Numerische Mathematik Bd. 1, S. 41—60 (1959)

Report on the Algorithmic Language ALGOL

by

the ACM Committee on Programming Languages and the GAMM Committee on Programming

edited by

A. J. PERLIS and K. SAMELSON

Editors' Note. In the interest of immediate circulation of the results of the ACM-GAMM committee work on an algebraic programming language, this preliminary report is presented. The language described naturally enough represents a compromise — but one based more on differences of taste than on content or fundamental ideas. Even so, it provides a natural and simple medium for the expression of a large class of algorithms. This report has not been thoroughly examined for errors and inconsistencies. It is anticipated that the committee will prepare a more complete description of the language for later publication.

Part 1. Introduction

In 1955, as a result of the Darmstadt meeting on electronic computers, the GAMM (Gesellschaft für angewandte Mathemathik und Mechanik), Germany, set up a committee on programming (Programmierungsausschuß). Later a subcommittee began to work on formula translation and on the construction of a translator, and a considerable amount of work was done in this direction.

A conference attended by representatives of the USE, SHARE, and DUO organizations and the ACM (Association for Computing Machinery) was held in Los Angeles on May 9 and 10, 1957 for the purpose of examining ways and means for facilitating exchange of all types of computing information. Among other things, these conferees felt that a single universal computer language would be very desirable. Indeed, the successful exchange of programs within various organizations such as USE and SHARE had proved to be very valuable to computer installations. They accordingly recommended that the ACM appoint a committee to study and recommend action toward a universal programming language.

By October 1957 the GAMM group, aware of the existence of many programming languages, concluded that rather than present still another formula language, an effort should be made toward unification. Consequently, on October 19, 1957, a letter was written to Prof. JOHN W. CARR III, president of the ACM. The letter suggested that a joint conference of representatives of the GAMM and ACM be held in order to fix upon a common formula language in the form of a recommendation.

An ACM Ad-Hoc committee was then established by Dr. CARR, which represented computer users, computer manufacturers, and universities. This committee held three meetings starting on January 24, 1958 and discussed many technical details of programming language. The language that evolved from these meetings was oriented more towards problem language than towards computer language and was based on several existing programming systems. On April 18, 1958 the committee appointed a sub-committee to prepare a report giving the technical specifications of a proposed language.

A comparison of the ACM committee proposal with a similar proposal prepared by the GAMM group (presented at the above-mentioned ACM-Ad-Hoc committee meeting of April 18, 1958) indicated many common features. Indeed, the GAMM group had planned on its own initiative to use English words wherever needed. The GAMM proposal represented a great deal of work in its planning and the proposed language was expected to find wide acceptance. On the other hand the ACM proposal was based on experience with several successful, working problem oriented languages.

Both the GAMM and ACM committees felt that because of the similarities of their proposals there was an excellent opportunity for arriving at a unified language. They felt that a joint working session would be very profitable and accordingly arranged for a conference in Switzerland to be attended by four members from the GAMM group and four members from the ACM committee. The meeting was held in Zurich, Switzerland, from May 27 to June 2, 1958 and attended by F. L. BAUER, H. BOTTENBRUCH, H. RUTISHAUSER and K. SAMELSON from the GAMM committee and by J. BACKUS, C. KATZ, A. J. PERLIS, and J. H. WEGSTEIN for the ACM Committee*.

It was agreed that the contents of the two proposals should form the agenda of the meeting, and the following objectives were agreed upon:

I. The new language should be as close as possible to standard mathematical notation and be readable with little further explanation.

II. It should be possible to use it for the description of computing processes in publications.

III. The new language should be mechanically translatable into machine programs.

There are certain differences between the language used in publications and a language directly usable by a computer. Indeed, there are many differences between the sets of characters usable by various computers. Therefore, it was decided to focus attention on three different levels of language, namely a *Reference Language*, a *Publication Language* and several *Hardware Representations*.

Reference Language

1. It is the working language of the committee.
2. It is the defining language.
3. It has only one unique set of characters.

* In addition to the members of the conference, the following people participated in the preliminary work of these committees: GAMM: P. GRAEFF, P. LÄUCHLI, M. PAUL, F. PENZLIN. — ACM: D. ARDEN, J. MCCARTHY, R. RICH, R. GOODMAN, W. TURANSKI, S. ROSEN, P. DESILETS, S. GORN, H. HUSKEY, A. ORDEN, D. C. EVANS.

4. The characters are determined by ease of mutual understanding and not by any computer limitations, coders notation, or pure mathematical notation.

5. It is the basic reference and guide for compiler builders.

6. It is the guide for all hardware representations.

7. *It will not normally be used stating problems.*

8. It is the guide for transliterating from publication language to any locally appropriate hardware representations.

9. The main publications of the common language itself will use the reference representation.

Publication Language (see Part III c)

1. The description of this language is in the form of permissible variations of the reference language (e.g., subscripts, spaces, exponents, Greek letters) according to usage of printing and handwriting.

2. It is used for stating and communicating problems.

3. The characters to be used may be different in different countries but univocal correspondence with reference representation must be secured.

Hardware Representations

1. Each one of these is a condensation of the reference language enforced by the limited number of characters on standard input equipment.

2. Each one of these uses the character set of a particular computer and is the language accepted by a translater for that computer.

3. Each one of these must be accompanied by a special set of rules for transliterating from Publication language.

Acknowledgements. The members of the conference wish to express their appreciation to the Association for Computing Machinery, the "Deutsche Forschungsgemeinschaft", and to the "Eidgenössische Technische Hochschule Zürich", for substantial help in making this conference and resultant report possible.

Part II. Description of the reference language

1. Structure of the language

As stated in the introduction, the algorithmic language has three different kinds of representation — reference, hardware, and publication — and the development described in the sequel is in terms of the reference representation. This means that all objects defined within the language are represented by a given set of symbols — and it is only in the choice of symbols that the other two representations may differ. Structure and content must be the same for all representations.

