

J.A.R. HOARE

ЗY

IMPLEMENTATIONS

AND

PROGRAMS

SPECIFICATIONS,

C 1982 C.A.R. Hoare

Programming Research Group Oxford University Computing Laboratory 45 Banbury Road Oxford OX2 6PE United Kingdom

Summary

A specification is a predicate describing all observations permitted of the system specified. Specifications of complex systems can be constructed from specifications of their components by connectives defined in the predicate calculus. A program is just a predicate expressed using only a restricted subset of such connectives, codified as a programming language. An implementation of the programming language is a mechanism that will accept any predicate of the language, and then behave as described by it. Given a proposed model of an implementation it is desirable to prove that every program expressible in the language is consistent and complete with respect to the model; furthermore, there should be no program which logically implies all the others. These points are illustrated by the design of a very simple programming language, describing the interactions of concurrent processes. It is suggested that the design of a realistic programming language requires, and is worthy of, the skills of a methematical logician.

Contents

P	aq	e

1.	Introduction	
2.	Interacting Processes	3
	2.1 No action	4
	2.2 Arbitrary Action	4
	2,3 First action	5
	2.4 Recursion	5
	2.5 Alternatives	6
	2.5 Concurrency	7
3.	Example: the Dining Philosophers	9
	3.1 Alphabets	9
	3.2 Behaviour	10
	3.3 Deadlock:	1 1
	3.4 Deadlock averted	11
٩.	Mathematical Properties	12
	4.1 Algebraic properties	12
	4.2 Ordering properties	14
	4.3 Continuity	15
	4.4 Unique solutions	17
5.	Programming	18
	5.1 Programming methodology	19
	5.2 Top-down development	19
	5.3 Introduction of recursion	20
6.	Implementation	
	6, [†] Consistency	22
	5.2 Completeness	23
	5.3 The excluded miracle	24
	6.4 Computability	26
	6.5 Continuity again	27
7.	Conclusion	
Ack	nowladgementa	29

SPECIFICATIONS, PROGRAMS AND IMPLEMENTATIONS

1. Introduction

A <u>variable</u> in a formula of applied mathematics stands for some directly or indirectly observable value. The correspondence between the variable and the observation must be established informally by some such phrase as "Let \underline{x} stand for the position (in metres) of a body at time \underline{t} (measured in seconds), and let \underline{v} stand for its velocity (in metres per second)".

"Let <u>coin</u> stand for the number of coins inserted into a vending machine up to a certain moment, and let <u>choc</u> stand for the number of chocolates it has dispensed".

The <u>alphapet</u> of a system is the set of variables denoting those observations which are of current interest. The selection of a user, 1 and relevant alphabet of observations is one of the primary characteristics of a successful scientific theory. A <u>specification</u> of a system S is defined by its alphabet \prec S together with a predicate, usually containing variables from \prec S, which describes all possible observations which may be made of the system. An observation <u>ascribas</u> a value to each variable in the alphabet; a specification <u>describes</u> the observation if the predicate evaluates to true when each variable is replaced by its ascribed value. We use S itself to stand for the predicate. Example 1.

LINEAR specifies a body that moves at a constant speed of 3 metres per second between time 1 and time 4. Some observations described by this specification are tabulated below:

×	t	v
0	1	2
3	2	3
6	3	3
12	٥	17
14	5	12
٥	5	0

The first three lines make the consequent of the specification true, and the last three make the antecedent false. When t is outside the specified range, the specification is indeterminate: it specifies nothing at all about the values of x and y.

Example 2.

PROFIT → (choc ≤ coin) BUFFER1 ← (coin ≼ choc+1) VM1 → PROFIT & BUFFER1 ≺VM1 = {coin,choc}

The predicate PROFIT specifies that a vending machine shall be profitable in the sense that it never dispenses more chocolates than there have been coins inserted. BUFFER1 states that the machine will buffer only one coin, so it is impossible to insert two coins and then extract two chocolates. Some possible observations are:

coin	choc	
0	0	
Т	٥	
1	1	
13	12	

By a bold abstraction, we now say that a system is fully defined by the strongest specification which describes its every possible behaviour. Thus two systems are regarded as the same if their alphabets are the same, and their specifications are logically equivalent. This is reasonable, for then there can be no observation of one of them which is not also a possible observation of the other, and by the principle of the identity of indiscernibles, they should be regarded as the same.

Let 5 be an arbitrary specification, and let A be the strongest specification of some actual system, and suppose that A logically implies 5. This means that every observation described by A is also described by 5. Thus we can claim that A is a <u>correct</u> implementation of the specification 5. The suggestion of this and the previous paragraph is that questions of equivalence and correctness of systems can be treated within the traditional framawork of mathematical logic without requiring any specialised axioms or proof rules. This suggestion is explored in the next section by defining predicates which describe the behaviour of systems composed from interacting concurrent processes.

2. Interacting Processes

We are interested in systems which engage in certain observable events, such as the insertion of coins into a vending machine, or extraction of chocolates. For each event individually, it is possible to record how many times that event has occurred up to any given moment. The alphapet of a process contains variables standing for these counts. For example, the alphabet of a vending machine may be declared as $\not\exists VM1 \Rightarrow \{coin, choc\}$ 2.1 No action

```
Let A be the alphabet \{a,b,c, \ldots\}
We define the prodicate
ZERO<sub>A</sub> \triangleq (a = b = c = ... = 0)
\prec ZERO<sub>A</sub> \triangleq A
```

This is the specification of a rather useless system, which never does anything at all, so thet its event counts remain forever zero.

```
Example VMO 4 ZERD {coin, choc}
```

This describes the behaviour of a broken vending machine. Its alphabet indicates that it is equipped with the physical organs for accepting Coine and dispensing chocolates, but its pradicate states that it never uses them.

2.2 Arbitrary action We define the predicate CHAOS_A ← true ≺CHAOS_A ← A

This is a specification which places no constraint whatsoever on the behaviour of the specified system. Every system is correct in accordance with this specification.

