

CRMS APL PROCESSOR

REFERENCE MANUAL

Paul McJones

February 2, 1973

Revised May 30, 1974

Systems Group

Technical Document

Center for Research in Management Science

University of California

Berkeley

This work was done as part of the Systems Development effort
under National Science Foundation Grant GS-32138.

This manual and the original microcode have undergone several modifications since the original publication date of this manual.

C. GRANT

W. GREINER

TABLE OF CONTENTS

1. Introduction	1
2. APL Objects	3
2.1. Floating-Point Numbers	3
2.2. Integers	5
2.3. Characters	5
2.4. Arrays	5
2.5. Indirect Parameter Words	7
2.6. The "Undefined Value"	8
3. The Data Segment	9
3.1. Processor State Area	9
3.2. Global Variables	13
3.3. Runtime Stack	13
3.4. Array Block Storage	16
4. The Procedure Segment	22
4.1. Function Blocks and the Function Directory	22
4.2. Components of a Function Block	22
4.2.1. Function Block Header	22
4.2.2. Code	25
4.2.3. Line Table	25
5. Controlling the APL Processor	26
6. The Instructions	30
6.1. Operands, Results, and Reference Counts	30
6.2. Traps	30
6.3. The Individual Instructions	31
The FETCH Family	33
The MEMREF Family	36
The GENSCALAR Family	41
INDEX	49
SHORT_CONSTANT	50
The ETC Family	52
Appendix 1: Index to the APL Instructions	71
Appendix 2: Summary of Trap Classes and Numbers	72
Appendix 3: Array Block Storage Allocation Algorithms	74
Appendix 4: The Handling of Exponent Underflow	76

1. Introduction

One of the major goals of the new CRMS computer system is to provide efficient, time-shared APL. The system contains two processors with common main (core) and auxiliary (disk) memories, as well as controllers for a variety of peripheral devices. One processor (the primary subject of this document) is specially tailored for the execution of APL programs. The other processor is optimized for executing programs in the SIMPLE programming language and is described elsewhere.

Rather than directly interpreting source language character strings, the APL processor executes an internal "machine language" form adapted for that purpose. Three types of changes are involved in the mapping from external to internal APL:

- (1) (literal notation $\Rightarrow$ binary notation) Each numeric or character literal is replaced by a binary encoding of the number or character.
- (2) (identifier $\Rightarrow$ variable address) An arbitrary correspondence between the identifiers of a program and unique consecutive integers is formed; then all occurrences of an identifier are replaced by its number (address).
- (3) (infix $\Rightarrow$ postfix) The expressions are reordered so that an operator follows its operands instead of preceding (monadic) or separating (dyadic) them.

These changes are made for technological rather than logical reasons. Arithmetic can be performed much faster on operands in binary notation than in decimal notation. Similarly, binding identifiers to addresses in advance

saves time for the processor, and the resultant addresses are in general more compact than the identifiers they replace.

Programs exist for the SIMPLE processor which translate external to internal APL and perform the concomitant editing, debugging, and "system" functions. Thus the sole responsibility of the APL processor is to execute internal APL programs, and, indeed, this processor is capable of nothing else. Since the two processors share core memory, programs and data prepared by the SIMPLE processor are directly accessible by the APL processor.

This manual assumes familiarity with APL\360; for example, see APL\360 Reference Manual by Sandra Pakin (Science Research Associates, 1972).

2. APL Objects

APL is a language for expressing algorithms which manipulate numbers, characters, and n-dimensional arrays. In this section we will describe how these objects are represented internal to the APL processor.

Every value occupies a single 32-bit word. Because of the dynamic nature of APL, stemming from its lack of compile-time type declarations, each value must carry explicit type information at run-time. Thus type bits and value are packed together into a 32-bit word or descriptor. In Figure 2.1 are illustrated the formats of the various types of values; some further explanation is in order.

2.1. Floating-Point Numbers

A floating-point number consists of a signed, fixed-point coefficient (with the radix point lying between the first and second significant bits) and an "exponent" or power-of-two scale factor. If F is a floating-point number whose sign, fraction, and exponent are, respectively, s , f and e (all taken as unsigned integers), then

$$F = (1 - 2s) \cdot \left(1 + \frac{f}{2^{22}}\right) \cdot 2^{(e-128)}.$$

As can be seen from Figure 2.1, all floating-point numbers are "normalized," that is, if $c = \left(1 + \frac{f}{2^{22}}\right)$ is the coefficient of a floating point number F , then $1 \leq c \leq 2 - 2^{-22}$. Equivalently, the "leading" bit of F (the first bit after the sign of F) is a one. The bit patterns corresponding to unnormalized numbers are thus available to encode other types of values; Figure 2.1 shows how the type field of a value not a floating-point number is used.

2.2. Integers

An alternative representation for integers of magnitude less than 2^{23} is available. In particular, zero must be represented in this format since it is not a normalizable floating-point number. A point concerning zero is that it is always represented with the sign bit $s = 0$; "minus zero" is not allowed.

2.3. Characters

A scalar character is represented by an 8-bit code chosen from an extended-ASCII character set which includes the special characters necessary for APL.

2.4. Arrays

A descriptor for a scalar (numeric or character) might be termed an "immediate descriptor" since it completely specifies the identity of its value. In contrast, an array descriptor merely points to an array block stored elsewhere which contains the elements and shape information of the array. This scheme has two benefits:

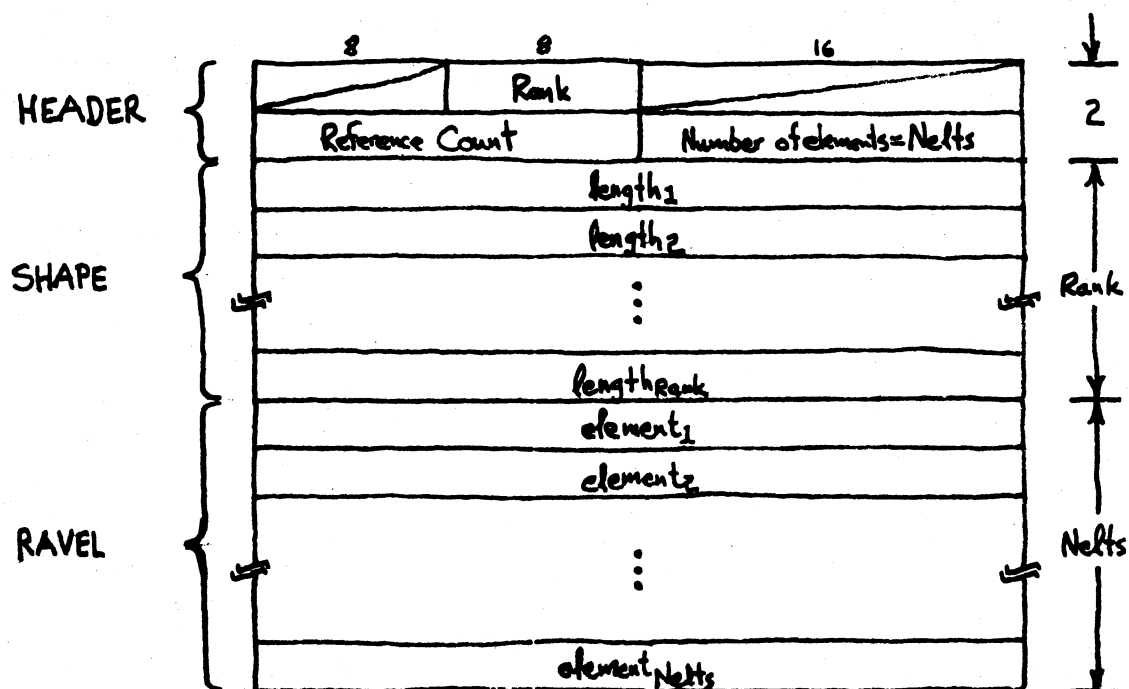
- (1) All descriptors are a uniform one word in size;
- (2) Several distinct array descriptors may point to the ~~same~~ array block.

The form of an array block is given in Figure 2.2. Following is a description of the various fields.

The rank of an array is the number of dimensions it has; the numbers $\text{length}_1, \text{length}_2, \dots, \text{length}_{\text{rank}}$ in the array block are called the shape of the array and must be regular integer-type descriptors for nonnegative numbers. If the value of length_1 of an array equals n , then there are n

12 January 73

6



 = allocator field

APL Array format

Figure 2.2

elements along the i^{th} coordinate of the array; the total number of elements in the array is

$$\text{Nelts} = \prod_{i=1}^{\text{rank}} \text{length}_i .$$

The cases $\text{Nelts} = 0$ and $\text{Nelts} = 1$ are both possible. $\text{Rank} = 0$ should never occur.

The elements of the array may be any scalar descriptors, and are stored in a linear sequence known in APL as the ravel of the array. Since it is possible for several array descriptors to point to the same array block, a reference count is used to determine when the storage for an array block may be released. This count is a field in the array block giving the number of array descriptors currently pointing to this array block; it is updated whenever such a descriptor is created or destroyed. If a reference count is reduced to zero, then the storage occupied by that array block is returned to a free pool.

2.5. Indirect Parameter Words

These descriptors are not for ordinary values, but are used in conjunction with pass-by-reference parameters to functions. Each locates a local or global variable acting as a pass-by-reference formal parameter variable; references by load and store instructions directed at the formal variable instead specify the variable located by the indirect parameter word. The variable to be specified is identified by one bit (labeled "seg" in Figure 2.1) signifying local or global ($0 \Rightarrow \text{local}$, $1 \Rightarrow \text{global}$) and sixteen bits specifying an address. For a local variable, this address is relative to the state variable SBASE (see Section 3.3); for a global variable this is just the ordinary address of that variable in the data segment.

2.6. The "Undefined Value"

This descriptor (which is defined!) is used to initialize global and local variables so that certain references[†] to variables which haven't been explicitly assigned a value will cause a value error trap. The processor automatically initializes local variables (excluding formal parameters with a descriptor for the "undefined value" during the function call instruction; such initialization for the global variables and local variables of the initial stack frame must be done in software on the SIMPLE processor before starting the APL program.

[†] SHORT_LOCAL, LONG_LOCAL, LONG_GLOBAL and ASSIGN_INDEXED .

3. The Data Segment

A program for the APL processor consists of a pair of procedure segments together with a data segment. The procedure segments (described in Section 4) contain reentrant code, and express the invariant algorithm of the program. The data segment, on the other hand, contains all the storage used to run the program: that occupied by the variables of the program as well as that used by the processor for its internal data structures.