The purpose of the algorithmic language is to describe computational processes. The basic concept used for the description of calculating rules is the well known arithmetic expression containing as constituents numbers, variables, and functions. From such expressions are compounded, by applying rules of arithmetic composition, selfcontained units of the language — explicit formulae — called arithmetic statements.

To show the flow of larger computational processes, certain nonarithmetic statements are added which may describe e.g., alternatives, or recursive repetitions of computing statements.

Statements may be supported by declarations which are not themselves computing rules, but inform the translator of certain properties of objects appearing in statements, such as the class of numbers taken on as values by a variable, the dimension of an array of numbers or even the set of rules defining a function.

Sequences of statements and declarations when appropriately combined, are called programs. However, whereas complete and rigid formal rules for constructing translatable statements are described in the following, no such rules can be given in the case of programs. Consequently, the notion of program must be considered to be informal and intuitive, and the question whether a sequence of statements may be called a program should be decided on the basis of the operational meaning of the sequence.

In the sequel explicit rules — and associated interpretations — will be given describing the syntax of the language. Any sequence of symbols to which these rules do not assign a specific interpretation will be considered to be undefined. Specific translators may give such sequences different interpretations.

2. *Basic Symbols*

The reference language is built up from the basic symbols listed in Part III a. These are

1. *Letters* λ (the standard alphabet of small and capital letters)
2. *Figures* ζ (arabic numerals 0, ..., 9)
3. *Delimiters* δ consisting of
   a) operators ω:
      arithmetic operators $+ \quad - \quad \times \quad /$
      relational operators $< \quad \leqq \quad = \quad \geqq \quad > \quad \neq$
      logical operators $\neg \quad \vee \quad \wedge \quad \equiv$
      sequential operators *go to* *do* *return* *stop* *for* *if* *or* *if either* *or if*
   b) separators σ: , : ; := =: → 10 .
   c) brackets β: ( ) [ ] ↑ ↓
   d) declarators φ: *procedure* *array* *switch* *type* *comment*

Of these symbols, letters do not have individual meaning. Figures and delimiters have a fixed meaning which for the most part is obvious, or else will be given at the appropriate place in the sequel.

Strings of letters and figures enclosed by delimiters represent new entities. However, only two types of such strings are admissible:

1. Strings consisting of figures ζ only represent the *(positive) integers* G (including *0*) with the conventional meaning.

2. Strings beginning with a letter λ followed by arbitrary letters λ and/or figures ζ are called *identifiers*.

They have no inherent meaning, but serve for identifying purposes only.

3. Expressions

Arithmetic and logical processes (in the most general sense) which the algorithmic language is primarily intended to describe, are given by arithmetic and logical expressions, respectively. Constituents of these expressions, except for certain delimiters, are numbers, variables, elementary arithmetic operators and relations, and other operators called functions. Since the description of both variables and functions may contain expressions, the definition of expressions, and their constituents, is necessarily recursive.

The following are the units from which expressions are constructed.

i) (positive) Numbers N. Form: $N \sim G \,.\, G_{10} \pm G$
where each G is an integer as defined above.

$G \,.\, G$ is a decimal number of conventional form. The scale factor ${}_{10} \pm G$ is the power of ten given by $\pm G$. The following constituents of a number may be omitted in any occurrence:

The fractional part $.00\cdots0$ of integer decimal numbers;
the integer *1* in front of a scale factor;
the $+$ sign in the scale factor;
the scale factor ${}_{10} \pm 0$.

Examples:
4711
137.06
$2.9997_{10}10$
${}_{10}-12$
$3_{10}-12$

ii) Simple Variables V are designations for arbitrary scalar quantities, e.g., numbers as in elementary arithmetic.

Form: $V \sim I$

where I is an identifier as defined above.

Examples:
a
x 11
PSI 2
ALPHA

iii) Subscripted Variables V designate quantities which are components of multidimensional arrays.

Form: $V \sim I\,[I]$

where $I \sim E, E, \cdots, E$ is a *list* of arithmetic expressions as defined below. Each expression E occupies one subscript position of the subscripted variable, and is called a *subscript.* The complete list of subscripts is enclosed in the subscript brackets [ ].

The array component referred to by a subscripted variable is specified by the actual numerical value of its subscripts (cf. arithmetic expressions).

Subscripts, however, are intrinsically integer valued, and whenever the value of a subscript expression is not integral, it is replaced by the nearest integer (in the sense of proper round off).

Variables (both simple and subscripted ones) designate arbitrary real numbers unless otherwise specified. However, certain declarations (cf. *type declarations*) may specify them to be of a special type, e.g., *integral*, or *Boolean*. Boolean (or logical) variables may assume only the two values "true" and "false".

iv) Functions F represent single numbers (function values), which result through the application of given sets of rules to fixed sets of parameters.

Form: $$F \sim I\,(P, P, \ldots, P)$$

where I is an identifier, and $P, P, \cdots, P$ is the ordered list of actual parameters specifying the parameter values for which the function is to be evaluated. A syntactic definition of parameters is given in the sections on *function declarations* and *procedure declarations*. If the function is defined by a *function declaration*, the parameters employed in any use of the function are expressions compatible with the type of variables contained in the corresponding parameter positions in the function declaration heading (cf. *function declaration*). Admissible parameters for functions defined by *procedure declarations* are the same as admissible input parameters of procedures as listed in the section on *procedure statements*.