Example

VMBR ♀ CHAOS {coin,choc}

This machine is even more badly broken than VMO; it accepts coins and dispenses chocolates with gay abandon.

4.

2.3 First action

Let $a \in A$ Let P(a) be a predicate with alphabet A $(a;P(a)) \cong ZERO_n \lor a > 0 \& P(a-1)$

This specifies a system which first engages in the event counted by the variable "a", and then behaves as specified by P. On the first observation, all the counts are zero; on all subsequent observations, the count of "a" is positive; furthermore, on reducing the count of "a" by one, we get an observation described by P.

Example

Recall that VM1 $rightarrow (\operatorname{cnoc} \leq \operatorname{coin} \leq \operatorname{choc} + 1)$ \therefore (coin; choc; VM1) \Leftrightarrow (coin = choc = 0 \checkmark coin > 0 & (coin - 1 = choc = 0 \checkmark choc > 0 & cnoc - 1 \leq coin - 1 \leq choc)) \Leftrightarrow VM1

where we allow ; to be right associative.

Thus VM1 specifies a machine that first accepts a coin, then dispenses a chocolate, after which it behaves again like VM1. Such a machine will alternately accept coins and dispense chocolates for as long as there is any call upon it to do so.

2.4 Recursion

Let P be the name of a predicate, and let F(P) be an expression denoting a predicate, and possibly containing occurrences of P. furthermore, let

$$\alpha P = \alpha F(P)$$

Thus the equation

may de regarded as a recursive definition of the predicate named by P. Problems of existence and uniqueness of this solution are postponed to section 4. 6.

Example

 $P \triangleq (coin; choc; P)$ $\alpha P \triangleq \{coin, choc\}$

We already know that a possible solution of this equation is:

 $P = VM1 = (choc \leq coin \leq choc + 1)$

In fact, this is the only solution, as will be shown in section 4.4.

2.5 Alternatives

Let P and Q be specifications, with

```
dP = dQ
```

We stipulate that $\alpha(P \lor Q) = \blacktriangleleft P$.

 $(P \lor 0)$ specifies a system which will behave like P or like Q (or like both). The choica between the alternatives is not determined; it may be made by the environment within which this system is embedded, as will be described in 2.6.

Example 1.

 $P \triangleq (choc; coin; P \lor coin; choc; P)$

This describes a vending machine which allows its customer to sample a chocolate, and trusts him to pay after. The solution to the defining equation is:

 $P = (choc \leq count+1 \leq choc + 2)$

Not surprisingly, the customer should pay for the privilege of using such a machine, as in the next example.

Example 2.

VM2 ≙ coin: P

←→ choc ≤ coin ≤ choc + 2

This machine allows its customer to insert up to two coins before extracting up to two chocolates.

Example 3.

CUST ♀ (choc; rejoice; CUST ∨ coin; choc; CUST ∨ kisswife; CUST) This describes the behaviour of a customer for the vending machine. He is, of course, very willing to take a chocolate without paying, after which he will greatly rejoice. He is also willing to kiss his wife, and whether he does so depends on whether she too is willing. He is also willing to pay for his chocolate in the normel way. The unique solution of the equation is

CUST = (choc ≤ coin + rejoice + 1 ≤ choc + 2 & rejoice ≤ choc)

Note that the specification of the customer does not prevent him from engaging in the following sequences of actions, since in each case, it correctly describes the values of the counts before and after each event of the sequence;

> coin, kisswife, choc, choc, rejoice, coin, coin, choc

The reason for these possibly unexpected sequences is that our specification language is too weak to describe such constraints as:

"he can't kiss his wife between insarting a coin and extracting a chocolate".

"he never inserts two coins in a row".

The weakness of the specification language is a daliberate decision that we do not wish to observe the exact relative timing of events; thus we allow events to occur "simultaneously", in the sense that any attempt to place them in order will laad to a non-determinate result. Some further consequences of this decision will become more approact later.

2,6 Concurrency

Let P and Q be specifications with disjoint alphabets, i.e.,

 $\alpha F n \alpha \hat{a} = \left\{ \right\}$

we then stipulate that

$$d(P \pm 3) = dP \cup dG$$
.

(P & 5) describes a system composed from the two subsystems separately described by F and Q, as they evolve together. Each event that occurs is in the alphabet of only one of the components; and its occurrence

requires participation of that component, leaving the other component unaffected. Each observation of the completa system splits uniquely into two parts; the values of variables in $\sim P$ are described by P and are not mentioned by Q; the value of variables in $\sim Q$ are described by Q and are not mentioned in P. Thus the full observation is exactly described by (P&Q).

Now let us relax the restriction of disjointness of alphabets. We stipulate that each event in the intersection of the two alphabets requires simultaneous participation of <u>both</u> subsystams deacribed by P and Q. Thus the count of events recorded by one subsystem must always be the same as that recorded for the other subsystem; so each possible value of this count must be described by both P and Q, and therefore by (PAQ).

Example 1.

Example 2.

VM1 🛆 PROFIT & BUFFER1

A simple vending machine can be constructed from two concurrent subsystems, one of which is profitable and the other refuses to allow insertion of a second conduction the first chocolate has been dispensed. Each event that occurs is an event in the life of both subsystems.

VM2 & CUST ⇔ cnoc ≤ coin ≤ choc + 2 & choc ≤ coin + rejoice + 1 ≤ choc + 2 & rejoice ≤ choc ⇔ rejoice = 0 & choc ≤ coin ≤ choc + 1 ∨ rejoice = 1 ≤ choc = coin

In this system, the customer can still kiss his wife at any time - that is of no concern to the vending machine. The customer never allows the humber of coins inserted to reach (choc + 2) and the vending machine never allows the number of chocs to exceed the number of coins inserted. It may therefore seem surprising that the customer should ever have cause to rejoice. This can happen when the events coin and choc occur simultaneously; the vending machine behaves as if the coin drooped first; but the customer thinks that choc was dispensed first and rejoices prematurely. If we wish to exclude such strange happenings, we shall have to strengthen our language for specifications, or use a more complex composition operator than & ; these topics will be pursued in later **sections**.