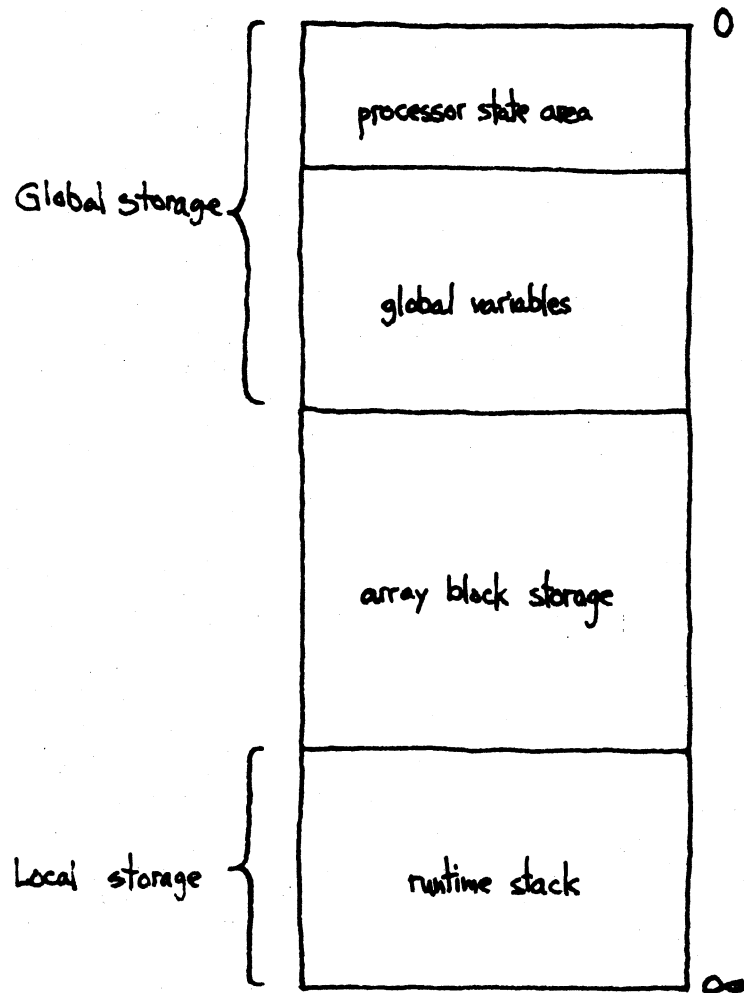
Figure 3.1 shows the overall layout of the data segment. (Higher addresses are toward the bottom of the figure, as is the case throughout this document.) The region labeled "global storage" holds information accessible throughout the execution of the program, and is statically allocated before execution begins. Local storage consists of information for each of a sequence of function activations. It is allocated using a last-in-first-out stack discipline dictated by the nested activations of functions resulting from the flow of control through the program. Finally, array block storage contains arrays each of which is referenced by one or more variables or temporary cells. Space in this area is dynamically allocated.

3.1. Processor State Area

An assortment of state variables used by the processor is gathered together at the beginning of the data segment (see Figure 3.2). The individual state variables will be described later: SBASE, LTOP and LBASE in Section 3.3 (Runtime Stack); TRAPCLASS and TRAPNUMBER in Section 6.2 (Traps); saved-NELTS, -OPNO, -APTR, -BPTR, and -RPTR in Section 6.3 (under The GENSCALAR Family); and ROVER and BLOCKPTR in

11 January 73

10

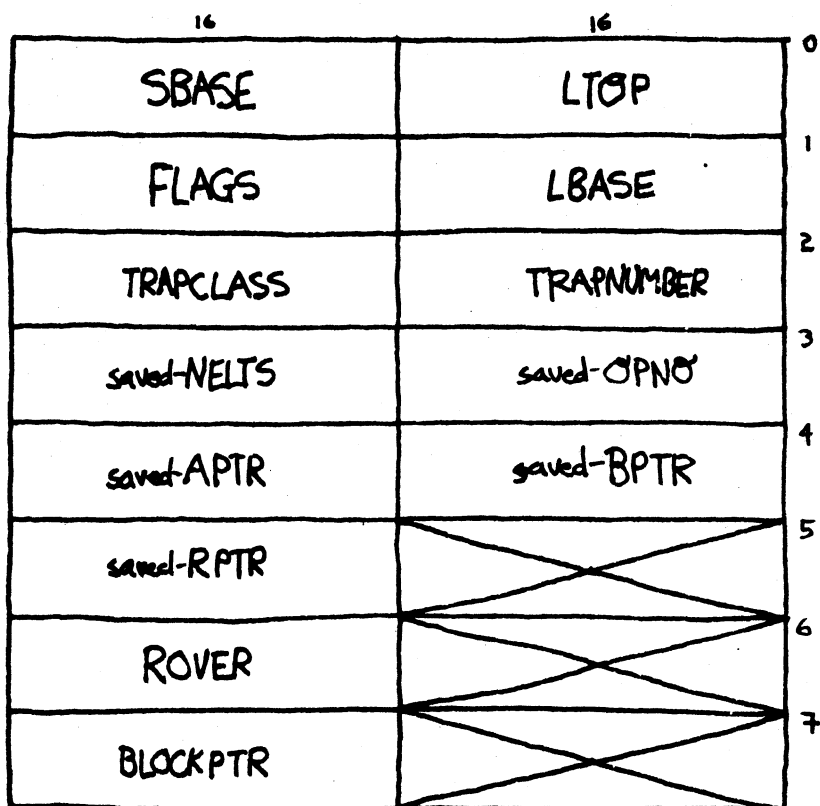


APL Data Segment

Figure 3.1

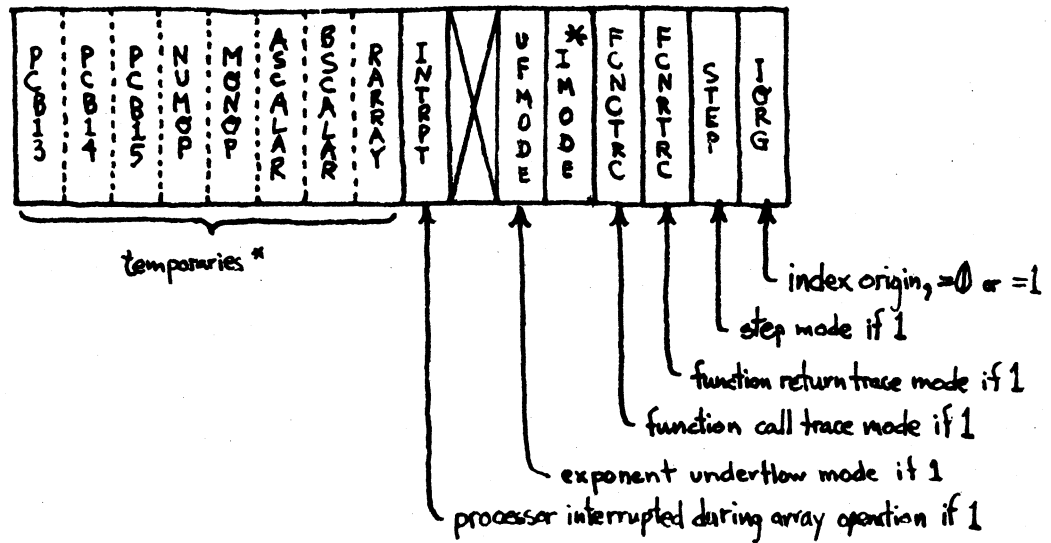
11 January 73

11



APL Processor State Area

Figure 3.2



* If the processor is stopped with $FLAGS\$INTRPT=1$, then these temporary bits must not be changed.

APL Processor State FLAGS

Figure 3.3

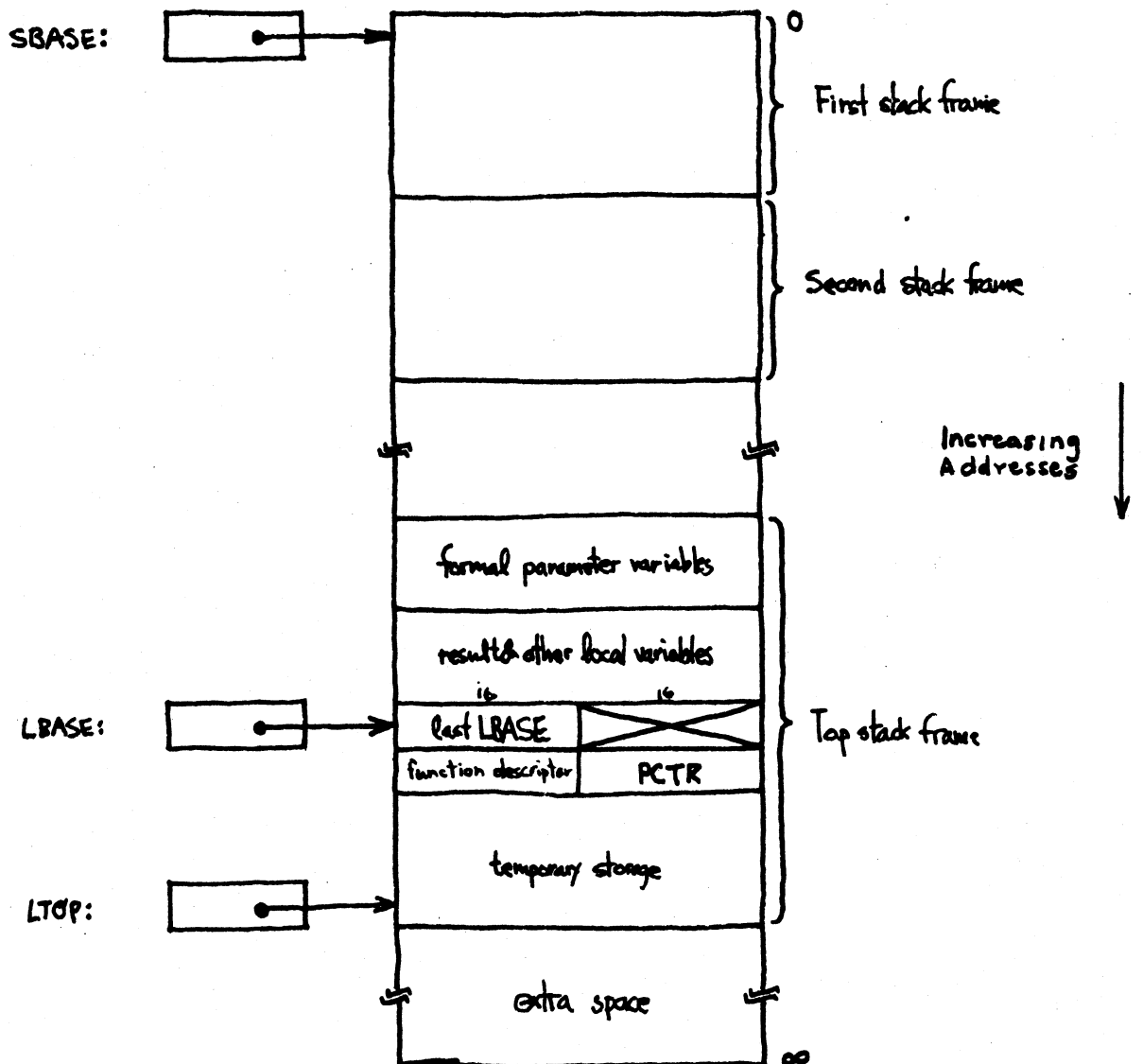
Section 3.4 (Array Block Storage). This leaves FLAGS, which is really a collection of one-bit variables, as shown in Figure 3.3. IORG is the APL index origin (see Section 6.3, various instructions). FCNCTRC and FCNRTRC control function call and return tracing (see Section 6.3, CALL and RETURN). STEP controls step tracing, another debugging tool (see Section 6.2, step trap). UFMODE is used to specify the action to be taken when exponent underflow occurs (see Appendix 4). INTRPT and the "temporary" flag bits are discussed in Section 6.3 under The GENSCALAR Family.