Identifiers designating functions, just as in the case of variables, may be chosen according to taste. However, certain identifiers should be reserved for the standard functions of analysis. This reserved list should contain:

abs (E)	for the modulus (absolute value) of the value of the expression E
sign (E)	for the sign of the value of E
entier (E)	for the largest integer not greater than the value of E
sqrt (E)	for the square root of the value of E
sin (E)	for the sine of the value of E

and so on according to common mathematical notation.

v) Arithmetic expressions E are defined as follows:

A number, a variable (other than Boolean), or a function is an expression.

Form: $$E \sim N$$
$$\sim V$$
$$\sim F$$

If E_1 and E_2 are expressions, the first of which are neither "$+$" nor "$-$", then the following are expressions:

1. $E \sim + E_1$
2. $\sim - E_2$
3. $\sim E_1 + E_2$
4. $\sim E_1 - E_2$
5. $\sim E_1 \times E_2$
6. $\sim E_1 / E_2$
7. $\sim E_1 \uparrow E_2 \downarrow$
8. $\sim (E_1)$

The operators $+, -, \times, /$ appearing in 1 through 6 have the conventional meaning. The parentheses $\uparrow \downarrow$ used in 7 denote exponentation, where the leading expression is the base and the expression enclosed in parentheses is the exponent.

Examples: $2 \uparrow 2 \uparrow n \downarrow \downarrow$ means $2^{(2^n)}$
$2 \uparrow 2 \downarrow \uparrow n \downarrow$ means $(2^2)^n$

The proper interpretation of expressions can always be arranged by appropriate positioning of parentheses.

An arithmetic expression is a rule for computing one real number by executing the indicated arithmetic operations on the actual numerical values of the constituents of the expression. This value is obvious in the case of numbers ***N***. For variables ***V***, it is the current value (assigned last in the dynamic sense), and for functions ***F*** it is the value arising from the computing rules defining the function (cf. *function declaration*) when applied to the current values of the function parameters given in the expression.

The sequence of operations within one expression is generally from left to right, with the following additional rules:

a) parentheses are evaluated separately

b) for operators, the conventional rule of precedence applies:

first: $\times$ /
second: $+$ $-$

In order to avoid misunderstandings, redundant parentheses should be used to express, for example, $\frac{a\,b}{c}$ in the form $(a \times b)/c$ or $(a/c) \times b$ rather than by $a \times b/c$, or $a/c \times b$ respectively, and to avoid constructions such as $a/b/c$.

Examples:

A
Alpha
Degree
A [*1,1*]
A [*j* + *k* − *2, j* − *k*]
A [*mu*[*s*]]
a × *sin* (*omega* × *t*)
0.5 × *a*[*N*×(*N*−*1*)/*2, 0*]

vi) Boolean expressions *B* are defined analogously to arithmetic expressions:

a) A truth value, a variable (Boolean by declaration), or a function (Boolean by declaration) is an expression.

Form: ***B*** $\sim$ *0* (the truth value "false")
$\sim$ *1* (the truth value "true")
$\sim$ ***V***
$\sim$ ***F***

b) If $\boldsymbol{E}_1$ and $\boldsymbol{E}_2$ are arithmetic expressions then the following arithmetic relations are expressions:

$\boldsymbol{B} \sim (\boldsymbol{E}_1 < \boldsymbol{E}_2)$
$\sim (\boldsymbol{E}_1 \leqq \boldsymbol{E}_2)$
$\sim (\boldsymbol{E}_1 \neq \boldsymbol{E}_2)$
$\sim (\boldsymbol{E}_1 \geqq \boldsymbol{E}_2)$
$\sim (\boldsymbol{E}_1 > \boldsymbol{E}_2)$
$\sim (\boldsymbol{E}_1 = \boldsymbol{E}_2)$

Such expressions take on the (current) value "true" whenever the corresponding relation is satisfied for the expressions involved, otherwise "false".

c) If $\boldsymbol{B}_1$ and $\boldsymbol{B}_2$ are expressions, the following are expressions:

$$\begin{aligned} \boldsymbol{B} &\sim \neg\, \boldsymbol{B}_1 \\ &\sim \boldsymbol{B}_1 \vee \boldsymbol{B}_2 \\ &\sim \boldsymbol{B}_1 \wedge \boldsymbol{B}_2 \\ &\sim \boldsymbol{B}_1 \equiv \boldsymbol{B}_2 \\ &\sim (\boldsymbol{B}_1) \end{aligned}$$

The operators $\neg$, $\vee$, $\wedge$, $\equiv$ have the interpretations "not", "or", "and", and "equivalent".

Interpretation of the binary operators will be from left to right. The scope of "$\neg$" is the first expression to its right. Any other desired precedence must be indicated by the use of parentheses.

Examples: $(x = 0)$
$(X > 0) \vee (y > 0)$
$(p \wedge q) \vee (x \neq y)$

4. Statements $\boldsymbol{\Sigma}$

Closed and selfcontained rules of operations are called Statements $\boldsymbol{\Sigma}$. They are defined recursively in the following way:

a) Basic statements $\boldsymbol{\Sigma}$ are those described in this section.

b) Strings of one or more statements* may be combined into a single (compound) statement by enclosing them within the "statement parentheses" *begin* and *end*. Single statements are separated by the statement separator ";".

Form: $\boldsymbol{\Sigma} \sim$ *begin* $\boldsymbol{\Sigma}; \boldsymbol{\Sigma}; \cdots; \boldsymbol{\Sigma}$ *end*

c) A statement may be made identifiable by attaching to it a label $\boldsymbol{L}$, which is an identifier $\boldsymbol{I}$, or an integer $\boldsymbol{G}$ (with the meaning of identifier). The label precedes the attached statement being labeled, and is separated from it by the separator colon (:). Label and statement together constitute a statement called "labeled statement".