3. Example: the Dining Philosophers

In ancient times, a wealthy philanthropist endowed a College to accommodate five eminent philosophers. Each philosopher had a room in which he could engage in his professional activity of thinking; there was also a common dining room, furnished with a circular table, surrounded by five chairs, each labelled by the name of the philosopher who was to sit in it. The names of the philosophers were Phil₀, Phil₁, Phil₂, Phil₃, Phil₄ and they were disposed in this order anticlockwise round the table. To the left of each philosopher there was leid a goldan fork, and in the centre a large bowl of spaghetti, which was constantly replenished.

A philosopher was expected to spend most of his time thinking; but ` when he felt hungry, he went to the dining room, sat down in his dwn chair, picked up his dwn fork on his left, and plunged it into the spaghetti. But such is the tangled nature of spaghetti that a second fork is required to carry it to the mouth. The philosopher therefore had also to pick up the fork on his right. When he was finished he would put down both his forks, get up from his chair, and continue thinking. Of course, a fork can be used by only one philosopher at a time. If the other philosopher wants it, he just has to wait until the fork is available again.

3.1 Alphabets

We shall now construct a mathematical model of this system. First we must select the relevant sets of events. For Phil, the set is defined:

Note that the alphabets of the philosophers are mutually disjoint. There is no event in which they can agree to participate jointly, so there is no way whatsoever in which they can interact or communicate with each other ~ a realistic reflection of the behaviour of philosophers of those days. The other actors in our little drame are the five forks, each of which bears the same number as the philosopher who owns it. A fork is picked up and put down either by this philosopher, or by his neighbour on the other side. Its alphabet is defined

where \bigcirc denotes subtraction modulo 5.

Thus each event except sitting down end getting up requires participation of exactly two adjecent actors, a philosopher and a fork.

3.2 Behaviour

Apart from thinking and esting which we have chosen to ignore, the life of each philosopher is described:

The role of a fork is a simple one; it is repeatedly picked up and put down by one of its adjacent philosophers:

Fork_i \triangleq (i picks up fork i; i puts down fork i; Fork_i v (i \bigcirc) picks up fork i; (i \bigcirc) puts down fork i; Fork_i)

 \iff (i \ominus 1) picks up fork i = (i \ominus 1) puts down fork 1

& i picks up fork i > i puts down fork i ≥ ((i picks up fork i) - 1) V i picks up fork i = i puts down fork i & i ⊖ 1 picks up fork i ≥ (i⊖1) puts down fork i > (((i⊖1) outs down fork i) - 1)

The Deraviour of the entire college of philosophers is simoly the conjunction of the behaviour of these components:

College $\triangleq \begin{pmatrix} 4 \\ 4 \\ 1 \\ i=0 \end{pmatrix} \triangleq \begin{pmatrix} 4 \\ 4 \\ i=0 \end{pmatrix} \triangleq \begin{pmatrix} 4 \\ 4 \\ i=0 \end{pmatrix}$

10.

3.3 Deadlock:

When this mathematical model had been constructed, it revealed a serious danger. Suppose all the philosophers get hungry at about the same time; they all sit down; they all pick up their own forks; and they all reach out for the other fork - which isn't there. In this undignified situation, they will all assuredly starve. Although each actor is capable of further action, there is no action which any pair of them can agree to do next.

A possible observation of this and outcome is

```
4
& i sitsdown ≠ i picks up fork i = 137
i≂0
```

 $\begin{array}{ccc} \overset{4}{\underline{&}} & i \text{ puts down fork } i = i \text{ puts down fork } (i \bigoplus 1) \\ \overset{1=0}{\underline{&}} & = i \text{ picks up fork } (i \bigoplus 1) = i \text{ gets up } = 136 \end{array}$

This observation is described by the specification of the College; but there is no way of adding unity to just one of the counts so that this successor observation is also described by the specification. If we are not willing to allow the simultaneous occurrence of several events in the life of a single actor, nothing further can happen.

3.4 Deadlock averted

However, our abory does not end so saoly. Once the danger was detected, there were suggested many ways to event it. For example, one of the philosophers could always pick up the wrong fork first - if only they could have agreed which one it should be! The purchase of a single additional fork was rulad out for similar reasone, whereas the purchase of five more forks was much too expensive.

The solution finally adopted was the appointment of a Footman, whose duty it was to assist aach philosopher into and out of his chair. His alphabet was defined:

This footman was given secret instructions never to allow more than four philosophers to be simultaneously seated;

Footman $\Delta \sum_{i=0}^{4} (i \text{ sits down - i gets up}) \leq 4$

Let F $_j$ (for 0 \leqslant j \leqslant 4) denote the behaviour of the footman with j philosophers seated. Then

$$F_{4} \stackrel{4}{\cong} \bigvee_{i=0}^{4} (1 \text{ gets up; } F_{3})$$
(for j = 1,2,3) $F_{j} \stackrel{4}{\cong} \bigvee_{i=0}^{4} (i \text{ gets up; } F_{j-1} \lor 1 \text{ sits down; } F_{j+1})$

$$F_{0} \stackrel{4}{\cong} \bigvee_{i=0}^{4} (1 \text{ sits down; } F_{1})$$

We need to prove that

$$F_j \implies \sum_{i}^{4} (i \text{ sits down - i gets up}) 4 4-j$$

and hance

This establishes that F $_{\rm O}$ is a correct implementation of the behaviour specified for the Footman.

The edifying tale of the dining philosophers is due to Edsger W. Dijkstra; the Footman is due to Carel Scholten.

Mathematical Proparties

A major advantage of specifying complex systems in terms of some familiar domain like the predicate calculus is that the operators enjoy a number of alegant mathematical properties. Algebraic properties are those that can be expressed as simple squations, or (in the case of predicates) as equivalances. Ordering properties are those which are expressed as inequalities, or (in the case of predicates) as implicatione.