3.2. Global Variables

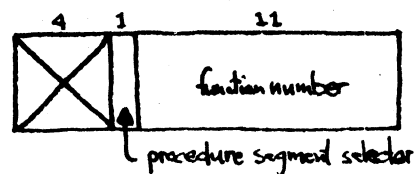
Each global variable in the program is allocated a single cell in a special area of the data segment. This cell contains a descriptor for the value of the variable, be it scalar, array, or "undefined." Global variables are addressed by integers larger than seven. The addresses zero through seven are invalid since they correspond to the cells of the processor state area. Note that when execution of a program begins, each of the global variables should contain the standard "undefined" value (see Section 2.6, The "Undefined Value"); the processor does not itself initialize these variables.

3.3. Runtime Stack

As was mentioned earlier, the processor maintains a pushdown stack for storage local to each function activation. As shown in Figure 3.4, a stack frame (entry) contains one word for each formal parameter or other local variable, two words of local processor state information, and a variable number of words containing descriptors for temporary values. At any moment the most recently activated function is being executed, and its stack frame, called the current frame, is on top of the stack.



Note: the function descriptor has the following form -



APL Runtime Stack

Figure 3.4

Several state variables are used to delimit the stack. SBASE contains the address in the data segment of the first word of the first (oldest) stack frame, and is never changed by the processor. LBASE points to the current stack frame, and LTOP points to the topmost temporary value cell of the current frame. These two state variables contain addresses relative to SBASE. The processor updates LBASE for each function call or return. LTOP is decremented by the various processor instructions when they remove their operand(s), and is incremented when they stack their result(s).

LBASE points not to the first word of the current frame, but to the first word of local state information. The local variables (including actual parameters) are addressed by 12 bit 2's complement negative integers (-1, -2, etc.), which the processor interprets relative to LBASE (see figure 3.4). If a function has a result variable, it is always assigned the address -1 (so the result variable is in the cell immediately preceding the first word of local state information).

The local state information in a stack frame merits some discussion. Since stack frames are not of constant length, some form of backpointer is necessary. This need is filled by the "last LBASE" field, which contains the value of LBASE for the function activation immediately preceding in the stack. The "function descriptor" and "PCTR" fields in a stack frame refer not to the preceding activation, but to the given one. These fields identify, respectively, the function activated and the location in that function of the next instruction executed. In the context of a given program, a function is identified by one bit to select between the two procedure segments together with an 11-bit integer designating one of the functions in that segment. The location of an instruction is always

specified as a byte address (in the corresponding function block) of the first byte of the instruction (which may be one or more bytes long). In this manual, unqualified mention of the state variable PCTR will always refer to the field of that name in the current stack frame.

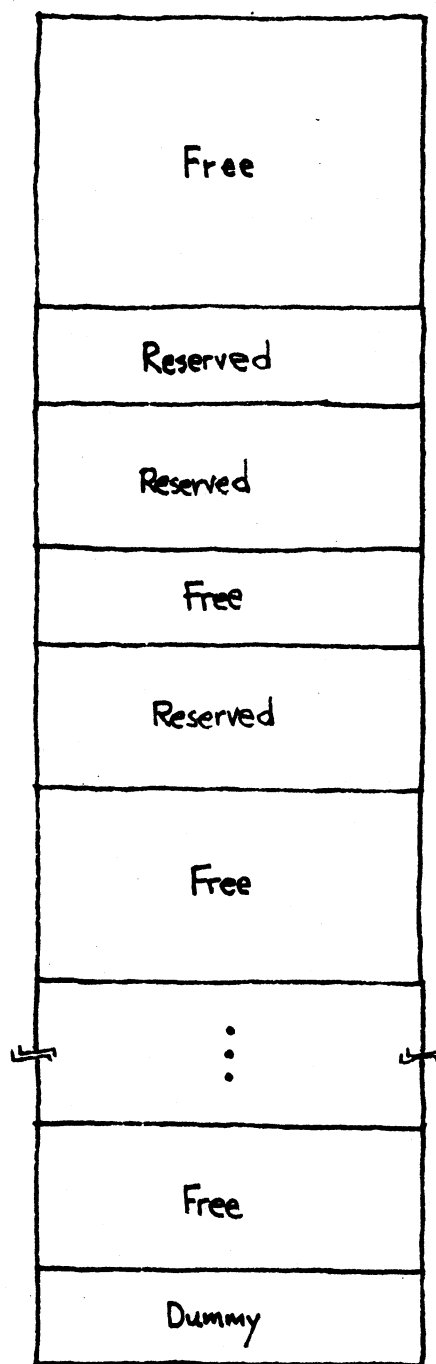
3.4. Array Block Storage

Descriptors for array values exist in various parts of the data segment: in global variables, in local (and formal parameter) variables, and in temporary storage cells. The actual arrays, however, are gathered together in array block storage. Space in this region is dynamically allocated using a scheme of "boundary tags and roving free pointer," as described in The Art of Computer Programming, Vol. 1, by Donald Knuth (Addison-Wesley, 1968).

Array block storage consists of a sequence of reserved and free blocks. Each reserved block contains an array, whose reference count field is by definition nonzero. (See Section 2.4 for the description of the structure of an array per se.) As the name suggests, free blocks contain space available for future use. In the interest of avoiding fragmentation of storage, no free block is allowed to exist adjacent to another free block (see Figure 3.5 for a typical configuration).

When the array in a reserved block is no longer needed, there are four cases to consider:

- (1) Both the preceding and following blocks are free; the block being freed should be coalesced with both its neighbors into one big free block.
- (2) Only the preceding block is free; it should be extended to include the following newly freed space.



Typical APL Array Storage Configuration

Figure 3.5

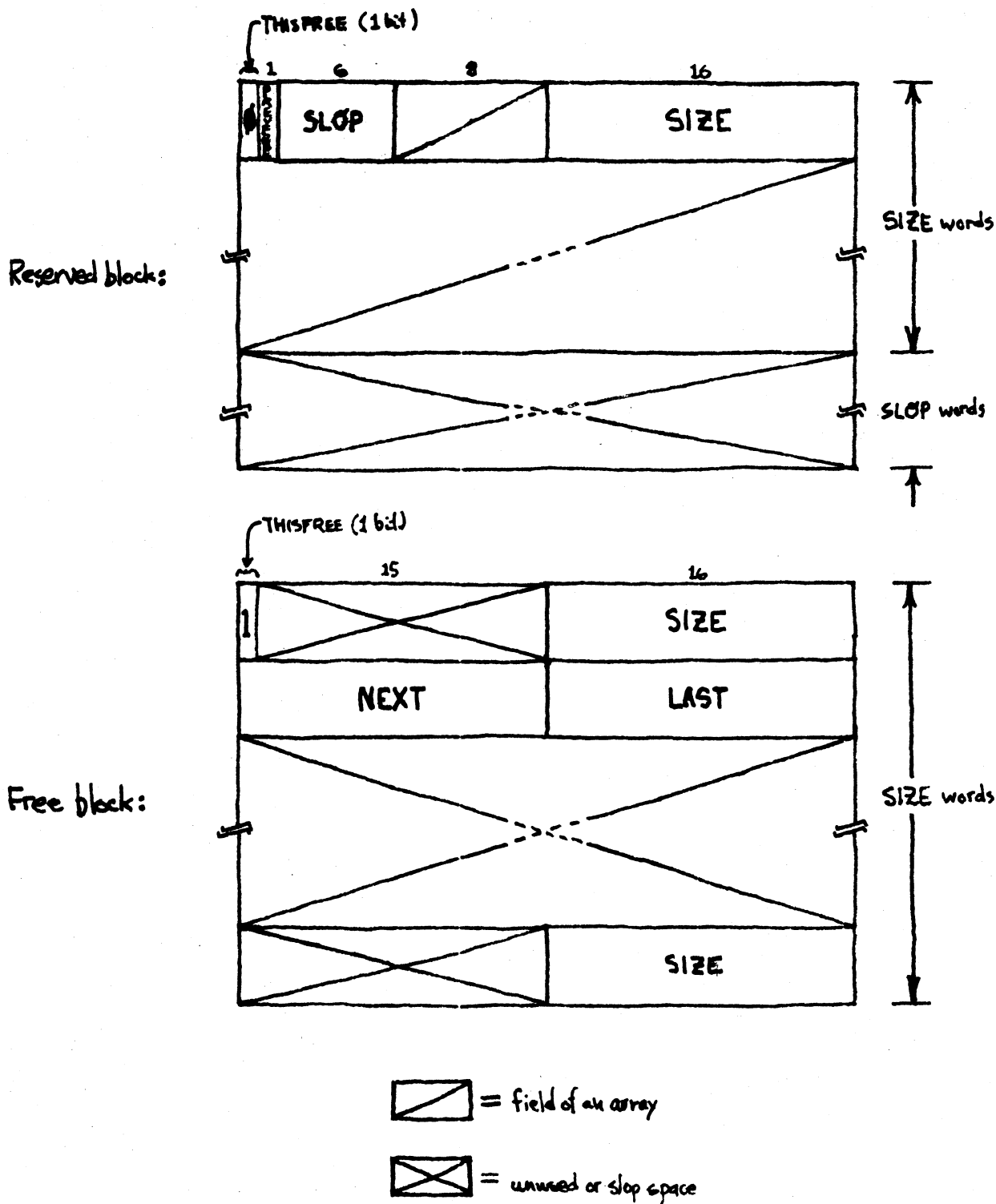
- (3) Only the following block is free; it should be extended to include the preceding newly freed space.
- (4) Neither the following nor the preceding blocks are free. The block being freed cannot be coalesced.

Figure 3.6 shows the format of reserved and free blocks from the allocator's point of view (thus leaving out the details of the array in a reserved block). Note that given the starting location of a block, it is possible to tell whether the block is free or reserved (since the Boolean field `THISFREE` occupies the left-most bit of the first word of every block). Also, every reserved block has another Boolean field, `PREVFREE`, indicating whether the preceding block is free. With these fields and the `SIZE` fields it is possible to discriminate between the cases (1) through (4) listed above. These cases assure that `THISFREE` and `PREVFREE` not happen simultaneously.

All of the free blocks in the array block storage area are joined into a doubly linked ring using the `NEXT` and `LAST` fields appearing in every free block. The state variable `ROVER` always points to some member of this ring. To ensure that the ring is never empty (so the value of `ROVER` is well defined), there is a dummy free block with an indicated size of zero which can never be allocated. This dummy free block serves another purpose. It lies at the very end of array block storage, and has its `THISFREE` bit set false (!) to inhibit coalescing with the last real block. Figure 3.7 gives a typical configuration of the free block chain.