Form: $\boldsymbol{\Sigma} \sim \boldsymbol{L} : \boldsymbol{\Sigma}$

A labeled statement may not itself be labeled. In the case of labeled compound statements, the closing parentheses *end* may be followed by the statement label (followed by the statement separator) in order to indicate the range of the compound statement:

Form: $\boldsymbol{\Sigma} \sim \boldsymbol{L} :$ *begin* $\boldsymbol{\Sigma}; \boldsymbol{\Sigma}; \cdots ; \boldsymbol{\Sigma}$ *end* $\boldsymbol{L};$

i) **Assignment statements** serve for assigning the value of an expression to a variable.

Form i): $\boldsymbol{\Sigma} \sim \boldsymbol{V} := \boldsymbol{E}.$

* *Declarations* which may be interspersed between statements have no operational (dynamic) meaning. Therefore, they have no significance in the definition of compound statements.

If the expression on the right hand side of the assignment delimiter $:=$ is arithmetical, the variable $\boldsymbol{V}$ on the left hand side must also be numerical, i.e., it must not be Boolean.

Generally, the arithmetic type of the expression $\boldsymbol{E}$ is determined by the constituents and operations of the expression $\boldsymbol{E}$. However $\boldsymbol{V}$ may be declared to be of a special type provided this declaration is compatible with the possible values of the expression $\boldsymbol{E}$.

Form ii): $\boldsymbol{\Sigma} \sim \boldsymbol{V} := \boldsymbol{B}$

If the expression on the right hand side of the assignment statement is Boolean, $\boldsymbol{V}$ may be any variable. This means that the truth values "true", and "false" of the Boolean expression may be interpreted arithmetically as integers "*1*" and "*0*", which may then be assigned to a numerical variable.

ii) **'Go to' statements.** Normally, the sequence of operations (described by the statement of a program) coincides with the physical sequence of statements. This normal sequence of execution may be interrupted by the use of *go to* statements.

Form: $\boldsymbol{\Sigma} \sim \underline{go\ to}\ \boldsymbol{D}$

$\boldsymbol{D}$ is a *designational expression* specifying the label of the statement next to be executed. It is either a label $\boldsymbol{L}$ or a switch variable $\boldsymbol{I}\,[\boldsymbol{E}]$ (cf. *switch declaration*), where $\boldsymbol{I}$ is an identifier and $\boldsymbol{E}$ a subscript expression. In the latter case, the numerical value of $\boldsymbol{E}$ (the value of the subscript) is an ordinal which identifies the component of the switch $\boldsymbol{I}$ (named by declaration). This element which is again a designational expression specifies the label to be used in the *go to* statement. This label determination is obviously a recursive process, since the elements of the switch may again be switch variables.

Examples:

go to hell

go to exit $[(i \uparrow 2\downarrow - j \uparrow 2\downarrow + 1)/2]$
where *exit* refers to the declaration
switch exit $:= [\boldsymbol{D}_1, \boldsymbol{D}_2, \cdots, \boldsymbol{D}_m]$

iii) **'If' Statements.** The execution of a statement may be made to depend upon a certain condition which is imposed by preceding the statement in question by an *if* statement.

Form: $\boldsymbol{\Sigma} \sim \underline{if}\ \boldsymbol{B}$

where $\boldsymbol{B}$ is a Boolean expression.

If the value of $\boldsymbol{B}$ is "true", the statement following the *if* statement will be executed. Otherwise, it will be bypassed, and operation will be resumed with the next statement following.

Example: In the sequence of statements

if $(a > 0)$; $c := a \uparrow 2\downarrow + b \uparrow 2\downarrow$;
if $(a < 0)$; $c := a \uparrow 2\downarrow + b \uparrow 2\downarrow$;
if $(a = 0)$; *go to bed*

one and only one of the three statements rightmost in each line will be executed.

iv) 'For' statements. Recursive processes may be initiated by use of a *for* statement, which causes the following statement to be executed several times, once for each of a series of values assigned to the recursing variable contained in the *for* statement.

Form: $\Sigma\sim$ a) *for* $V := l$

b) *for* $V := E_{i_1}(E_{c_1})E_{e_1}, \cdots, E_{i_k}(E_{s_k})E_{e_k}$

where l is a list of k expressions $E_1, E_2, \ldots, E_k$; and E_{i_j}, E_{s_j}, E_{e_j} are expressions. In Form a) the intent is to assign to V in succession the value of each expression of the list (expressions taken in order of listing) and the statement following the *for* statement is executed immediately following each such assignment.

In Form b) each group of expressions E_i (E_s) E_e determines an arithmetic progression. The value of E_i is the initial value, E_s gives the value of the increment (step), and E_e determines the end value which is the last term of the progression contained in the interval $[E_i, E_e]$. The intent is to assign to V each value of every progression (these again taken in the order of listing from left to right), and the statement following the *for* statement is executed immediately following each such assignment.

The effect of a *for* statement may be precisely described in terms of "more elementary" statement forms. Thus the form (a) is precisely equivalent to

$$V := E_1; \Sigma; \quad V := E_2; \Sigma; \quad \cdots \quad V := E_k; \Sigma$$

where Σ is the statement following the *for* statement.

The form (b) is precisely equivalent to

$$\begin{aligned} &V := E_{i_1}; L_1 : \Sigma^{\star}; \quad V := E_{i_1} + E_{s_1}; \quad \textit{if} \quad (V \leqq E_{e_1})^{\star\star}; \quad \textit{go to}\ L_1; \\ &\vdots \\ &V := E_{i_k}; L_k : \Sigma; \quad V := E_{i_k} + E_{s_k}; \quad \textit{if} \quad (V \leqq E_{e_k}); \quad \textit{go to}\ L_1; \end{aligned}$$

where Σ is the statement following the *for* statement.

Examples: a) *for* $I := 1\,(1)\,n$; $p := p \times y + A\,[I]$

b) *for* $a := 1, 3, 5, 9.76, \cdots, -13.75$;
begin — — — — — — — —
— — — — — — — — *end*

v) Alternative statements. An alternative statement is one which has the effect of selecting execution for one from a set of given statements in accordance with certain conditions which exist when the statement is encountered.