4.1 Algebraic properties

(1) "&" and "v" are associative, commutative, and idempotent.

(2) "&", and "v" and "c;" distribute through v.

However, because the operands of "v" must have the same alphabet, it is not generally true that "v" distributes through "&". Nevertheless any predicate which uses only "&" and "v" (and constants ZERO and CHAOS) can be reduced to disjunctive normal form, in which & is the innermost operator. The "&" operator may then be eliminated, using:

(3)
$$ZERO_A & ZERO_B = ZERO_A \cup B$$

 $CHAOS_A & CHAOS_B = CHAOS_A \cup B$
 $ZERO_A & CHAOS_B = ZERO_A \cup B$

We would now like to develop a normal form for predicates containing the prefixing operator "c;". Since prefixing distributes through "v", it is necessary only to show how it distributes through "&".

$$(4) (a;P) & CHAOS_{A} = a; (P & CHAOS_{A})$$

$$(a;P) & ZERO_{A} \cong a; (P & ZERO_{A}) \text{ if } a \in A$$

$$= ZERO_{A} \text{ otherwise}$$

$$(5) (a;P) & (b;Q) \iff b; ((a;P) & Q) \text{ if } \begin{cases} a \in dQ \\ b \in dP \end{cases}$$

$$\iff a; (P & (b;Q)) \text{ if } \begin{cases} a \in dQ \\ b \in dP \end{cases}$$

$$\iff a; (P & (b;Q)) \text{ if } \begin{cases} a \in dQ \\ b \in dP \end{cases}$$

$$\iff a; (P & (b;Q)) \text{ if } \begin{cases} a \in dQ \\ b \in dP \end{cases}$$

$$\iff a; (P & (b;Q)) \text{ if } \begin{cases} a \in dQ \\ b \in dP \end{cases}$$

$$\iff a; (P & (b;Q)) \text{ if } \begin{cases} a \in dQ \\ b \in dP \end{cases}$$

$$\implies a; (P & (b;Q)) \text{ if } \begin{cases} a \in dQ \\ b \in dP \end{cases}$$

$$\implies a; (P & Q) \text{ if } a \in (dP \land Q)$$

$$\implies a; (P & Q) \text{ if } a \in (dP \land Q)$$

This theorem permits prefixing to be moved outside "&" in all cases except for (a;P) & (b;Q) when $a,b \in dP \land dQ$ and $(a \neq b)$. In this case each operand is attempting an action which requires simultaneous carticipation of the other, but they disagree on which action it shall ba. As a result, one might expect that nothing can happen, and the conjunction should reduce to ZERO. But this fails to take into account the possibility of simultaneous occurrence of events; for example:

-Q)

(a;b;ZERO) & (b;a;ZERO) $\iff (a = b = 0 \quad \forall \ a = b = 1)$

If we wish to eliminate this possibility, we must introduce the restriction that the alphabets of the two operands of "&" may have at most one evant in common.

Under this restriction, every expression can be reduced to a normal form in which "v" is the outermost operator, and the "&" operator has been eliminated.

4.2 Ordering properties

The set of all predicates is partially ordered by the relation of logical implication, which we write " \implies "; this is shown by the follow-ing familiar metatheoreme:

 $P \Longrightarrow P$ if $P \Longrightarrow Q$ and $Q \Longrightarrow P$ then $P \Longleftrightarrow Q$ if $P \Longrightarrow Q$ and $Q \Longrightarrow R$ then $P \Longrightarrow R$.

A function F is said to be <u>monotonic</u> if it "respects" the ordering of its operand, i.e.

The definition extends to functions with many operands, if they are monotonic in each of their operands separately, e.g.,

If
$$P \implies Q$$
 then $F(P,R) \implies F(Q,R)$
and $F(R,P) \implies F(R,Q)$ for all P,Q,R .

Any operator that distributes through "v" is monotonic, so all the operands we have defined so fer enjoy this property. Furthermore any composition of monotonic operators is also monotonic, so any function or expression defined in terms of these operators will be monotonic in all its operands.

There is a good reason why operators used in recursive definitions should be monotonic - it ensures the <u>existence</u> of a predicate satisfying each recursive definition. This assurance is given by the Tarski-Knaster theorem, provided we accept the assumption that the space of predicates is <u>complete</u>. A partial ordering is complete if <u>every</u> set 5 of predicates has an infinite conjunction \forall S, such that (for any Q):

if and only if $Q \longrightarrow \forall S$.

The <u>limit points</u> \forall S are related to ordinary pradicates in much the same way as the reals are to the rationals. They may be uncountable in number, but that should not be taken as a reason for denying their existence.

If there is more than one solution for a recursion equation, it is necessary to decide which one is meant. The usual tachnique have is to take the <u>weekest</u> solution, i.e. the disjunction of all possible solutions. The fact that this is itself a solution is also assured by the Tarski-Knaster theorem.

4.3 Continuity

Let $S = \left\{ S_n \mid n \ge 0 \right\}$ be a countable set of predicates such that $S_{n+1} \longrightarrow S_n$ for all n.

Such an S is known as a <u>chain</u>, and its limit VS is written

∀n ≽ 0. S

If f is a monotonic function of predicates, and S is a chain, then

 $\left\{ \mathbb{P}(S_n) \mid n \ge 0 \right\}$

is also a chain, and has as limit

$$\forall n, F(S_n)$$
.

Because F is monotonic, we have the implication:

 $F(\forall n, S_n) \longrightarrow \forall n, F(S_n)$ for all chains S.

If this implication can be strengthened to equivalence, then F is said to be <u>continuous</u>.

A function of several pradicates is continuous if it is continuous in each argument separately.