`BLOCKPTR` always points to the most recently allocated block.

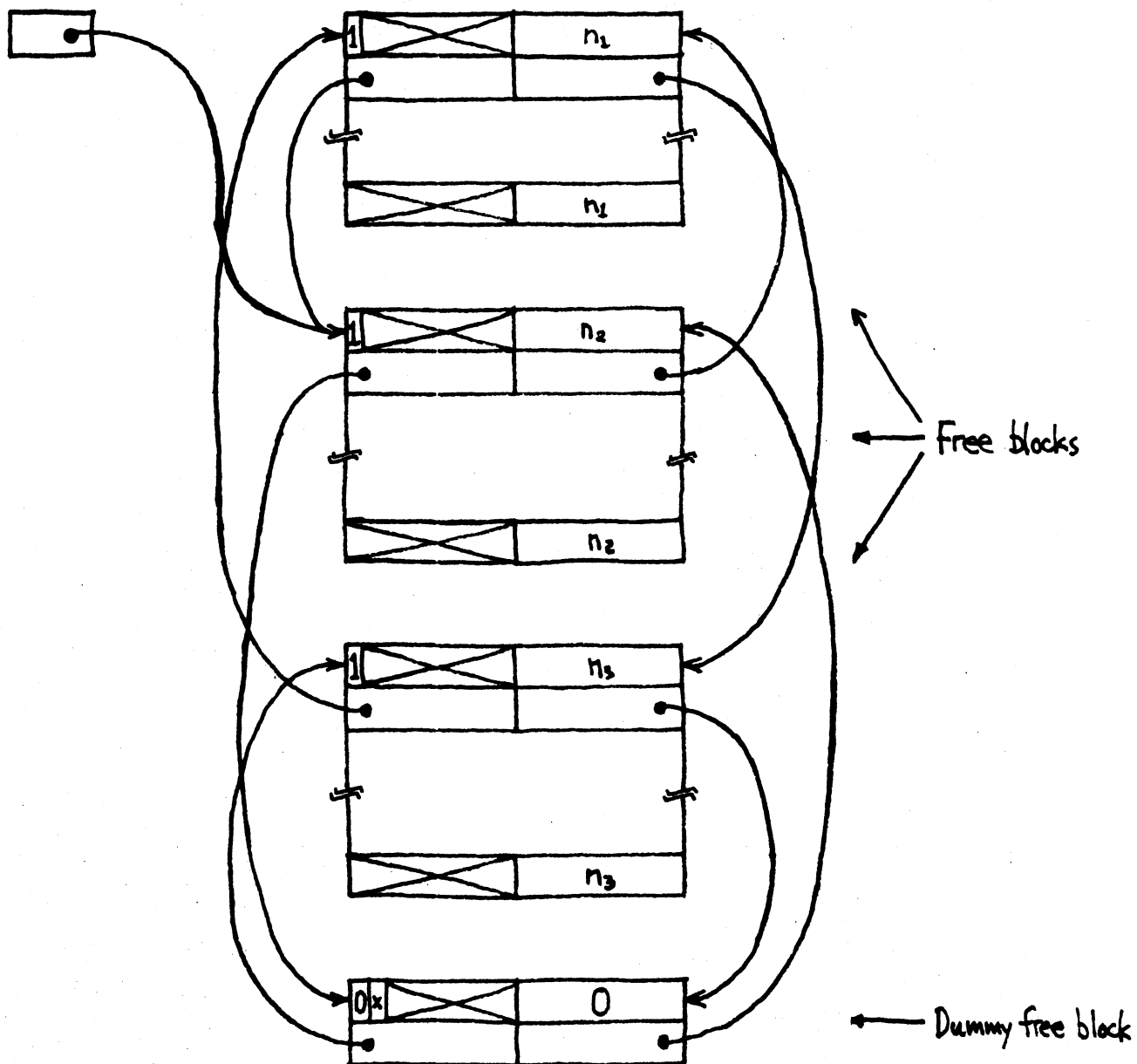
Note (in figure 3.6) that a reserved block may end with up to 4 words of "slop" which are not necessary to hold the actual array. A field named



APL Array Storage Reserved & Free Block Formats

Figure 3.6

ROVER:



APL Array Storage -
Typical Free Chain Configuration

Figure 3.7

SLOP in the first word of the block holds the number of slop words, which incidentally are not counted in the SIZE field. The main reasons for slop words are that (1) a free block must be at least three words long (If a smaller packet of unused words is created by breaking a larger free block, it must be attached to the preceding reserved block as slop.) and (2) a reserved block housing an APL array (See figure 2.2.) takes up at least three words. The smallest array, the null vector, needs exactly three words.

4. The Procedure Segment

4.1. Function Blocks and the Function Directory

Just as an APL source program consists of a set of function definitions, a procedure segment contains a sequence of function blocks. In the interest of function block relocatability, the procedure segment begins with a directory each of whose entries contains the starting address in the segment of the corresponding function block. Thus each function can always be referred to by a unique ordinal number, even if the lengths of functions change. Figure 4.1 shows the over-all layout of a procedure segment.

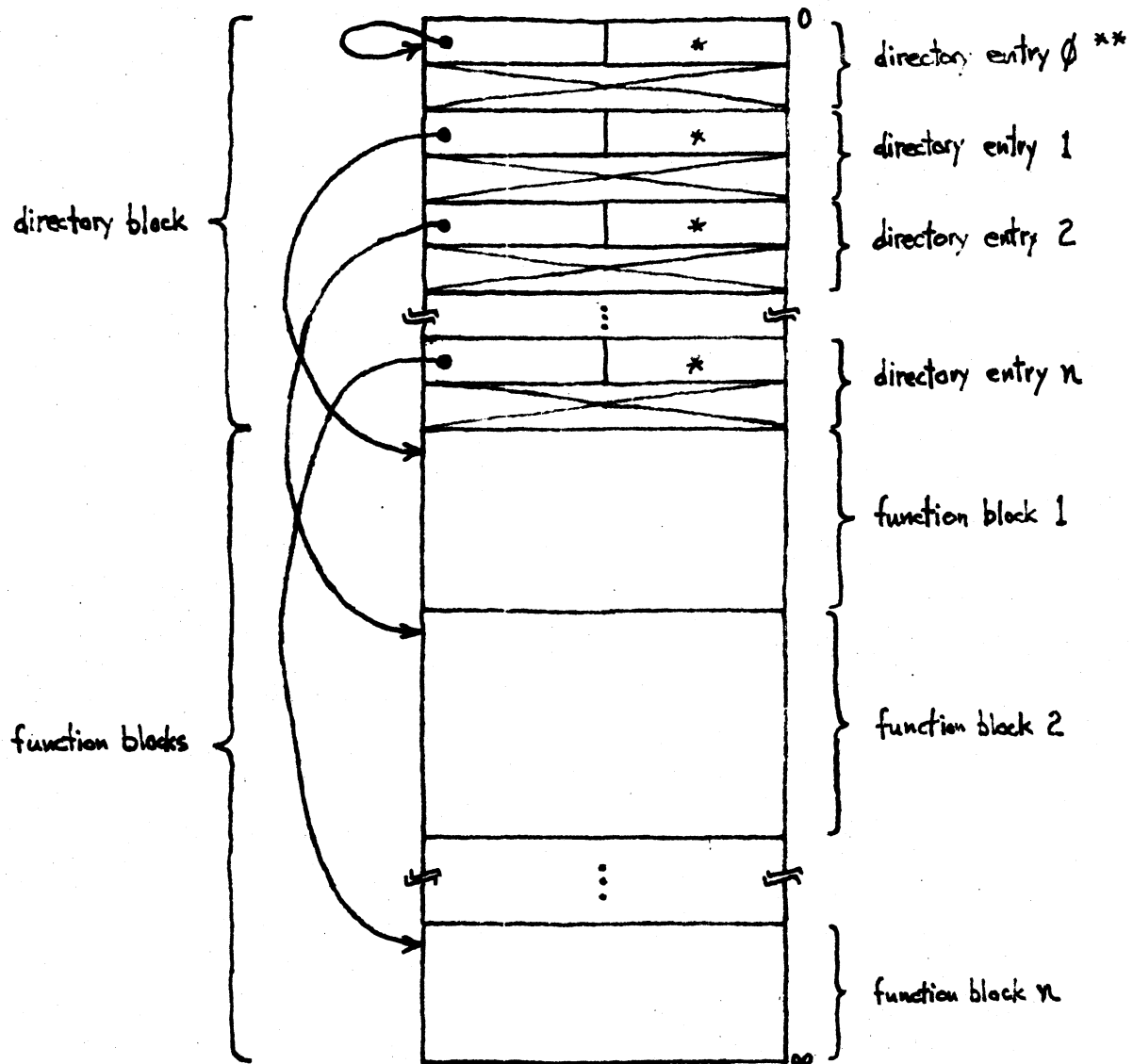
4.2. Components of a Function Block

A function block is the compiled version of an APL source function definition, and is diagrammed in Figure 4.2. The two-word header contains various flags and fields; it corresponds roughly to the header or "line zero" of a source function. Next in a function block is the actual code, a sequence of variable-length instructions. Completing a function block is its line table, which serves to map APL source statement numbers into addresses in the code.

4.2.1. Function block header. The fields in the header of a function block have the following meanings:

CALLTRACE--if this bit is one and the state flag FCNCTRC is also one, the processor will trap just after executing a call instruction designating this function;