Form: *if either* B_1; Σ_1; *or if* B_2; ... ; *or if* B_k; Σ_k; *end*

where Σ_i is any statement other than a quantifier, i.e., *if*, *for*, or *or if*, and B_i is any Boolean expression.

---

* If Σ is a labeled statement L_1 is that label. If not the effect is as though it had a (unique) label L_1.

** This relational form obtains if the progression is increasing; if decreasing, the relation $\geqq$ is understood to be employed.

The effect of an alternative statement may be precisely described in terms of "more elementary" statement forms. Thus the above form is precisely equivalent to the sequence of statements:

if $\boldsymbol{B}_1$; *begin* $\boldsymbol{\Sigma}_1$; *go to next end*; *if* $\boldsymbol{B}_2$; *begin* $\boldsymbol{\Sigma}_2$; *go to next end*; $\cdots$; *if* $\boldsymbol{B}_k$; $\boldsymbol{\Sigma}_k$ where "*next*" is the label of the statement following the alternative statement.

Example: *if either* $(a > 0)$; $y := a + 2$; *or if* $(a < 0)$; $y := a/2$; *or if* $(a = 0)$; $y := 0.57$ *end*.

vi) 'Do' Statements. A statement, or string of statements, once written down, may be entered again (in the sense of copying) in any place of the program by employing a *do* statement which during copying permits substitution for certain constituents of the statement reused.

Form: $$\boldsymbol{\Sigma} \sim \underline{do}\ \boldsymbol{L}_1, \boldsymbol{L}_2\ (\boldsymbol{S}_{\rightarrow} \rightarrow \boldsymbol{I}, \ldots, \boldsymbol{S}_{\rightarrow} \rightarrow \boldsymbol{I})$$

where $\boldsymbol{L}_1$ and $\boldsymbol{L}_2$ are labels, the $\boldsymbol{S}_{\rightarrow}$ are strings of symbols not containing the separator $\rightarrow$ and the $\boldsymbol{I}$ are identifiers, or labels, and the list enclosed by parentheses is a substitution list.

The *do* statement operates on the string of statements from, and including, the one labeled $\boldsymbol{L}_1$ through the one labeled $\boldsymbol{L}_2$, which statements constitute the range of the *do* statement. If $\boldsymbol{L}_1$ is equal to $\boldsymbol{L}_2$, i.e., if the range is just the one statement $\boldsymbol{L}_1$, the characters ", $\boldsymbol{L}_2$" may be omitted.

The *do* statement causes itself to be replaced by a copy of the string of statements constituting its range. However, in this copy all identifiers or labels, listed on the righthand side of a separator "$\rightarrow$" in the substitution list of the *do* statement, (and which are utilized in these statements) are replaced by the corresponding strings of symbols $\boldsymbol{S}_{\rightarrow}$ on the left hand side of the separators "$\rightarrow$". These strings $\boldsymbol{S}_{\rightarrow}$ may be chosen freely with the one restriction that the substitutions produce formally correct statements in the copy*.

Whenever a *do* statement contains in its range another *do* statement, the copying, and substituting process for this second innermost *do* will be executed first.

Therefore the (actual) copy induced from a *do* statement never contains a *do* statement.

Declarations within the range of a *do* statement are not reproduced in the copy

Examples: *do 5,12* $(x[i] \rightarrow y$, *black label* $\rightarrow$ *red label*, $\cdots$, $f(x, y) \rightarrow g)$

do 12 $A, ABC\ (x \uparrow 2 \downarrow + 3/y \rightarrow A, \cdots)$

The range of a *do* statement should contain complete statements only i.e., if the *begin* (*end*) delimiter of a compound statement lies in the range of the *do*, then so should the matching *end* (*begin*). If this rule is not complied with the result of the *do* statement may not be the one desired.

* Thus, in the copy produced any designational expression whose range is a statement within the range of the *do* statement must be transformed so that its range refers to the copy produced.

vii) Stop statements. *Stop* is a delimiter which indicates an operational (dynamic) end of the program containing it. Operationally, it has no successor statement.

Form: $\Sigma \sim \underline{stop}$

viii) Return statements. *Return* is a delimiter which indicates an operational end of a *procedure*. It may appear only in a *procedure declaration* (cf. *procedure declaration*).

Form: $\Sigma \sim \underline{return}$

ix) Procedure statements. A procedure statement serves to initiate (call for) the execution of a *procedure*, that is, a closed, selfcontained process with a fixed ordered set of input and output parameters, permanently defined by a *procedure declaration*. (cf. *procedure declaration*)

Form: $\Sigma \sim \boldsymbol{I}(\boldsymbol{P}_i, \boldsymbol{P}_i, \cdots, \boldsymbol{P}_i) =: (\boldsymbol{P}_0, \boldsymbol{P}_0, \cdots, \boldsymbol{P}_0)$

Here $\boldsymbol{I}$ is an identifier which is the name of some procedure i.e., it appears in the heading of some procedure declaration (cf. *procedure declaration*), $\boldsymbol{P}_i, \boldsymbol{P}_i, \cdots, \boldsymbol{P}_i$ is the ordered list of actual input parameters specifying the input quantities to be processed by the procedure.

