments which would have the same effect for <u>any</u> initial assignment to x, y, q, s.

e.g. the answer to example 4 is $q := q + x \div y$; x := x modulo y

WEAKEST PRECONDITIONS

(13)

c achieves r is defined:

$$c \underline{a} r = \{(m,m) \mid \exists m' (m,m') \in c \& m'*(r) = \underline{r} \}$$

i.e. the inverse image of r under c. It is the condition satisfied by exactly those initial machine states in which c can successfully be obeyed, and can end in a machine state satisfying r.

Theorem $b \underline{a} r = b n r$

b&r

where b is a condition.

Proof. LHS =
$$\{ mm \mid \exists m' \ (m,m') \in b \& m'*(r) = \underline{T} \}$$

= $\{ mm \mid (m,m) \in b \& m*(r) = \underline{T} \}$
= $\{ mm \mid (m,m) \in B \& (m,m) \in r \}$
= $\{ mm \mid (m,m) \in B \& (m,m) \in r \}$

We usually identify a condition with the set of machine state pairs in which the condition is true.

$$\underline{\mathbf{T}} \mathbf{a} \mathbf{r} = \mathbf{r}$$

Strict
$$e \underline{a} \underline{F} = \underline{F}$$
, $\underline{F} \underline{a} \underline{r} = \underline{F}$

Distributive $c = (\bigcup_{n} r_{n}) = \bigcup_{n} (c = r_{n})$

(Additive)
$$(\bigcup_{n} c_{n}) \ge r = \bigcup_{n} (c_{n} \ge r)$$

A command c is <u>deterministic</u> if for each initial state m, there is at most <u>one</u> final state m' such that $(m,m') \in c$. All commands defined so far are deterministic.

If c is deterministic.

Multiplicative
$$c_{\underline{a}}(\bigcap_{n} r_{n}) = \bigcap_{n} (c_{\underline{a}} r_{n})$$

This is not true if c is nondeterministic,

e.g.
$$c = x := 0 \cup x := 1$$

then
$$c \underline{a} x = 0 = \underline{T} & \underline{a} x = 1 = \underline{T}$$

 $c \underline{a} (x = 0 & x = 1) = c \underline{a} \underline{F} = \underline{F}$

$$x := e \underline{a} b(x) = b(e)$$

where b(e) is the result of replacing all occurrences of x in b(x) by e.

e.g.
$$(X := 37 \underline{a} X)12 = \underline{T}$$

 $(X := Y*3 \underline{a} X)12 = Y*3 > 12 = \underline{Y}$
 $(X := X - 1 \underline{a} X)12 = (X - 1 > 12) = (X > 13)$

Proof. The value of x after the assignment is by definition equal to the value of e before the assignment. So b(x) is true (of x) after the assignment if any only if b(e) is true (of the value of e) before the assignment. The values of all other variables of b remain unchanged by the assignment.

Theorem of composition

(16)

Proof. c1; c2 arrives at a state satisfying r if c1 arrives at a state from which c2 achieves r, and conversely

LHS =
$$\begin{cases} mn & \exists m' (m,m') \in (c1; c2) & \&r*(m') = T \end{cases}$$

= $\begin{cases} mn & \exists m' (m,m') \in (c1; c2) & \&r*(m') = T \end{cases}$
= $\begin{cases} mn & \exists m' (m,m') \in c1 & \&(m'',m') \in c2 & \&r*(m') = T \end{cases}$
= $\begin{cases} mn & \exists m''(m,m'') \in c1 & \& \exists m''(m'',m') \in c2 & \&r*(m') = T \end{cases}$
= $\begin{cases} mn & \exists m''(m,m'') \in c1 & \& m'' \in (c2 a r) \end{cases}$ = RHS