It is easy to see that the operators "c;", "&" and "v" are continuous. It follows that every function of predicates defined in terms of these operators is also continuous. The main advantages of continuity will emerge later; but here we note that continuity of a function F greatly simplifies the search for a solution to a recursion equation of the form:

Define $F^{0}(true) = true$

$$F^{n+1}(true) = F(F^{n}(true))$$
$$= \underbrace{F(F(\dots, F(true)))}_{n+1 \text{ times.}}$$

Now the weakest solution of the equation is

Proof. (1) that it is a fixed point:

$$F(\forall n F^{n}(true)) = \forall n(F^{n+1}(true)) \text{ by continuity}$$

$$= true \& \forall n.F^{n}(true)$$

$$= \forall n.F^{n}(true)$$
(2) that it is the weakest fixed point:
Let P = F(P)
... P = F^{n}(P) by induction
... P = \forall n F^{n}(P) by monotonicity

Examples

(1)
$$P \triangleq I(P)$$
 where I is the identity function.
 $I^{n}(true) \iff true$
 $\therefore P \iff \forall n, I^{n}(true) \iff true$

(2)
$$P \mathrel{\triangleq} (P \lor Q)$$

 $F^{n}(true) \iff true \lor Q \lor Q \lor \dots \lor Q$
 $\iff true$
 $\therefore P \iff true$

(3)
$$p \Leftrightarrow p \& q$$

 $f^{n}(true) \longleftrightarrow q$
 $\therefore P = q$
(4) $P = a; P$ where $dP = \{a, b\}$
 $F^{1}(true) = (a = b = 0 \lor a > 0)$
 $F^{2}(true) = (a = b = 0 \lor a = 1 \& b = 0 \lor a > 1)$
 $F^{n}(true) = (b = 0 \lor a \ge n)$
 $\therefore p \longleftrightarrow \forall n.(b = 0 \lor a \ge n)$
 $\longleftrightarrow (b = 0)$

As shown by these examples, the explicit solution of each equation requires an induction to recast $F^{n}(true)$ as a predicate explicitly containing n as a free variable.

4.4 Unique solutions

In all our examples of sections two and three, the solutions of the recursive equations have been unique. It is useful to recognise cases of unique solution, because they are simpler to reason about. Consider the equation

$$P = F(P)$$
.

F(P) is said to be <u>guarded</u> if it can be written in the form

 $a;G(P) \vee b;H(P) \vee \dots$ (1)

where G(P), H(P) are expressed using only the notations "c;" and "v". If F(P) is guarded, then the equation above has an unique solution.

Proof. An expression is said to be guarded to depth N+1 if each of G(P), H(P), in formula (1) is guarded to depth n. Using distributivity of "c;" such formula can be written

a;b; ...; G'(P) ∨ c;d; ...; H'(P) ∨ ...

where the length of each of the prefix sequences "c;d;...;" is greater than n.

If F(P) is guarded, then $F^{n}(P)$ is guarded to depth n. For any observation, we define its <u>length</u> as the sum of all the values ascribed to its variables; thus the length is a count of the total number of events that have occurred, and is necessarily finite. If an observation is of length n, and a predicate $F^{n}(P)$ is guarded to depth n, then one can determine whether the predicate describes the observation by merely looking at the prefix sequences, independent of the value of P.

Now let

P = F(P) and Q = F(Q) $\therefore P = F^{n}(P)$ and $Q = F^{n}(Q)$ by induction.

Consider an observation described by P. Let its length be n. This observation is also described by $F^{n}(P)$ and therefore by $F^{n}(Q)$, which has all the same prefixes of length n. Thus all observations described by P are also described by Q, and vice versa. P and Q are therefore equivalent as predicates.

5. Programming

As discussed in the introduction, a specification is an arbitrary predicate describing all possible observations of some system. The task of the scientist is to <u>discover</u> the atrongest specification describing some aspect of the behaviour of the natural universe. The task of the engineer is different: he has to <u>construct</u> some mechanism which meets a specification describing the needs of a potential user of the mechanism. The engineer's duties would be much simplified if he had at his disposal a stock of <u>universal</u> mechanisms. An universal mechanism is one that will first accept the text of any oesired specification, and will then automatically transform itself into a mechanism which behaves in accordance with that specification, for as long as it is wanted. Such a marvellous mechanism might be a bit expensive; and if so, it should have a switch to turn it back into a universal mechanism when there is no need for it to continue to satisfy its current specification.

For the civil engineer or naval architect, an universal mechanism of this kind is nothing but a pipe-dream. But for the computar programmer, it is a reality which he takes for granted - it is the stored program digital computer. The only problem is that the computer will not accept

18.

an arbitrary predicate as a specification; the predicate must be written in a highly restricted notation known as a programming language. Such predicates are known as <u>programs</u>. In the remainder of this paper we shall define a program as a pradicate expressed solely in terms of the constants and operators introduced in section 2 together with recursion. Later, we shall consider some further restrictions on this notation.

5.1 Programming methodology

The task of the programmer is now clear. First, he must use his best judgement to formulate a specification of the desired product. The specification should be expressed as a predicate, taking advantage of the full power of the concepts and notations of logic and mathematics to keep the formalisation simple and clear. Clarity is of the utmost concern, since a misunderstanding at this stage can have a severe impact on the quality of the product. The programmer now has to reformulate the specification as a program P, expressed in the restricted notations of his programming language; and this may involve a gross expansion in the size of the text. Furthermore, he must find a program that is adequately afficient when executed on a mechanism of affordable capacity and speed. His task may be slightly simplified by the fact that it is sufficient to find a program P that merely implies its specification 5; there is no need to achieve exact equivalence.

Thus the task of programming is rather like that of finding an explicit definition of a function which satisfies given differential equations. Just as some equations have no explicit solutions, some oredicates cannot be programmed because they are incomputable or even inconsistent. For soluble equations, mathematicians have discovered many techniques for finding and checking proposed solutions, though their application usually demands some mathematical skill and insight. A collection of methods for constructing a program to meet an arbitrary specification is known as a programming methodology.