The list of actual output parameters $\boldsymbol{P}_0, \boldsymbol{P}_0, \ldots, \boldsymbol{P}_0$, specifies the variables to which the results of the procedure will be assigned, and alternate exits if any. The procedure declaration defining the procedure called contains in its heading a string of symbols identical in form to the procedure statement, and the formal parameters occupying input and output parameter positions there give complete information concerning the admissibility of parameters employed in any procedure call shown by the following replacement rules:

formal parameters in procedure declaration	admissible parameters in procedure statement
input parameters	
single identifier (formal variable)	any expression (compatible with the type of the formal variable)
array, i.e., subscripted variable with k ($\geqq 1$) empty parameter positions	array with n ($\geqq k$) parameter positions k of which are empty
function with k empty parameter positions	function with n ($\geqq k$) parameter positions k of which are empty
procedure with k empty parameter positions	procedure with k empty parameter positions
parameter occurring in a procedure (added as a primitive to the language)*	every string of symbols S, which does not contain the symbol "," (comma)

* Within a program certain procedures may be called which are themselves not defined by procedure declarations in the program, e.g., input — output procedures. These procedures may require as parameters quantities *outside* the language, e.g., a string of characters providing input — output format information.

output parameters	
single identifier (formal variable)	simple or subscripted variable
array (as above for input parameters)	array (as above for input parameters)
(formal) label	label

If a parameter is at the same time an input and output parameter this parameter must obviously meet the requirements of both input and output parameters.

Within a program, a procedure statement causes execution of the procedure called by the statement. The execution, however, is effected as though all formal parameters listed in the procedure declaration heading were replaced, throughout the procedure, by the actual parameters listed, in the corresponding position, in the procedure statement.

This replacement may be considered to be a replacement of every occurence within the procedure of the symbols, or sets of symbols, listed as formal parameters, by the symbols, or sets of symbols, listed as actual parameters in the corresponding positions of the procedure statement, after enclosing in parentheses every expression not enclosed completely in parentheses already.

Furthermore, any *return* statement is to be replaced by a *go to* statement referring, by its label, to the statement following the *procedure* statement, which, if originally unlabeled, is treated as having been assigned a (unique) label during the replacement process.

The values assignable to, or computable by, the actual input parameters must be compatible with type declarations concerning the corresponding formal parameters which appear in the procedure.

For actual output parameters, only type declarations duplicating given type declarations for the corresponding formal parameters may be made.

Array declarations concerning actual parameters must duplicate, in corresponding subscript positions, array declarations referring to the corresponding formal parameters.

5. Declarations Δ

Declarations serve to state certain facts about entities referred to within the program. They have no operational meaning and within a given program their order of appearance is immaterial. They pertain to the entire program (or procedure) in which they occur, and their effect is not alterable by the running history of the program.

i) Type declarations Δ. Type declarations serve to declare certain variables, or functions, to represent quantities of a given class, such as the class of integers, or class of Boolean values.

Form: $\Delta \sim \underline{\boldsymbol{type}}\,(\boldsymbol{I}, \boldsymbol{I}, \cdots \boldsymbol{I}, \boldsymbol{I}[\,], \cdots \boldsymbol{I}[,], \cdots \boldsymbol{I}[\,,\,,], \cdots)$

where ***type*** is a symbolic representative of some type declarator such as *integer* or *boolean* and the ***I*** are identifiers.

Throughout the program, the variables, or functions named by the identifiers ***I***, are constrained to refer only to quantities of the type indicated by the declaration. On the other hand, all variables, or functions which are to represent other than arbitrary real numbers must be so declared.

ii) Array declarations Δ. Array declarations give the dimensions of multi-dimensional arrays of quantities.

Form: $\Delta \sim \underline{array}\,(I, I, \cdots I[l : l'],\quad I, I, \cdots, I[l : l'], \cdots)$

where *array* is the array declarator, the I are identifiers, and the "l", and "l'" are lists of integers separated by commas.

Within each pair of brackets, the number of positions of l must be the same as the number of positions of l'.

Each pair of lists enclosed in brackets $[l : l']$ indicates that the identifiers contained in the list $I, I, \ldots, I$ immediately preceding it are the names of arrays with the following common properties:

a) the number of positions of l is the number of dimensions of every array.

b) the values of l, and l' are the lower and upper bounds of values of the corresponding subscripts of every array.

An array is defined only when all upper subscript bounds are not smaller than the corresponding lower bounds.

iii) Switch declarations Δ. A switch declaration specifies the set of designational expressions represented by a switch variable. If used in a *go to* statement, its value specifies the label of the statement called by the *go to* statement (cf. *go to* statements)

Form: $\Delta \sim \underline{switch}\quad I := (D_1, D_2, \cdots, D_n)$

where *switch* is the switch declarator, I is an identifier, and the D_i are designational expressions (cf. *go to* statement).

The switch declaration declares the list $D_1, D_2, \ldots D_n$ to be a symbolic vector (the "switch"), the designational expression D_k being the k^{th} component. Reference is made to the switch by the switch variable $I[E]$, where I is the switch identifier and E is a subscript expression. The switch variable when used in *go to* statements selects, by the actual value of its subscript, that component of the switch determining the label called for by the *go to* statement. A switch variable being a designational expression, may appear as a component of a switch.

iv) Function declarations Δ. A function declaration declares a given expression to be a function of certain of its variables. Thereby, the declaration gives (for certain simple functions) the computing rule for assigning values to the function (cf. *functions*) whenever this function appears in an expression.

Form: $\Delta \sim I_N\,(I, I \ldots, I) \quad := E$

where the I are identifiers and E is an expression which, among its constituents, may contain simple variables named by identifiers appearing in the parentheses.

The identifier I_N is the function name. The identifiers in parentheses designate the formal parameters of the function.

Whenever the function $I_N\,(P, P, \ldots, P)$ appears in an expression (a *function call*) the value assigned to the function in actual computation is the computed value of the defining expression E. For the evalution, every variable V which

is listed as a parameter I in the *function declaration*, is assigned the current value of the actual parameter P in the corresponding position of the parameter list of the function in the function call. The (formal) variables V in E which are listed as parameters in the declaration bear no relationship to variables possessing the same identifier, but appearing elsewhere in the program. All variables other than parameters appearing in E have values as currently assigned in the program.