(17)

if $b \rightarrow c1$, $c2 \underline{fi} \underline{a} \underline{r} = \underline{if} \underline{b} \rightarrow (c1 \underline{a} \underline{r}), (c2 \underline{a} \underline{r}) \underline{fi}$ = $b \cap (c1 \underline{a} \underline{r}) \cup b \cap (c2 \underline{a} \underline{r})$

e.g.
$$(\underline{\text{if }} x \langle y \rightarrow y := y - x, x := x - \underline{yfi} \underline{a} GCD(x,y) = K) \equiv \underline{\text{if }} x \langle y \rightarrow GCD(x,y-x) = K, GCD(x-y, y) = K \underline{fi}$$

$$GCD(x,y) = K \& x \neq y$$

Proof. LHS = (b; c1
$$\cup$$
 b; c2) a r
= (b; c1)a r \cup (b; c2) a r
= b a (c1 a r) \cup b a(c2 a r)
= b \(\cap (c1 a r) \cup b \(\overline{0}\) (c2 a r)
= RHS.

Theorem of repetition.

 $\frac{do b \Rightarrow c \text{ od a } r = 0}{n = 0} P_n$

where Po = \overline{b} r

Pn+1 = $b \cap c\underline{a}$ Pn $\bigcup \overline{b} \cap r$ e.g. \underline{do} y+1 $\langle x \rightarrow y := y+1 \underline{od} \underline{a} y = x - 1 = \bigcup_{n=0}^{\infty} Pn$

where $Po = \overline{y+1} \langle x \cap y = x - 1 \equiv y = x - 1$ $P1 \equiv y+1 \langle x \cap (y := y+1 \ge Po) \cup y = x - 1$ $\equiv y = x - 2 \forall y = x - 1$ $\equiv x - 1 - 1 \langle y \langle x \rangle$

Pn
$$\equiv x - n - 1 \leqslant y \leqslant x$$

$$Pn \equiv \exists n \quad x - n - 1 \leqslant y \leqslant x$$

$$\equiv y \leqslant x$$

LHS =
$$\bigcup_{n} e_{n} \underline{a} r$$
 where $e_{0} = \overline{b}$, $e_{n+1} = b; e; e_{n} \cup \overline{b}$
= $\bigcup_{n} (e_{n} \underline{a} r)$ where $e_{0} \underline{a} r = \overline{b} \cap r$, $e_{n+1} \underline{a} r = (b; e; e_{n} \underline{a} r) \cup \overline{b} \cap r$
= $b \cap (e \underline{a} (e_{n} \underline{a} r)) \cup \overline{b} \cap r$
= $\bigcup_{n} (e_{n} \underline{a} r) \cup \overline{b} \cap r$

EXERCISES (18)

1. Derive and simplify the following preconditions:

a.
$$(y := 0; \underline{do} y+1 < x \rightarrow y := y+1 \underline{od}) \underline{a} (y = x - 1)$$

b.
$$x := X$$
; $y := Y$; $do y > 0 \Rightarrow x := x+1$; $y := y - 1 od a x = X + Y + Z$

c.
$$y := Y$$
; $s := 0$; $do y > 0 \rightarrow y := y - 1$; $s := s + X od a s = X*Y$

d. x := X; q := 0; $do x \ge y \rightarrow x := x - y$; q := q+1 od a X = q*y+x & x < y (all variables are nonnegative integers).

We can now solve problems of the form:

(19)

given command S and postcondition r,

$$? = S a r$$

But programmers must solve a different problem:

given postcondition r and precondition p,

$$p \Rightarrow ? \underline{a} r$$

e.g. using only +1 and < , write a command c which does not change x, and which satisfies:

$$0 < x \Rightarrow c = y = x - 1$$

This is usually more difficult! We shall need some more theory.