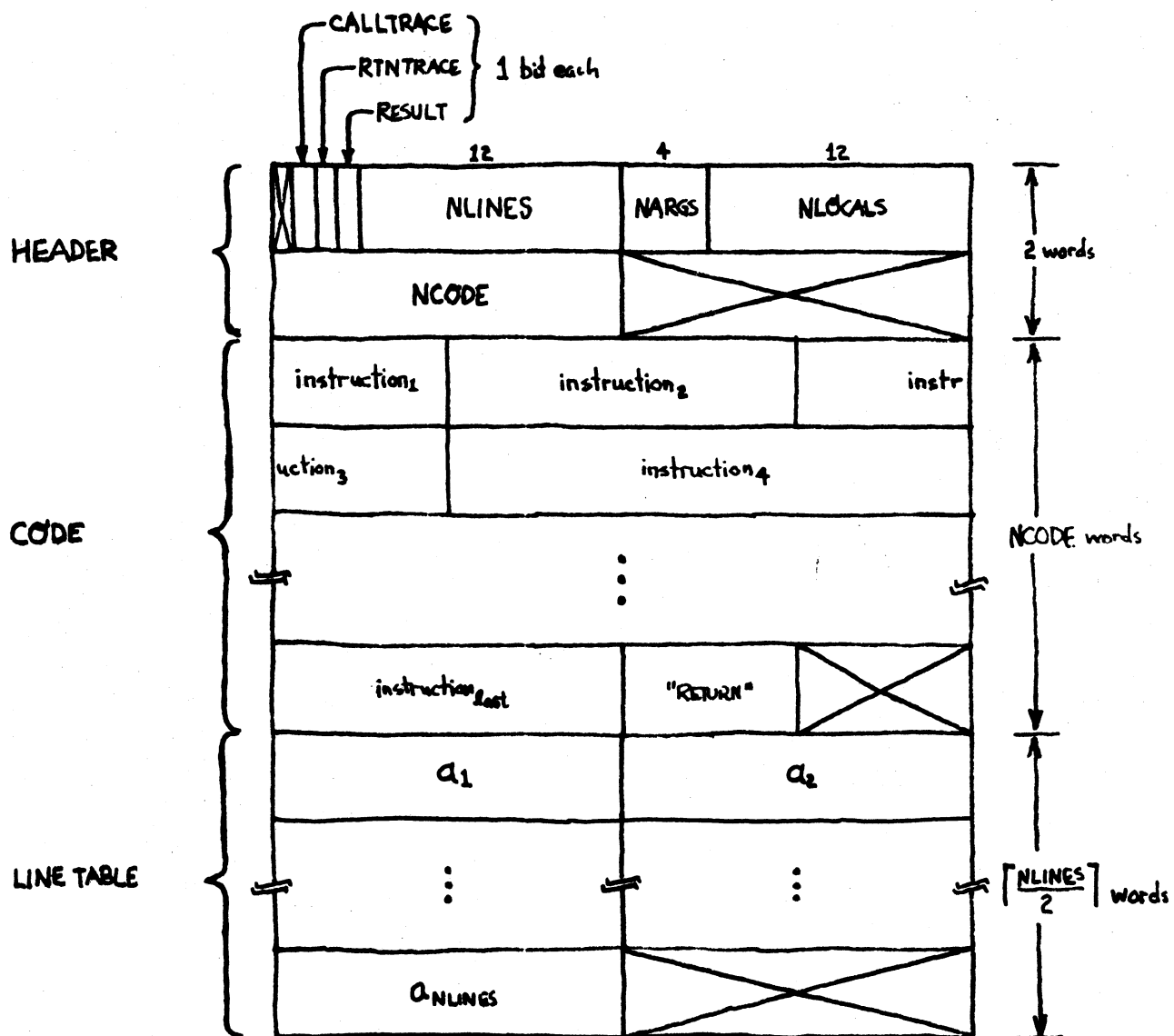
RTNTRACE--if this bit is one and the state flag FCNRTRC is also one, the processor will trap just before returning from this function;



* This field, not used by the APL processor, contains the length of the corresponding block.
 ** Directory entry $\emptyset$ corresponds to the directory block itself.

APL Procedure Segment

Figure 4.1



APL Function Block

Figure 4.2

RESULT--if this bit is one, the function returns a value (the contents of the local variable with address -1);

NLINES--the number of lines in the source function from which this function block was compiled (also the number of entries in the line table);

NARGS--the number of actual parameters with which this function should be called;

NLOCALS--the number of local variables excluding the actual parameters (including the result variable, if any) required by this function;

NCODE--the number of 32-bit words occupied by the code in this function block.

4.2.2. Code. Instructions for the APL processor are not of fixed length. Most instructions are one, two, or three bytes long. One instruction is five bytes long, and others are $2 + 4n$ bytes long (for $0 \leq n \leq 255$). Detailed descriptions of each of the instructions is the subject of Section 6, but it would be appropriate here to state that the last instruction in the code for every function should be a "RETURN" instruction (see Section 6.3). This way a function will return if the flow of control "runs off the end." Following this last instruction will be zero to three bytes of garbage to fill out the last word of code.

4.2.3. Line table. The sole use of the line table is made by the "BRANCH" instruction (i.e., APL's monadic right arrow). This table consists of a sequence of 16-bit elements (half-word aligned), ostensibly one per line of source program. Each element is the function block-relative byte address of an instruction (presumably the first instruction compiled from the corresponding source line). Note that the function block-relative byte address of the very first instruction in a code body is 8, due to the preceding two words (at four bytes per word) of header.

5. Controlling the APL Processor

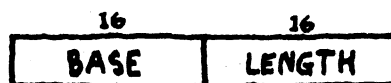
In Sections 3 and 4 a description of the static states of the data and procedure segment components of a program is given. Sections 5 and 6 are devoted to the dynamic aspects of executing a program. The effects of the individual instructions are given in Section 6; this section describes how the processor is commanded to start and stop executing a particular program.

In an over-all view of the CRMS system, the APL processor appears as an i/o device to the SIMPLE processor. Five preassigned cells in central memory and the two external communications flipflops are associated with the APL processor "device." The flipflops will be referred to as FF_{SA} and FF_{AS} , and have the following properties:

- (i) The APL processor can read (sense state = $\emptyset$ or = 1) and reset (set state to $\emptyset$) FF_{SA} , while it can set FF_{AS} (to state 1);
- (ii) The SIMPLE processor can set FF_{SA} , and can read and reset FF_{AS} .

Here are the preassigned memory cells and their usage:

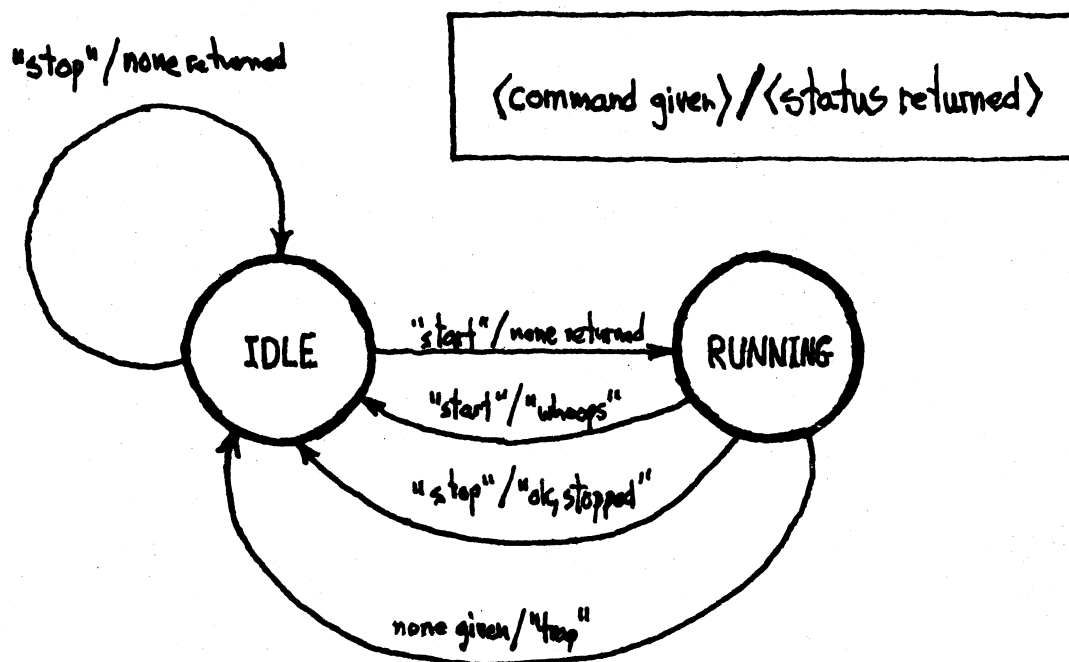
- (i) APLCOMWD (457B)--the SIMPLE processor places a command (either 1 for "start" or 0 for "stop") here for the APL processor;
- (ii) APLSTATWD (460B)--the APL processor places a status code here whenever it stops;
- (iii) APLDSEGWD, APLPSEG0WD, APLPSEG1WD (461B, 462B, 463B, respectively) --before the SIMPLE processor commands the APL processor to start, it sets up these words to describe the data segment, procedure segment, and alternate procedure segment, respectively, of the program to be run. Figure 5.1 gives the format of these "segment descriptors."



BASE is the absolute address in central memory of the first word of the segment.
LENGTH is the length, in words, of the segment.

Segment Descriptor Format

Figure 5.1



APL Processor States

Figure 5.2

The APL processor is always in one of two states, idle or running a program--see Figure 5.2. A transition from idle to running occurs only when the SIMPLE processor gives a start command. A transition from running to idle occurs either when the SIMPLE processor gives a stop command or spontaneously when the APL processor encounters a trap condition in the program being executed, whichever happens first. These start and stop commands are given in the following way: First, the SIMPLE processor sets the low order (right-most) bit of APLCOMWD: zero means stop; one means start (the other bits of APLCOMWD have no significance). Then, the processor sets FF_{SA} . Some time later the APL processor will notice that FF_{SA} is set, will reset it, and will obey the command indicated by the current setting of APLCOMWD.

Whenever the APL machine goes idle, it issues a status code. These codes are issued with a method like that used for giving start/stop commands. The APL processor stores the status code in APLSTATWD and sets FF_{AS} . Eventually the SIMPLE processor will notice that FF_{AS} is set, reset it, and then look at the current value of APLSTATWD. Here are the code values:

- 1 = "whoops" ("start" command received while running; APL processor stopped)
- 2 = "ok, stopped"
- 3 = "trap" (APL processor encountered a trap condition in program being run)

Some notes on this interface to the APL processor are in order. First of all, when the processor is running and is given a stop command, it could take as long as a few milliseconds before the shutdown occurs and the "ok, stopped" status is issued. (Sometimes the APL processor really gets wrapped up in its work!) Second, the SIMPLE processor should be sure the APL

processor is idle before it writes into the segment descriptors APLDSEGWD, APLPSEG0WD, and APLPSEG1WD. This is because the APL processor looks at these cells from time to time as long as it is running.

6. The Instructions

This section details the execution of each APL processor instruction, including operands, results, and possible traps. We begin with some general information.

6.1. Operands, Results, and Reference Counts

Instructions remove their operands from the top of the stack and replace them with their results. If an instruction takes more than one operand, then the first operand is in the stack cell addressed by LTOP, the second operand is addressed by (LTOP-1), etc.

As was mentioned in section 2.4, every array block has a reference count field giving the number of array descriptors currently pointing to that array. This field is automatically incremented and decremented by the various instructions as they create and destroy array descriptors. For example the SHAPE instruction, which takes a single operand, decrements the reference count of its operand (providing the operand is an array) before replacing the operand by the result on the stack. Whenever a reference count becomes zero, the space occupied by that array is automatically returned to the free chain in array block storage.