5.2 Top-down development

One rather obvious programming method is the technique of "topdown development" or "divide-and-conduer"; it is rather similar to integration by parts, in that its injudicious use can require solution of subproblems more difficult than the original problem. Let S be the specification of the desired program. Then

- Let F(T,U) be an expression containing the predicates T and U, but otherwise expressed wholly within the programming language.
- (2) Prove that $F(T,U) \longrightarrow S$.
- (3) Find programs P and Q such that

$$P \longrightarrow T$$

and $Q \longrightarrow U$,

(4) The result you want is

F(P.Q).

The validity of this method depends on the fact that all operators of the programming language are monotonic. Its utility derives from the fact that F is proved correct before P and Q are programmed.

5.3 Introduction of recursion

One of the most significant tasks of the orogrammer is to construct correct loops or recursions. Suppose 5 is the specification which is believed to require a recursively defined program. Then the following steps are recommended:

(1) Find an expression E which maps the observations of the alphabet onto non-negative integers. This is known as a "variant function".
(2) Find a program f(P), containing the predicate name P, such that

$$F(S \lor E \geqslant n) \Longrightarrow (S \lor E > n)$$

where n is a fresh variable.

Thus if the recursive call P establishes S in all circumstances when $\mathcal{E} \leq n$, then F(P) is <u>better</u> than P, in that it ensures S in the case when $\mathcal{E} = n+1$ as well.

(3) Then the program you want is defined recursively:

□ ▲ F(P)

20.

Proof. From step 2, by an easy induction

$$F^{n}(\underline{true}) \iff F^{n}(S \lor E \ge 0) \implies (S \lor E \ge n) \text{ for all } n$$

$$\therefore \quad (\forall n.F^{n}(\underline{true})) \implies (S \lor \forall n.E > n)$$

$$\iff S$$

Since F is expressed wholly in notations which are known to be continuous, $\forall n.F^n$ (true) is the weakest solution of

P 4 F(P)

6. Implementation

The main motivation for writing specifications in a restrictive programming language is the existence of a universal mechanism that will automatically implement the program. In practice, such an implementation is constructed from silicon chips, boards, wires, stc., together with loaders, compilers, or interpreters for the given programming language. But for our present purposes, it is more convenient to construct an abstract mathematical machine, passing through a series of states; each state corresponds to a possible observation described by the program initially given to the machine. Here is an informal description of the behaviour of such a machine, intended to model the actions of an interacting system:

(0) Input the specification with given alphabet.

 Declare an integer variable corresponding to each event in the alphacet. Initialise all these variables to zero.

(2) Repeat the following steps as long as possible (or until switched off):

(2.1) For each variable in turn, sdd one to its value, test if the specification is true, and subtract one again.

(2.2) For all variables which have passed test (2.1), wait until the user/environment of the mechanism has selected which of the events is to occur. If no variable has passed the test, this wait will last forever.

(2.3) Add one to the variable selected in step 2.2.

We assume that an observation of this machine can be made just before each iteration of the loop (2); at that time, the current values of all its variables are observable.

6.1 Consistency

Clearly, this implementation does not correctly implement every predicate; in fact the predicate "a > 3" is false immediately after step (1). However, any predicate which is true after step (1) will remain true forever - that is the purpose of the tests of step (2.1). So, to show that the implementation correctly implements every <u>program</u>, it suffices to show that every program describes the observation when all the counts are zero, i.e., that

 $ZERO_A \longrightarrow P$ for all programs P with alphabet A. This can be proved by structural induction on P, using the following lemmas.

- (1) $ZERO_1 \implies ZERO_1$
- (2) ZERD_a \implies CHAOS_a

- (4) If $ZERO_A \longrightarrow P$ and $ZERO_A \longrightarrow Q$ then $ZERO_A \longrightarrow P \checkmark Q$
- (5) If $ZERO_A \longrightarrow P$ and $ZERO_B \longrightarrow O$ then $ZERO_A \longrightarrow B \implies P \land O$
- (6) If for all n ZERO, $\Longrightarrow P_{n}$

The last clause is required to show that all programs defined by recursion are implied by ZERG; this conclusion depends also on continuity of all the connectives of the programming language, a property which has already been established in 4.3.

Thus we have shown that the proposed implementation works correctly for all programs expressed the language. This result may be reformulated: we have shown that the programming language is <u>consistent</u> with the given model implementation, in the sense that every poservation produced by the implementation will be correctly described by its program.

6.2 Completeness

Suppose that a set I of one or more correct implementations has been proposed for a given programming language. Let P be a program submitted to these implementations. Now it may be that there is an observation described by P which can never in fact occur on any of the implementations. Thus the predicate P is not the <u>strongest</u> possible predicate describing the behaviour of its proposed implementations. In this case, we say the programming language is <u>incomplete</u> with respect to I, since it is not capable of expressing every true fact about I. Incompleteness is a serious fault, because it means that some correct programs cannot be proved correct solely in terms of their definition as predicates; their proof would require operational reasoning based on the implementations.

Unfortunately, our programming language is not complete with respect to the implementation described above. In that implementation, each observation except the first (all zeroes) has a <u>predecessor</u> which can be derived from it by subtracting unity from just one of its component variables. Consider the specification

coin = choc.

This is satisfied by the observation

coin = chac = 5

but not by the observation

```
coin = 4 \& choc = 5
```

or by the only other predecessor abservation

coin = 5 & choc = 4.

Thus the specification describes an observation which can never be reached by the intanded implementation.

This would not matter if such a specification could never be expressed as a program. Unfortunately it can. Consider a silly customer of the vending machine

One solution to the problem would be to expand the set of implementations to allow any number of events to occur simultaneously. Another solution is to restrict the programming language still further, so that no specification expressible as a program will describe an unreachable observation.

Let us attempt a solution of the second kind. A predicate is said to be <u>grounded</u> if every described observation has a predecessor also described by it. Each observation therefore has a chain of predecessors reaching back to the initial zero observation. The intended implementation can follow this chain in the reverse direction, and can thus reach any observation described by a grounded predicate. So we need to restrict our programming language in such a way that it can exprass only grounded predicates.