If p and b are conditions

$$p \cap b = p$$
; $b = p \underline{a} b = b \underline{a} p = b$; p

$$c; (c1 \cup c2) = c; c1 \cup c; c2$$

$$(c1 \cup c2); c = c1; c \cup c2; c$$

if
$$p(b) \Rightarrow c1 \underline{a} r$$

and
$$p(l) \stackrel{-}{b} \implies c2 \underline{a} r$$

then
$$p \implies \underline{if} b \rightarrow c1, c2, \underline{fi} \underline{a} r$$

THEOREM OF INVARIANCE Let c be \underline{do} b \rightarrow c1 \underline{od} (for deterministic c1) and let b/p clap ... (1) (i.e. p is an invariant of then $p \cap (c \underline{a} \underline{T}) \subseteq c \underline{a} (\overline{b} \cap p)$ Proof. define $c_0 = \overline{b}$, $c_{n+1} = b$; c1; $c_n \cup \overline{b}$ so $c = \bigcup_{n} c_n$ LHS = $p \cap (\bigcup c_n) \underline{a} \underline{T} = p \cap \bigcup (c_n \underline{a} \underline{T})$ distribution $= \bigcup_{n} p \cap (c_{n} \underline{a} \underline{T}) \subseteq \bigcup_{n} (c_{n} \underline{a} (\overline{b} \cap p))$ (by lemma) $= (\bigcup c_n) \underline{a} (\overline{b} \cap p) = RHS.$ distribution LEMMA (23)If c is deterministic (1)and $p(h) b \subseteq cap$ (2)and $c_0 = \overline{b}$ and $c_{n+1} = b; c; c_n \cup \overline{b}$ (3)then $\forall n \ p \cap (c_n \underline{a} \underline{T}) \subseteq c_n \underline{a} (\overline{b} \cap p)$ (4)Proof case-1 n = 0: $p \cap (\overline{b} \underline{a} \underline{T}) = p \cap \overline{b} = \overline{b} \cap (p \cap \overline{b}) = \overline{b} \underline{a}(\overline{b} \cap p)$ $p \cap (c_{n+1} \underline{a} \underline{T}) = p \cap (b; c; c_n) \underline{a} \underline{T} \bigcup p \cap (\overline{b} \underline{a} \underline{T})$ by(3) \subseteq b((c a p) ((c a(c, a I))) p(b by(2)

=
$$b \cap (c_{\underline{a}}(p \cap (c_{\underline{n}} \underline{a} \underline{T}))) \cup p \cap \overline{b}$$
 by (1)

$$\stackrel{\subseteq}{=} b \cap (c \underline{a} (c_{\underline{n}} \underline{a} (\overline{b} \cap \underline{p}))) \underline{b} \underline{a} (\overline{b} \cap \underline{p}) \qquad by(4)$$

$$= (b; c; c_{\underline{n}}) \underline{a} (\overline{b} \cap \underline{p}) \underline{u} \underline{b} \underline{a} (\overline{b} \cap \underline{p})$$

$$= ((b; c; c_{\underline{n}}) \underline{u} \overline{b}) \underline{a} (\overline{b} \cap \underline{p})$$
(3)

$$= c_{n+1} \underline{a} (\overline{b} \mathbf{a} p)$$

Let t be an expression, (always defined)

$$c \underline{dec} t =_{df} (k := t; c) \underline{a} (0 \leqslant t \leqslant k)$$

(where k is a fresh variable).

Theorem 2. Let $c = \underline{do} b \Rightarrow c1 \underline{od}$ (deterministic)

p∩b ⊆ clap

p∩b ⊆ c1 dec t

then $p \subseteq c_{\underline{a}}(\overline{b}(p))$

Proof define $c_0 = \overline{b}$, $c_{n+1} = b$; c; $c_n \cup \overline{b}$ by Lemma 2 $p \cap (t \le n) \subseteq c_n \underline{a} \underline{T}$

$$p = \bigcup_{n} (p \cap t \leq n) \leq p \cap \bigcup_{n} (e_{n} \underline{a} \underline{T}) = p \cap e \underline{a} \underline{T}$$

 $\subseteq c \underline{a} (\overline{b} \cap p)$

by theorem (1)

LEMMA

(25)