Example:

$$I(Z) := Z + 3 \times y$$
$$\cdots\cdots\cdots\cdots\cdots\cdots$$
$$alpha := q + I(h + 9 \times mu)$$

In the statement assigning a value to alpha the computation is:

$$alpha := q + \big((h + 9 \times mu) + 3 \times y\big)$$

v) Comment declarations Δ. Comment declarations are used to add to a program informal comments, possibly in a natural language, which have no meaning whatsoever in the algorithmic language, and no effect on the program, and are intended only as additional information for the reader.

Form:

$$\Delta \sim \underline{comment}\ \mathbf{S}_{;};$$

where *comment* is the comment declarator, and $\mathbf{S}_{;}$ is any string of symbols not containing the symbol " ;".

vi) Procedure declarations Δ. A procedure declaration declares a program to be a closed unit (a procedure) which may be regarded as a single compound operation (in the sense of a generalized function) depending on a certain fixed set of input parameters, yielding a fixed set of results designated by output parameters, and having a fixed set of possible exits defining possible successors.

Execution of the procedure operation is initiated by a *procedure statement* which furnishes values for the input parameters, assigns the results to certain variables as output parameters, and assigns labels to the exits.

Form:

$$\Delta \sim \underline{procedure}\ \boldsymbol{I}(\boldsymbol{P}_i) =: (\boldsymbol{P}_0), \boldsymbol{I}(\boldsymbol{P}_i) =: (\boldsymbol{P}_0), \cdots, \boldsymbol{I}(\boldsymbol{P}_i) =: (\boldsymbol{P}_0)$$
$$\Delta\,;\Delta\,;\cdots;\Delta\,;\ \underline{begin}\ \Sigma\,;\Sigma\,;\cdots;\Delta\,;\Delta\,;\cdots;\Sigma\,;\Sigma\ \underline{end}$$

Here, the I are identifiers giving the names of the different procedures contained in the procedure declaration. Each P_i represents an ordered list of formal input parameters, each P_0 a list of formal output parameters which include any exits required by the corresponding procedures.

Some of the strings "$=: (P_0)$" defining outputs and exits may be missing in which case corresponding symbols "$I(P_i)$" define a procedure that may be called within expressions.

The Δ in front of the delimiter *begin* are declarations concerning only input and output parameters. The entire string of symbols from the declarator *procedure* (inclusive) up to the delimiter *begin* (exclusive) is the *procedure heading*. Among the statements enclosed by the parentheses *begin* and *end* there must be, for each identifier I listed in the heading as a procedure name, exactly one statement

labeled with this identifier, which then serves as the entry to the procedure. For each "single output" procedure $\boldsymbol{I}(\boldsymbol{P}_i)$ listed in the heading, a value must be assigned within the procedure by an assignment statement "$\boldsymbol{I} := \boldsymbol{E}$", where $\boldsymbol{I}$ is the identifier naming that procedure.

To each procedure listed in the heading, at least one *return* statement must correspond within the procedure. Some of these *return* statements may however be identical for different procedures listed in the heading.

Since a procedure is a self-contained program (except for parameters), the defining rules for statements and declarations within procedures are those already given. A formal input parameter may be

a) a single identifier $\boldsymbol{I}$ (formal variable),

b) an array $\boldsymbol{I}$ [, , $\cdots$,] with k ($k = 1, 2, \ldots$) empty subscript positions,

c) a function $\boldsymbol{F}$ (, , $\cdots$,) with k ($k = 1, 2, \ldots$) empty parameter positions,

d) a procedure $\boldsymbol{P}$ (, , $\cdots$,) with k ($k = 1, 2, \ldots$) empty parameter positions,

e) an identifier occurring in a procedure which is added as a primitive to the language.

A formal output parameter may be

a) a single identifier (formal variable)

b) an array with k ($k = 1, 2, \ldots$) empty subscript positions

A formal (exit) label may only be a label.

A label is an admissible formal exit label if, within the procedure, it appears in *go to* statements or *switch* declarations.

An array declaration contained in the heading of the procedure declaration, and referring to a formal parameter, may contain expressions in its lists defining subscript ranges. These expressions may contain

1. numbers
2. formal input variables, arrays, and functions.

All identifiers and all labels contained in the procedure have identity only within the procedure, and have no relationship to identical identifiers or labels outside the procedure, with the exception of the labels identical to the different procedure names contained in the heading.

A procedure declaration, once made, is permanent, and the only identifiable constituents of the declaration are the procedure declaration heading, and the entrance labels. All rules of operations and declarations contained within the procedure may be considered to be in a language different from the algorithmic language. For this reason, a procedure may even initially be composed of statements given in a language other than the algorithmic language, e.g., a machine language may be required for expressing input-output procedures.

A tagging system may be required to identify the language form in which procedures are expressed. The specific nature of such a system is not in the scope of this report.

Thus by using procedure declarations, new primitive elements may be added to the algorithmic language at will.