An effective restriction is the one introduced in section 4.1 to obtain a normal form: forbid the use of the operator "i" except when the alphabets of its operands contain at most one event in common. The fact that all programs satisfying this restriction are grounoed is proved by structural induction, based on the following lemmas:

- (1) ZERO, and CHAOS, are grounded.
- (2) If P and Q are grounded, so are (a;P) and $(P \lor Q)$.
- (3) Furthermore, if size $(\exists P \land \exists Q) \leq 1$ then (P & Q) is grounded.
- (4) If for all n P_n is grounded and $P_{n+1} \longrightarrow P_n$ then $\forall n.P_n$ is grounded.

Only the last clause requires any subtlaty of proof. Consider an observation described by $\forall n.P_n$. The maximum possible number of its predecessors is finite (equal to the size of the alphabet). Since P_{n+1} logically implies P_n , the set of predecessors (of the given observation) described by P_{n+1} is a subset of those described by P_n . These finite subsets form a descenoing chain, whose intersection is therefore non empty. Furthermore this intersection is described by $\forall n.P_n$.

6.3 The excluded miracle

In section 3.1 we defined the task of the programmer as the distionery of a program ρ which logically implies a given specification 5.

24.

Now suppose he could find a program P that logically implies every other program Q. This P would be a miraculous program, since it could be used to implement <u>every</u> implamentable specification. With such a program, he would never need any other; and each of his tasks would be trivial. An example of such a miraculous program would be one that expresses the predicate <u>false</u>.

In practice, we suspect that a programming language in which there exists a mireculous program would be unrealistic or useless in some other way; it would certainly be unworthy of serious mathematical study. Unfortunately, our little programming language (as proudly proved in section 5.1) contains the miracle ZERD_A. So we need to exclude this constant from the lenguage; but we also need to ensure that it can never be expressed in some other way, for example:

(choc;SILLY) & (coin;VM1)

(coin = choc = 0 v choc > 0 & coin < choc-1 < coin+1)
& (coin = choc = 0 v coin > 0 & choc < coin+1 < choc+1)
(coin = choc = 0)</pre>

This problem can be solved by imposing the same restriction on the alphabets of the operands of & as has already been recommended in 4.1 and 6.2. Then no program is equal to the miracle ZERO_A. A proof of this week structural induction based on the lemmas.

- (1) CHAOS \neq ZERO
- (2) If P and Q are programs distinct from ZERO_A then so are (a;P) and (P \vee Q), and (P & Q) provided size (\checkmark P \land Q) \leq 1.
- (3) If for all $n \ P_n \neq ZERO_A$ and $P_{n+1} \Longrightarrow P_n$ then $(\forall n, P_n) \neq ZERO_A$.

The proof of the last clause uses the fact that an observation has only a finite number of possible successors.

Now we need to prove that there are no other miracles besides $ZERO_{\mu}$, which we have already excluded. Unfortunately there is. If the alphabet A contains only one event, then all expressible programs are equivalent to CHAOS_A. We must therefore insist that every alphabet A contains at least two variables, say "a" and "b". Then there ere at least two oifferent programs with elphabet A:

$$P \stackrel{\bullet}{\bullet} a; P \iff b = 0$$

$$Q \stackrel{\bullet}{\to} a = 0$$

Furthermore the only observation they both describe is

 $(a = b = 0) \iff ZERO_{A}$

But since we have shown that no program implies $ZERD_A$, no program can imply both P and Q. Thus there is no program which implies every other program.

6.4 Computability

The universal mechanism described above requires the ability to determine the truth or falsity of a predicate for an arbitrary value of its free variables. It is well known that for some predicates this is impossible. We must therefore ensure that our programming language can express only predicates which are in some sense computable.

In general, the appropriate sense of computability seems to be that the <u>complement</u> of the predicate should be recursive enumerable. Thus if an observation <u>falsifies</u> the predicate, it will be possible to <u>prove</u> that it does so. Such predicates are said to be <u>falsifiable</u>. According to Karl Popper, falsifiability is also a required property of all scientific theories. In accordance with this definition all our programs are falsifiable:

- (1) ZERO, and CHAOS, are falsifiable.
- (2) If P and J are falsifiable, then so are (a;P), $(P \lor Q)$ and (P & Q).
- (3) If for all n, P is falsifiable, then so is $\forall n.P_{\perp}$.

Another vary desirable property of a general purpose orogramming language is that its programs should be able to opmpute every computable function. The usual method of proving this is to program in the language a simulation of some known universal machine such as a Turing machine. But our language cannot do this; in fact the language can be implemented on a finite state machine. In a social purpose language, such a limitation may be an advantage, if it permits a mechanical check against certain undesirable occurrences such as non-terminetion or deadlock. A more practical defect of our language is that it cannot even describe the simple examplas used to illustrate this paper. To solve these defects we would need to introduce more operators and perhaps variables denoting different kinds of observation.

6.5 Continuity again

The description of an universal implementation of our language essentially regards each complete program as a predicate which can be tested. In practice, programs are implemented in a structured feshion, and the complete implementation of a complex program is constructed from implementations of its parts, for example, a test of $(P \lor 0)$ or $(P \And 0)$ or $(P \And 0)$ or (a; P) can be constructed from tests of P and Q separately. However, the position with recursively defined predicates is less clear.

Following Scott, we identify an implementation of the predicate

```
P 全 F(P)
```

as the infinite sequence of "finite approximations":

$$\left\{ F^{n}(true) \mid n \geqslant 0 \right\}$$

such that the desired program P is the universal quantification of the sequence. Now if G is any <u>continuous</u> function of predicetes to predicates, an implementation of G(P) can be constructed from the implementation of P, thus

 $\left\{ G(F^{n}(true)) \middle| n \geqslant Q \right\}.$

Because G is continuous, this is an infinite sequence of finite approximations to G(P). Thus we have constructed an "implementation" of G(P)out of an implementation of P.