If
$$p \cap b \subseteq (k := t; c) \underline{a} (p \cap t < k)$$
 for fresh k (1)

and
$$c_0 = \overline{b}$$
, $c_{n+1} = b; c; c_n \cup \overline{b}$ (2)

then
$$\forall n \ p \cap t \leq n \subseteq c_n \underline{a} \underline{T}$$
 (3)

Proof. $p \cap b \subseteq (k ;= t;c) \underline{a} t < k \subseteq t > 0$

induction step.

$$p \cap t \leq n+1 = p \cap t = n+1$$

vp∩t≤n

$$\subseteq p \cap b \cap t = n+1 \cup \overline{b}$$

) i

$$\subseteq b \cap t = n+1 \cap (k := t; c) \underline{a} (p \cap t < k) \cup \overline{b} \cup$$

$$\subseteq b \cap c \underline{a}(p \cap t < n+1) \cup \overline{b}$$

U potsn

U cna I

$$= c_{n+1} \underline{a} \underline{T}$$

since $c_n \subseteq c_{n+1}$

Using only<, successor as operators, find c s.t.

$$0 < x \implies c \underline{a} y = x - 1 \& c$$
 changes only y.

Solution: reformulate postcondition as

$$y + 1 < x \cap y < x$$
 = $\overline{b} \cap p$

and find c1, t s.t.

$$y + 1 \langle x \wedge y \langle x \Longrightarrow c1 \underline{dec} t$$

try c1 = y := y +1

$$t = x - y$$

check
$$y + 1 \le x \implies (k := x - y; y := y + 1) \underline{a} (x - y \le k \cap y \le x)$$

$$RHS = x - y - 1 \le x - y$$

... by theorem 2.

$$y < x \implies \underline{do} y + 1 < x \implies y := y + 1 \underline{od} \underline{a} y = x - 1$$

It remains to find co s.t.

$$0 < x \implies c0 a y < x$$
.

Using theorem of assignment, c0 is y := 0.

Using only (, successor, and predecessor as operators find c s.t.

 $c \underline{a} x = X + Y$ where c changes only x and y.

Solution: reformulate postcondition as

$$\overline{b} \cap p = \overline{0 < y} \cap x + y = X + Y$$

check that
$$x := X$$
; $y := Y \underline{a} p$. (1)

we need to find c1, t s.t.

$$0 < y \cap x + y = X + Y \Longrightarrow (k := t; c1) \underline{a} (t < k \cap p)$$

an obvious choice is t = y.

$$c1 = c2; y := y - 1$$
 ... (2)

where $0 < y \cap p =$ c2 <u>a</u> $(y := y - 1 \underline{a} \times + y = X + Y)$ i.e. $0 < y \cap (x + y = X + Y) =$ c2 <u>a</u> (x + y - 1 = X + Y)obviously c2 = x := x + 1 (3) collecting (1), (2), (3), we get: $x := X; y := Y; \underline{do} 0 < y \rightarrow x := x + 1; y := y - 1 \underline{od}$

EXAMPLE 3

(28)

Using only (, +, - , find c s.t.

 $Y > 0 \implies c = q = X - Y \cap r = X \text{ modulo } Y \cdot c \text{ changes only } q \text{ and } r$ Reformulate postcondition as

$$\overline{Y \leqslant r} \cap X = q * Y + r \quad (= \overline{b} \cap p)$$
note that $(q :=0; r := X) \underline{a} p$ (1)

choose r as variant function.

$$b \implies r := r - Y \underline{dec} \quad r \tag{2}$$

we need to find c1 changing only q, s.t.

$$b \cap p \implies c1 = (X = q * Y + r - Y)$$

this is solved by c1 = q := q + 1.

...
$$c = q := 0$$
; $r := X$; do $Y \le r \rightarrow q := q + 1$; $r := r - y$ od Exercise: using only $(, +, -, f)$ find $c = x$.

 $c \underline{a} s = X * Y$, and c changes only s and y.