6.2. Traps

A trap results when the processor encounters some abnormal condition during the execution of a program. Here is a conceptual view of trap generation. Just before it executes each instruction, the processor consults an infallible oracle to find out whether execution would fail for any reason. If not, the processor executes the instruction (successfully, of course), and proceeds to the next instruction. Otherwise, the processor stops before executing the instruction. The processor sets the state

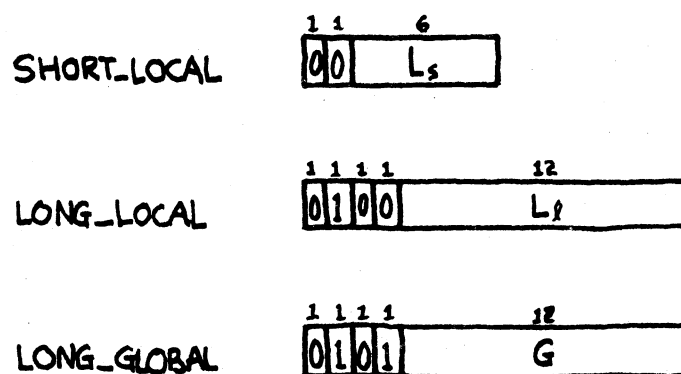
variables TRAPCLASS and TRAPNUMBER to a code describing what caused the trap, but leaves the rest of the program state untouched. (Thus PCTR points to the instruction which cannot be executed, and that instruction's operands are still on top of the stack.) Finally, the processor issues a "trap" status to the SIMPLE processor and enters its idle state, as described in section 5.

After each instruction in the following descriptions is a list of traps which could result; Appendix 2 summarizes all the traps by class and number. Note that traps of class 3 (a trap denoted (n,m) means the trap of class n and number m) are all due to a system error (compiler and/or processor), and are not normally expected to occur!

A few traps are not explicitly listed under the individual instructions since they apply to virtually any instruction. These are the (3,1), (3,2), and (3,3) traps, signifying the processor encountered a bad (too large) address supposedly pointing into the procedure segment, the runtime stack, and the rest of the data segment, respectively. The other special trap is the step (2,1) trap, which is intended as a debugging aid. It occurs when the processor is in step trace mode (i.e., the state flag STEP is equal to one) and completes the execution of an instruction. In this case it is the next instruction which is considered to have generated the trap (so, for instance, the PCTR points to this next instruction).

6.3. The Individual Instructions

Following are the actual descriptions of the instructions. (Appendix 1 lists all the instructions in alphabetical order.)



L_s and L_l are local addresses: negative integers in two's complement notation and truncated to six and twelve bits, respectively. Thus $-64 \leq L_s \leq -1$ and $-4096 \leq L_l \leq -1$ are satisfied.

G is a global address: a twelve-bit unsigned integer, which should be larger than seven.

The FETCH Family

Figure 6.1

The FETCH FAMILY

(SHORT_LOCAL, LONG_LOCAL, and LONG_GLOBAL)

Execution of a FETCH instruction pushes a copy of the descriptor which is the value of a designated variable onto the runtime stack. The first two forms, SHORT_LOCAL and LONG_LOCAL, always designate a local variable (or formal parameter variable) of the current function activation; a LONG_GLOBAL instruction always refers to a global variable. A local variable is addressed by a negative integer, which the processor interprets relative to the value of the state variable LBASE. As shown in Figure 6.1, this negative integer, in two's complement representation and truncated to six or twelve bits, is explicitly contained in the instruction, SHORT_LOCAL or LONG_LOCAL, respectively. A global variable is addressed by an integer larger than seven, which the processor interprets relative to the start of the data segment. (Addresses zero through seven are invalid since they correspond to the words of the processor state area.) Each LONG_GLOBAL instruction has a twelve-bit field holding the address of the variable to be affected.

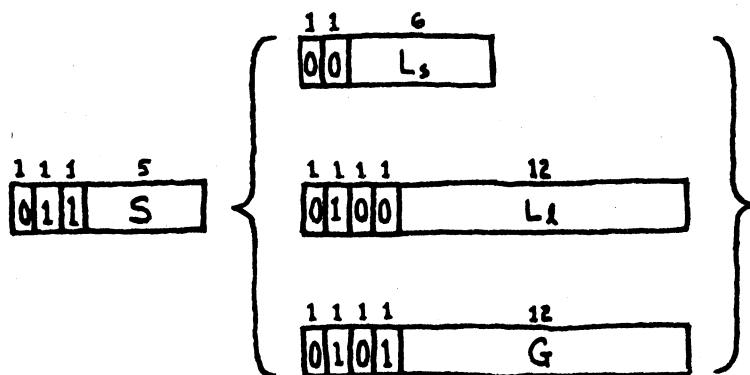
As an aid in implementing call-by-reference function arguments, one level of indirect addressing is possible. If the variable addressed by any FETCH instruction contains an indirect parameter word descriptor (see Section 2.5), then the instruction loads the value of the variable indirectly addressed by this descriptor. Indirect parameter word descriptors are produced by the REFERENCE instruction (q.v.).

Traps:

(3,6)--Illegal indirect chain if an indirectly addressed variable contains another indirect parameter word descriptor.

(10,0)--Value error if the variable directly or indirectly addressed by a FETCH instruction contains a descriptor for the "undefined value."

(1,1)--Runtime stack overflow if $SBASE+LTOP$ = address of last cell in data segment.



S is a suboperation code:

ASSIGN_INDEXED, rank S,	if $1 \leq S \leq 15$
ASSIGN	if $S = 16$
ASSIGN_NORESULT	if $S = 17$
REFERENCE	if $S = 18$
TEST_DEFINED	if $S = 19$

L_s , L_1 , and G are local or global addresses, as defined in Figure (SHORTLOCAL, LONGLOCAL, and LONGGLOBAL Instructions).

The MEMREF Family

Figure 6.2

The MEMREF Family

These instructions share a common format; the execution of each involves a variable designated by a field of the instruction itself. As shown in Figure 6.2, a MEMREF-family instruction is two or three bytes long. The first byte contains the opcode for the MEMREF family together with a subopcode selecting a member of the family. The last byte or pair of bytes addresses the variable to be affected; this address takes a form identical to the SHORT_LOCAL, LONG_LOCAL, or LONG_GLOBAL instruction which would be used to load the value of the same variable onto the stack.

One level of indirect addressing is provided by the MEMREF instructions, just as with the FETCH instructions. If the variable directly addressed by a MEMREF instruction contains an indirect parameter word descriptor, then the instruction (ASSIGN, etc.) applies to the variable indirectly addressed by this descriptor.

Traps from MEMREF instructions (see also descriptions of the individual instructions):

(3,7)--Malformed address field.

(3,6)--Illegal indirect chain if an indirectly addressed variable contains another indirect parameter word descriptor.

(3,8)--Undefined suboperation code.

ASSIGN_INDEXED

The first sixteen of the thirty-two possible MEMREF suboperation codes all stand for the ASSIGN_INDEXED operation, which is compiled for the APL assignment operator when the variable on the left-hand side is subscripted. Let RANK be the value of the suboperation code of the instruction, $1 \leq \text{RANK} \leq 15$. Then the instruction takes (RANK+1) operands. The first

RANK operands are the subscripts, which are limited to scalars. The last operand is the value computed by the code compiled from the right-hand side of the assignment, and must also be scalar. The variable addressed by the ASSIGN_INDEXED instruction must hold a descriptor for an array of rank RANK. If S_i is the value of the i^{th} subscript operand and L_i is the length of the i^{th} coordinate of this array, then $0 \leq (S_i - \text{IORG}) < L_i$ must hold for $1 \leq i \leq \text{RANK}$. The execution of this instruction substitutes the value of its last operand for the element in the array indexed by

$$\sum_{i=1}^{\text{RANK}} \left[(S_i - \text{IORG}) \cdot \prod_{j=i+1}^{\text{RANK}} L_j \right] \cdot$$

ASSIGN_INDEXED returns on the stack the value of its last operand.

ASSIGN_INDEXED is the only instruction that conditionally copies an array to assure a reference count of one.

If any of the operands is an array or the "undefined value," alternate procedure function number $32 + \text{RANK}$ is automatically called. This function is assumed to (1) expect $\text{RANK} + 2$ parameters and (2) return onto the stack the value of the last parameter.

Traps:

(3,4)--Missing operand(s) if $\text{LTOP} < (\text{LBASE} + \text{RANK} + 2)$.

(3,5)--Unknown descriptor type if the value of the addressed variable or of any of the operands is not a scalar or an array.

(5,0)--Rank error if the value of the addressed variable is a scalar.

(7,0)--Type error if any of the subscripts (first RANK operands) is a character.

(8,5)--Domain error if any of the subscripts is a nonintegral floating-point number.

(9,1)--Wrong number of subscripts if $RANK \neq$ rank of the array which is the value of the addressed variable (unless the addressed variable is a scalar).

(9,2)--Index error if $0 \leq (S_i - IORG) < L_i$ fails to hold for some i , $1 \leq i \leq RANK$.

ASSIGN, ASSIGN_NORESULT

These two instructions perform assignment to an unsubscripted variable. Each takes a single operand from the stack, the value to be assigned. Execution of an assignment stores this value into the variable directly or indirectly addressed by the instruction (see The MEMREF Family). ASSIGN also returns on top of the stack the value assigned to the variable. ASSIGN_NORESULT, as its name implies, returns no value. It is equivalent to an ASSIGN instruction followed by an EAT1 instruction, and could be compiled for an assignment which is not nested in a larger expression.

Traps:

(3,4)--Missing operand if $LTOP < (LBASE + 2)$.

REFERENCE

This function returns on the stack an indirect parameter word descriptor for the variable it directly or indirectly addresses. It is compiled for a variable reference used as an actual parameter to a user-defined function specifying pass-by-reference for the corresponding

formal parameter. The indirect parameter word descriptor (See section 2.5.) contains one bit to designate a local or global variable, and a sixteen-bit address field. This address is just a data segment address for a global variable, and is an SBASE-relative address for a local variable. (Unlike LBASE-relative addresses, which are always negative numbers, SBASE-relative addresses are always positive.) REFERENCE takes no operands from the stack.

Traps:

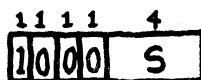
(3,1)--Runtime stack overflow if (SBASE+LTOP) = address of last word in data segment.

TEST_DEFINED

This instruction returns an integer-type value of zero or one depending on whether the variable it directly or indirectly addresses contains or does not contain a descriptor for the "undefined value." TEST_DEFINED takes no operands from the stack.

Traps:

(3,1)--Runtime stack overflow if (SBASE+LTOP = address of last word in data segment.



S is a suboperation code:

	S	operator
monadic	0	IDENTITY
	1	NEGATIVE
	2	FLOOR
	3	CEILING
	4	MAGNITUDE
	5	NOT
dyadic	6	COEFBEXP
	7	SUM
	8	DIFFERENCE
	9	PRODUCT
	10	QUOTIENT
	11	LOGICAL_PRODUCT
	12	LOGICAL_SUM
	13	LESS
	14	EQUAL*
	15	INDICATE\TRANSLATE*

* These operators accept character operands.

The GENSCALAR Family

Figure 6.3

The GENSCALAR Family

The instructions of this family (see Figure 6.3) correspond to (a subset of) the APL "primitive scalar functions." Some are monadic, and the rest are dyadic. Operands of any rank are allowed; the APL\360 rules of conformity are exactly followed. Since operands of rank greater than zero are treated elementwise, the descriptions of the individual GENSCALAR instructions are in terms of scalar operands.