There remains one outstanding question. A programming lenguage defined by the methods described in the previous sections is a notation for specifying <u>falsifiable</u> predicates. However, the set of observations generated by any mechanism must be recursive enumerable. Since the complement of a recursive enumerable set is not necessarily recursive enumerable, there will in general be abservations which cannot be generated by a correct implementation, even though they are described by the program which is being implemented. Thus no general-purpose programming language can be complete in the sense of 6.2. The problem is most ecute in the case of a non-terminating recursion

According to our definition

 $P \longleftrightarrow CHADS_{a}$.

However any "reasonable" implementation is likely to produce only the all-zero observation, so that it will behave as if

p 🔶 ZERO, .

Thus <u>all</u> observations (except one) fall into the gap between what is described by the program and what is actually gamerated by its implementation.

I do not know whether this problem has a solution, or even whather it needs one.

7. Conclusion

The first step in the construction of a mechanism to meet some requirement is to motain a very good understanding of that requirement. This understanding can be formalised as a predicate describing all possible acceptable observations of the behaviour of the mechanism. Because misunderstanding at this stage is so dangsrous, a specification should be short, simple, and well-structured; and to this end it should take full advantage of all available concepts and notations of mathematics and logic. For effective description of certain kinds of dynamic system, it is convenient also to use certain concepts akin to those of computer programs; and this requires they be defined in terms of conventional predicates. Even so, the construction of a successful specification requires the same human skills and insights as are characteristic of an applied mathematician of engineer.

The task of the programmer remains to find some predicate, expressed only in the restricted notations of his programming language, which logically implies the specification. The reason for the restrictions is to enable the program to be run on some available implementation of the language.

This view of programs and their specifications is highly relevant for the designer of a programming language. Firstly, he must have a wide understanding of the application area of his language. Based on this, he should define a range of concepts to assist in the specification of programs within that area. These concepts should enjoy nice algebraic

28.

properties, preferably admitting a simple normal form. If recursion or repatition is required, then all the propositional connectives involved should be at least monotonic.

The designer should then define a clase of model implementations, which will automatically conform to a subclass of specifications submitted to them. The implementations should be a reasonable abstraction of what can be built, perhaps in silicon, at acceptable cost. This will require some restriction on the generality of the notations used in the specification. These restrictions should not be so severe that they prohibit an exact description of the behaviour of the model implementations; but they should be severe enough to exclude the uniformity false predicate, or any program which logically implies all other programs. Finally, there should be some reasonably easy stepwise method for designing a program from its specification. All these tasks are considerably simplified if all the operators of the programming languege are not only monotonic but continuous.

This account of the nature of programming and of programming language design suggests that both activities require and deserve the techniques and skills of the mathematical logician.

Acknowledgements

The original iosa of equating a computer program with its strongest specification is due to $\hat{\epsilon}.\hat{\epsilon}.\hat{\epsilon}.$ Hehmer. The treatment of Communicating Systems was inspired by the work of A.J.R.G. Milner. The technique of predicate transformers is due to $\hat{\epsilon}.w$. Dijkstra. The development of specifications as predicates together with their alphabets is due to $\hat{J}.-\hat{R}.$ Abrial. The treatment of recursion by fixed points and continuity is one to D.S. Scott. Valuable and continuing mathematical support has been provided by $\hat{\epsilon}.-\hat{R}.$ Olderog.

29,

OXFORD UNIVERSITY COMPUTING LABORATORY PROGRAMMING RESEARCH GROUP TECHNICAL MONOGRAPHS

JULY 1982

This is a series of technical monographs on topics in the field of computation. Copies may be obtained from the Programming Research Group. (Technicat Monographs), 45 Banbury Road, Oxford, OX2 6PE, England.

- PRG-2 Dana Scott Outline of a Mathematical Theory of Computation
- PRG-3 Dana Scott The Lattice of Flow Diagrams
- PRG-5 Dana Scott Data Types as Lattices
- PRG-6 Dana Scott and Christopher Strachey Toward a Mathematical Semantics for Computer Langueges
- PRG-7 Dana Scott Continuous Lattices
- PRG-8 Joseph Stoy and Christopher Strachey OS6 - an Experimental Operating System for a Small Computer
- PRG-9 Christopher Strachey and Joseph Stoy The Text of OSPub
- PRG-10 Christopher Stracney The Varieties of Programming Language
- PRG-11 Christopher Strachey and Christopher P. Wadsworth Continuations: A Mathematical Semantics for Handling Full Jumps
- PRG-12 Peter Mosses The Mathematical Semantics of Aigol 60
- PRG-13 Robert Milne The Formal Semantics of Computer Languages and their Implementations
- PRG-14 Shan S Kuo, Michael H, Linck and Sohrab Saadat A Guide to Communicating Sequential Processes
- PRG-15 Joseph Stoy The Congruence of Two Programming Language Definitions
- PRG-15 C. A. R. Hoare, S. D. Brookes and A. W. Roscoe A Theory of Communiceting Sequential Processes
- PRG-17 Andrew P. Black Asport on the Programming Notation 3R

- PRG-18 Elizabeth Fielding The Specification of Abstract Mappings and their Implementation as B⁺-trees
- PRG-19 Dana Scott Lectures on a Mathematical Theory of Computation
- PRG-20 Zhou Chao Chen and C. A. R. Hoare Partial Correctness of Communicating Processes and Protocols
- PRG-21 Bernard Sufrin Formal Specification of a Display Editor
- PRG-22 C. A. R. Hoare A Model for Communicating Sequential Processes
- PRG-23 C. A. R. Hoare A Calculus of Total Correctness for Communicating Processes
- PRG-24 Bernard Sulrin Reading Formal Specifications
- PRG-25 C. B. Jones Development Methods for Computer Programs including a Notion of Interference
- PRG-26 Zhou Chao Chen The Consistency of the Calculus of Total Correctness for Communicating Processes
- PRG-27 C. A. R. Hoare Frogramming is an Engineering Prolession
- PRG-28 John Hughes Graph Reduction with Super-Combinators
- PRG-29 C. A. R. Hoare Specifications, Programs and Implementations
- PRG-30 Alejandro Teruel Case Studies in Specification. Four Games