Part III

a) Basic symbols

delimiters δ:

Operators		*Separators*	*Brackets*	*Declarators*
$\omega \sim +$	*go to*	$\sigma \sim$,	$\beta \sim$ (	$\varphi \sim$ *procedure*
$=$	*do*	:	)	*switch*
$\times$	*return*	;	[	*array*
$/$	*stop*	: =	]	***type****
$\neg$	*for*	= :	$\uparrow$	*comment*
$\vee$	*if*	$\rightarrow$	$\downarrow$	
$\wedge$	*or if*	10	*begin*	
$=$	*if either*	.	*end*	
$\neq$				
$\equiv$				
$>$				
$\geqq$				
$\leqq$				
$<$				

non-delimiters μ:

letters

$\lambda \sim A$ through Z
a through z

digits

$\zeta \sim 0$ through 9

b) Syntactic skeleton

Syllables:

list
$\boldsymbol{l \sim E, E, \cdots, E}$

simple variable
$\boldsymbol{V \sim I}$

subscripted variable
$\boldsymbol{V \sim I\,[E\,,E\,,\cdots,E]}$

function
$\boldsymbol{F \sim I\,(P\,,P\,,\cdots,P)}$

expression and Boolean expression

$\left.\begin{matrix}\boldsymbol{E}\\ \boldsymbol{B}\end{matrix}\right\}$ For the composition rules see the appropriate sections in Part II.

statement label
$\boldsymbol{L \sim I}$
$\boldsymbol{G}$

* Representant

designational expression
$D \sim L$
$\quad I\,[E]$

parameters
P For the composition rules see the appropriate sections in Part II

identifier
$I \sim \lambda\,\mu\,\mu\,\mu \cdots \mu$

integer
$G \sim \xi\,\xi\,\xi \cdots \xi$

number
$N \sim G\,.\,G_{10} \pm G$
(the first G: may be empty; the second G: may be empty)

string of symbols
$S_\alpha \sim xxx \cdots xx$ where x is not α, and α is a particular delimiter

Statements Σ:

assignment statement
$\Sigma \sim V := E$
$\quad V := B$

compound statement
$\Sigma \sim$ *begin* $\Sigma\,;\,\Sigma \cdots ;\,\Sigma$ *end*
(at least one Σ)

labelled statement
$\Sigma \sim L : \Sigma$ where Σ is unlabeled.

go to — statement
$\Sigma \sim$ *go to* D

do — statement
$\Sigma \sim$ *do* $L\,,L \quad (S_{\rightarrow} \rightarrow I, S_{\rightarrow} \rightarrow I, \cdots, S_{\rightarrow} \rightarrow I)$
(*do* L , L: may be empty; $(S_{\rightarrow} \rightarrow I, \cdots)$: may be empty)

quantifier statements
$\Sigma \sim$ *if* B
$\quad$ *for* $V := l$
$\quad$ *for* $V := E\,(E)\,E, E\,(E)\,E, \cdots, E\,(E)\,E$

alternative statement
$\Sigma \sim$ *if either* B_1; Σ_1; *or if* B_2; Σ_2; $\cdots$; *or if* B_k; Σ_k *end*

stop- and return-statement
$\Sigma \sim$ *stop*
$\quad$ *return*

procedure statement
$\Sigma \sim I(R) =: (R)$ where $R \sim P, P \ldots, P$

Declarations Δ:

Function declaration
$\Delta \sim I(R) := E$

Procedure declaration
$\Delta \sim$ *procedure* $I(R) \underbrace{=: (R)}_{\text{may be empty}} I(R) \quad \underbrace{=: (R)}_{\text{may be empty}} \ \ldots \ I(R) \underbrace{=: (R)}_{\text{may be empty}}$
$\Delta\,;\, \Delta; \cdot\cdot\,;\, \Delta;$ *begin* $\Sigma;\, \Sigma; \cdots \Delta; \Delta; \cdots;\, \Sigma;\, \Sigma$ *end*
where $R \sim P, P, P, \cdots, P, P$

Switch declaration
$\Delta \sim$ *switch* $I := (D, D, \cdots, D)$

Array declaration
$\Delta \sim$ *array* $(I, I, \cdots, I([l:l], I, \cdots, I[l:l], I, \cdots)$

Symbol classification declaration
$\Delta \sim$ *type* $(I, I, \cdots, I)$

Comment declaration
$\Delta \sim$ *comment* S;

c) *Publication language*

As stated in the introduction, the reference language is a link between hardware languages and handwritten, typed or printed documentation.

For transliteration between the reference language and a language suitable for publications*, the following

transliteration rules
may be used

reference language		*publication language*
subscript brackets	[ ]	lowering of the line between the brackets
exponentiation parentheses	↑ ↓	raising of the line between the arrows
parentheses	( )	any form of parentheses, brackets, braces
basis of ten	10	raising of the ten and of the following integral number, inserting of the intended multiplication sign
statement separator	;	*line convention*: each statement on a separate line *may be used*

Furthermore, if line convention is used, the following changes may be simultaneously used:

multiplication cross	×	multiplication dot .
decimal point	.	decimal comma ,
separation mark	,	any common separation mark that will not be ambiguous

* For example, for lectures in numerical analysis.

Example

Integration of a function $F(x)$ by Simpson's Rule. The values of $F(x)$ are supplied by an assumed existent function routine. The mesh size is halved until two successive Simson sums agree to within a prescribed error. During the mesh reduction $F(x)$ is evaluated at most once for any x. A value V greater than the maximum absolute value attained by the function on the interval is required for initializing.

abs (absolve value) is the name of a standard procedure always available to the programmer so that it need not be supplieo as an input parameter.

```
procedure    Simps (F( ), a, b, delta, V)
comment      a, b are the min. and max. resp. of the points def. interval of integ.
             F( ) is the function to be integrated.
             delta is the permissible difference between two successive Simpson sums.
             V is greater than the maximum absolute value of F on a, b;
begin
Simps: Ibar  := V × (b − a)
       n     := 1
       h     := (b − a)/2
       J     := h × (F(a) + F(b))
Jl:    S     := 0
       for   k := 1 (1) n
             S :=
             S + F(a + (2 × k − 1) × h)
             I := J + 4 × h × S
       if    (delta < abs (I − Ibar))
begin  Ibar  := I
             J := (I + J)/4
             n := 2 × n ; h := h/2
       go to Jl
end    Simps := I/3
return
integer (k, n)
end    Simps
```

Computation Laboratory Carnegie Institute of Technology
Pittsburgh, Pennsylvania
and
Mathematisches Institut der Universität Mainz

(Received October 20, 1958)