Normally the APL processor obeys stop commands from the SIMPLE processor only between instruction executions (see Section 5). The only exceptions are the GENSCALAR instructions, some of which spend as long as 15 microseconds per element of the result (which could have thousands of elements). If a command from the SIMPLE processor arrives during the execution of a GENSCALAR instruction, the APL processor sets the INTRPT state flag to one and stores some internal quantities in the state variables saved-NELTS,..., and saved-RPTR (see Figure 3.2). Then the processor stops just as it would between instructions. If the interrupted program is to be restarted later, then the "saved-XX" state variables must be preserved unmodified, as must the state flag INTRPT and the state flags marked as temporaries in Figure 3.3.

Traps from GENSCALAR instructions (also see descriptions of individual instructions):

(3,4)--Missing operands if $LTOP < LBASE + 3$ for dyadic operators, or if $LTOP < LBASE + 2$ for monadic operators

(3,5)--Unknown descriptor type if either operand is not a scalar or an array.

(5,0)--Rank error if the ranks of the operands for a dyadic operator differ and neither operand is a one-element array or scalar.

(6,0)--Length error if the ranks of the operands for a dyadic operator are the same, but the operand shapes are different and neither is a one-element array or scalar.

(1,N)--Array block storage overflow if there is insufficient free space to create the result array. N is the number of contiguous words (minus 1) that was not possible to allocate.

(3,9)--Bad freechain loop.

IDENTITY

This monadic instruction returns its (integral or floating-point) numeric operand unchanged.

Traps:

(7,0)--Type error if the operand is of character type.

NEGATIVE

This monadic instruction returns the same value as its numeric operand, except with the opposite sign, except that zero always has a positive sign.

Traps:

(7,0)--Type error if the operand is of character type.

FLOOR

This monadic instruction returns the greatest integer not larger than its numeric operand. The result is of integer type unless the operand is a floating-point number greater than or equal to 2^{23} in magnitude, in which case only a floating-point representation is possible.

Traps:

(7,0)--Type error if the operand is of character type.

CEILING

This monadic instruction returns the least integer not smaller than its numeric operand. The result is of integer type unless the operand is a floating-point number greater than or equal to 2^{23} in magnitude, in which case only a floating-point representation is possible.

Traps:

(7,0)--Type error if the operand is of character type.

MAGNITUDE

This monadic instruction returns the same value as its numeric operand, except with positive sign.

Traps:

(7,0)--Type error if the operand is of character type.

NOT

The single operand for this instruction must be logical, i.e., equal to zero or one (either integer or floating-point). The result is equal to one minus the operand, and is always of integer type.

Traps:

(7,0)--Type error if the operand is of character type.

(8,5)--Domain error if the operand is a number other than zero or one.

COEFBEXP

This dyadic instruction performs three operations depending on the

second operand. If the second operand is

- (1) negative, the result is in the range $(-1, -0.5]$ union $[0.5, 1)$.

The result is defined as the first operand multiplied by an integral power of 2. This operation is called `EXTRACT_COEFFICIENT`.

- (2) zero, the result is an integer between 1 and 256, inclusively.

The result is defined as 128 plus the 2's power exponent of first operand. This operation is called `EXTRACT_BIASED_EXPONENT`.

- (3) between 1 and 256, inclusively, the pseudo-first operand is in the range $(-1, -0.5]$ union $[0.5, 1)$. The pseudo-first operand is the first operand multiplied by an integral power of 2. The result is defined as 2^{-128} multiplied by the pseudo-first operand, multiplied by 2 to the second operand. This operation is called `CREATE_FLOATING_NUMBER`.

The result is undefined if the second operand is greater than 256.

Traps:

(7,0)--Type error if either operand is character.

(8,5)--Domain error if the first operand is zero or if the second operand is of type floating point.

`SUM`, `DIFFERENCE`, `PRODUCT`

These dyadic instructions return, respectively, the sum, difference, or product of two numeric operands. The result is of integer type if the result is zero or if the operands are of integer type and the magnitude of the result is less than 2^{23} . In every other case the result is of floating-point type.

Traps:

(7,0)--Type error if either of the operands is of character type.

(8,2)--Floating-point underflow if UFMODE is turned on and the magnitude of the true result is less than 2^{-128} (but is greater than zero).

(8,3)--Floating-point overflow if the magnitude of the result is not less than 2^{128} .

QUOTIENT

This dyadic instruction returns the quotient of its first operand divided by its second operand, both of which must be numeric. The result is always of floating-point type or zero. Zero divided by zero is one.

Traps:

(7,0)--Type error if either operand is of character type.

(8,2)--Floating-point exponent underflow if UFMODE is turned on and the magnitude of the result is less than 2^{-128} (but is greater than zero).

(8,3)--Floating-point exponent overflow if the magnitude of the result is not less than 2^{128} .

(8,4)--Zero divisor if the second operand is zero and the first is not.

LOGICAL_PRODUCT, LOGICAL_SUM

These dyadic instructions return, respectively, the logical conjunction and disjunction of their operands, as shown in the truth table below. Each operand must be either zero or one (integer or floating point); the result

will always be of integer type.

A	B	A L_PRODUCT B	A L_SUM B
0	0	0	0
0	1	0	1
1	0	0	1
1	1	1	1

Traps:

(7,0)--Type error if either operand is of character type.

(8,5)--Domain error if either operand is a number other than zero or one.

LESS

This dyadic instruction returns one if its first operand is smaller than its second operand, and returns zero otherwise; the result is always of integer type.

Traps:

(7,0)--Type error if either operand is of character type.

(8,2)--Floating-point underflow if UFMODE is turned on and the magnitude of the difference of the operands is less than 2^{-128} but greater than zero.

EQUAL

This dyadic instruction returns one if its operands are equal, and zero otherwise. The result is always of integer type, but the operands may be any scalars (character, integer, or floating point). The definition of equality used is as follows: A character is equal only to another character

possessing the same character code. Two numbers are equal if their difference (as would be returned by the DIFFERENCE instruction) is zero.

Traps:

(8,2)--Floating-point underflow if both operands are floating point, UFMODE is turned on, and the magnitude of their difference is between zero and 2^{-128} , exclusive.

INDICATE\TRANSLATE

This dyadic instruction implements one of two functions depending on the value of the second operand.

- (1) If the second operand is a negative number, then the result is one for a numeric (integer or floating-point) first operand and zero for a character first operand.

Traps: None.

- (2) Otherwise, the result is a translation of the operand: If the first operand is a character, the result is the corresponding integer between zero and 255 inclusive. And if the first operand is an integral number between zero and 255, the result is the character-type value possessing that code.

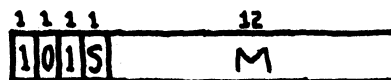
Traps:

(8,5)--Domain error if the operand is a number which is nonintegral, negative, or greater than 255.



The INDEX Instruction

Figure 6.4



S is the sign (1 $\Rightarrow$ negative, 0 $\Rightarrow$ nonnegative).

M is the magnitude, an unsigned integer.

The SHORT_CONSTANT Instruction

Figure 6.5

INDEX

This instruction corresponds to the indexing operation of APL\360 restricted to scalar subscripts. The first operand is the array to be indexed; its rank must be equal to the value of the RANK field of the INDEX instruction itself (see Figure 6.4). The rest of the operands are the integer or integral floating-point subscripts in left-to-right order. If S_i is the i^{th} subscript and L_i is the length of the i^{th} coordinate of the first operand, then $0 \leq (S_i - \text{IORG}) < L_i$ must hold for $1 \leq i \leq \text{RANK}$. The result returned by INDEX is the value of the ravel of its first operand indexed by

$$\sum_{i=1}^{\text{RANK}} \left[(S_i - \text{IORG}) \cdot \prod_{j=i+1}^{\text{RANK}} L_j \right] .$$

If any operand besides the first is an array or the "undefined value," alternate procedure function number $16 + \text{RANK}$ is automatically called. This function is assumed to (1) expect $\text{RANK} + 1$ parameters and (2) return a result onto the stack. (Note that RANK is always between 1 and 15, inclusively. $\text{RANK} = 0$ always results in trap (5,0).)

Traps:

(3,4)--Missing operand(s) if $\text{LTOP} < (\text{LBASE} + \text{RANK} + 2)$.

(3,5)--Unknown descriptor type if the first operand is not a scalar or an array, or if any other operand is not a scalar or an array or the "undefined value."

(5,0)--Rank error if the first operand is a scalar.

(7,0)--Type error if any subscript is a character.

(8,5)--Domain error if one of the subscript operands is a nonintegral floating-point number.

(9,1)--Wrong number of subscripts if $RANK \neq$ rank of first operand, unless the first operand is scalar.

(9,2)--Index error if $0 \leq (S_i - IORG) < L_i$ fails to hold for some i , $1 \leq i \leq RANK$.

SHORT_CONSTANT

This instruction takes no operands from the stack, and places one integer-type result on the stack. The sign and magnitude of the result, say S and M , respectively, together constitute the 13 right-most bits of the SHORT_CONSTANT instruction itself (see Figure 6.5).

Traps:

(1,1)--Runtime stack overflow if room for one more cell on the stack does not exist.



S is a suboperation code, and E is extra information present for some values of S:

S	operator	E present ?
0	UNDEFINED	No
1	EAT1	
2	INTERCHANGE	
3	SETORIGIN	
4	GETORIGIN	
5	SHAPE	
6	RESHAPE	
7	RAVEL	
8	CATENATE	
9	INDEX_GENERATOR	Yes, see Figure 6.7
10	CONSTANT_SCALAR	
11	CONSTANT_VECTOR	Yes, see Figure 6.8
12	BRANCH	No
13	GO	Yes, see Figure 6.9
14	GO_TRUE	
15	GO_FALSE	No
16	RETURN	
17	CALL	Yes, see Figure 6.10
18	NO_OPERATION	No
19	BREAKPOINT	
20	ATTENTION	Yes, see Figure 6.8.5
21	CHAR_VECTOR	

The ETC Family

Figure 6.6

The ETC Family

The instructions in this family (see list in Figure 6.6) are not actually related, except that the first two bits of the opcode of all of them are the same.

Traps (see also the descriptions of the individual instructions):

(3,11)--Undefined suboperation code in ETC-class instruction.

UNDEFINED

This instruction takes no operands. Its sole effect is to place a descriptor for the "undefined value" on top of the stack.

Traps:

(1,1)--Runtime stack overflow if room for one more cell on the stack does not exist, i.e., if $SBASE+LTOP$ = address of last word in data segment.

EAT1

This instruction takes one operand, of any type, and produces no result. It is useful at the end of a statement to remove the value of the statement from the stack.

Traps:

(3,4)--Missing operand if $LTOP < LBASE+2$.

INTERCHANGE

This instruction takes two operands, of any type, and returns two results by simply interchanging its operands on the top of the stack. The main use for INTERCHANGE is to reverse the operands to LESS to get the effect of a "GREATER" instruction. Neither operand need be a valid APL data type.

Traps:

(3,4)--Missing operand(s) if $LTOP < LBASE + 3$.

SETORIGIN

This instruction takes one operand, produces one result, and has a side effect. The operand must be either zero or one (either integer or floating-point); the state flag IORG (see section 3.1) is set accordingly. The result of this instruction is the previous value of IORG, zero or one, and is always of integer type. As a special case, the operand to SETORIGIN may be a suitable one-element array.

Traps:

(3,4)--Missing operand if $LTOP < LBASE + 2$.

(3,5)--Unknown descriptor type if the operand is not a scalar or an array.

(6,0)--Length error if the operand is an array with other than one element.

(7,0)--Type error if (the element of) the operand is a character.

(8,5)--Domain error if (the element of) the operand is a number not equal to zero or one.

GETORIGIN

This instruction takes no operands; it returns as its single result the current value of the IORG state flag, zero or one. The result is always of integer type.

Traps:

(1,1)--Runtime stack overflow if room for one more cell on the stack

is lacking.

SHAPE

This instruction corresponds exactly to the "monadic rho" operator of APL\360. Its single operand may be any scalar or array. The result is always an array of rank one whose length is equal to the rank of the operand; the integer type elements of this array are just the lengths of the coordinates of the operand. For example, SHAPE returns the empty vector when applied to any scalar.

Traps:

(3,4)--Missing operand if $LTOP < LBASE + 2$.

(3,5)--UNKNOWN descriptor type if the operand is not a scalar or an array.

(1,N>1)--Array block storage overflow if insufficient free space to create the result array exists.

(3,0)--Bad freechain loop.

RESHAPE

This instruction corresponds exactly to the "dyadic rho" operator of APL\360. The first operand is a vector of non-negative integral numbers (special case: a single such scalar) which is passed by-reference, not by-value; the result is an array with these numbers as its shape, and elements taken from the ravel of the second operand, which is passed by-value. Every floating-point element of the first operand is converted into an integer.

Traps:

(3,4)--Missing operand if $LTOP < LBASE + 3$.

(3,5)--Unknown descriptor type if either operand is other than a scalar or an array.

(5,0)--Rank error if the first operand has rank larger than one.

(7,0)--Type error if (any element of) the first operand is a character.

(8,5)--Domain error if (any element of) the first operand is a number which is nonintegral, negative, or not less than 2^{16} , or if the partial products of the elements of the first operand are not less than 2^{16} .

Partial products are computed right to left. Thus, the partial products of 2 3 4 are 24 12 4.

(6,0)--Length error if the second operand is an empty array but the shape for the result (as given by the first operand) is nonempty, or if the first operand is a vector with more than 255 elements.

(1,N>1)--Array block storage overflow if insufficient free space exists to allocate the result array, where N is the number of words (-1) that couldn't be allocated.

RAVEL

This instruction corresponds exactly to the "monadic comma" operator of APL\360. Its single operand may be scalar or array; the result is always a vector (rank one array) having the same number of elements as the operand has. The elements of the result are the same as the elements of the operand, taken in the standard APL ravel order.

Traps:

(3,4)--Missing operand if $LTOP < LBASE + 2$.

(3,5)--Unknown descriptor type if the operand is not a scalar or an array.

(1,N>1)--Array block storage overflow if insufficient free space exists to allocate the result array.

(3,9)--Bad freechain loop.

CATENTATE

This instruction corresponds to the "dyadic comma" operator of APL restricted to operands of rank zero or one. The result is always a vector (rank one array) whose length is equal to the sum of the number of elements in both operands. The elements of the result are just the elements of the first operand followed by the elements of the second operand. Note that the elements of the operands for CATENTATE may be any mixture of characters and numbers.

If either operand's rank is greater than one, alternate procedure function number 16 is automatically called. This function is assumed to (1) expect two parameters and (2) return a result onto the stack.

Traps:

(3,4)--Missing operand(s) if LTOP<LBASE+3.

(3,5)--Unknown descriptor type if either operand is other than a scalar or an array.

(6,0)--Length error if the sum of the lengths of the operands is not less than 2^{16} .

(1,N>1)--Array block storage overflow if insufficient free space exists for the result array

(3,9)--Bad freechain loop.

INDEXGEN

This instruction corresponds exactly to the "monadic iota" operator of APL. Its operand should be a nonnegative integral number or one-element array; the result is a rank one array whose length is equal to the value of the operand. The elements of the result are consecutive integers beginning with the current index origin (i.e., the value of the state flag. IORG).

Traps:

(3,4)--Missing operand if $LTOP < LBASE + 2$.

(3,5)--Unknown descriptor type if the operand is not a scalar or an array.

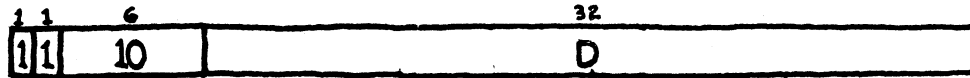
(6,0)--Length error if the operand is an array with length other than one.

(7,0)--Type error if the operand (or its element if it is a one-element array) is a character.

(8,5)--Domain error if the operand (or its element) is a number which is negative, nonintegral, or not less than 2^{16} .

(1, $N > 1$)--Array block storage overflow if sufficient free space is lacking; where N is the size of the block (minus 1) which couldn't be allocated.

(3,9)--Bad freechain loop.



D is any scalar descriptor (see Figure 2.1).

The CONSTANT_SCALAR Instruction.

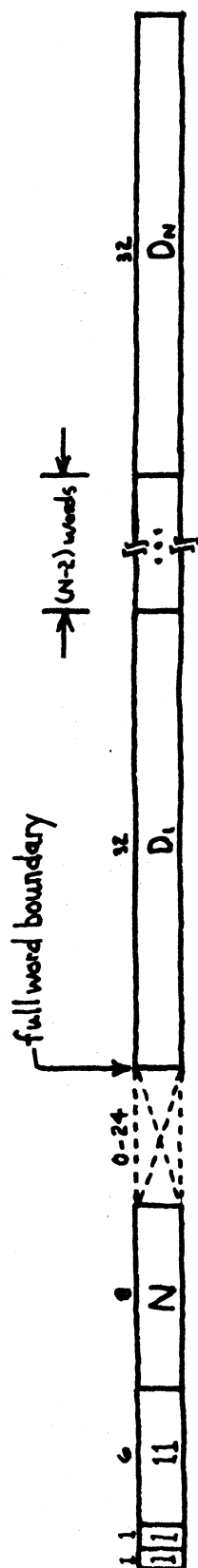
Figure 6.7

CONSTANT_SCALAR

This instruction is five bytes long (see Figure 6.7). The first byte is the opcode and the other four bytes (32 bits) constitute a descriptor for some scalar (character, integer, or floating-point) to be returned on top of the stack. (CONSTANT_SCALAR takes no operands from the stack.)

Traps:

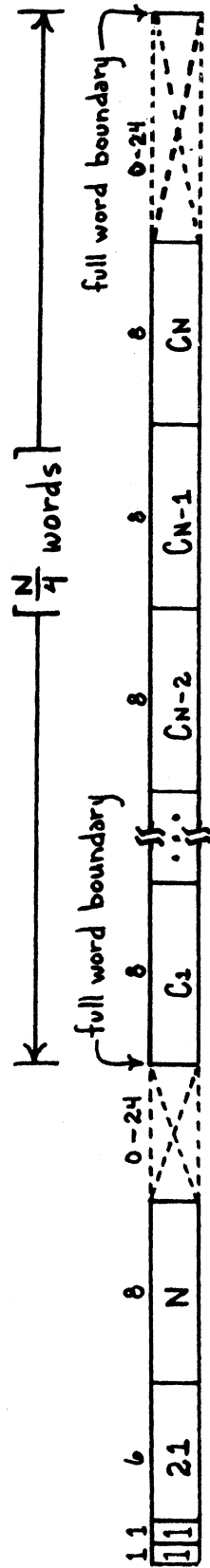
(1,1)--Runtime stack overflow if (SBASE+LTOP)= address of last word of data segment.



N is a count, $0 \leq N \leq 255$.
 D_i is any scalar descriptor, for $1 \leq i \leq N$.

The `CONSTANT_VECTOR` Instruction

Figure 6.8



N is a count, $0 \leq N \leq 255$

C_i is any character code, $1 \leq i \leq N$, $0 \leq C_i \leq 255$

The CHAR_VECTOR Instruction

Figure 6.8.5

CONSTANT_VECTOR

The CONSTANT_VECTOR format is as follows (see Figure 6.8):

- (1) The ETC-class opcode for CONSTANT_VECTOR;
- (2) A byte containing a nonnegative integer count less than 256.
- (3) Zero, one, two, or three bytes which are completely ignored to fill out the current code word;
- (4) "Count" 32-bit descriptors (of any scalar type), one to a code word.

The effect of this instruction is to return a vector (rank one array) with those scalars as elements. (CONSTANT_VECTOR takes no operands from the stack.)

Traps:

- (1,N>1)--Array block storage overflow if insufficient contiguous free space exists.
- (3,9)--Bad freechain loop.
- (1,1)--Runtime stack overflow if (SBASE+LTOP)= address of last word of data segment.

CHAR_VECTOR

This instruction is similar to CONSTANT_VECTOR (see Figure 6.8.5). But, instead of having "count" 32-bit descriptors, this instruction assumes "count" 8-bit character codes. CHAR_VECTOR expands each character code into a 32-bit character descriptor of a vector in the data segment.

Traps:

See `CONSTANT_VECTOR`

BRANCH

This instruction exactly corresponds to the "monadic right arrow" operator of APL\360. This instruction returns no result value, but has a side effect which depends on the value of its single operand:

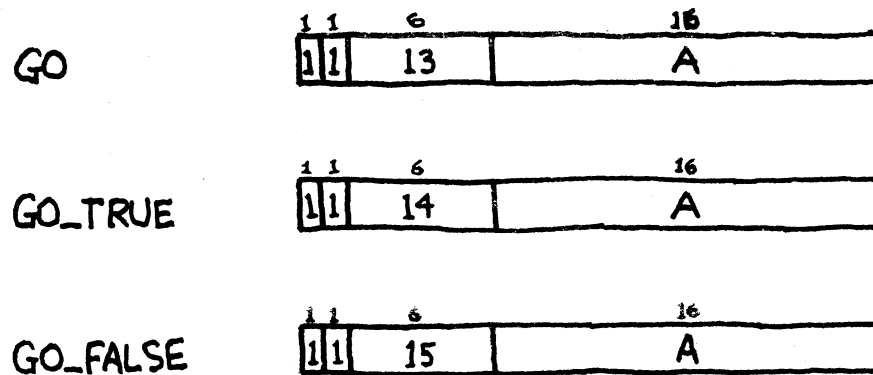
- (1) If the operand is a positive integral number (integer or floating point) not greater than the value of the current function block header's `NLINES` field, then `PCPTR` in the current stack frame is set equal to the contents of the line table entry in the current function block indexed by the value of the operand;
- (2) If the operand is an integral number outside the range mentioned in case (1) above, the processor returns from the current function just as if it were executing a `RETURN` instruction (q.v.);
- (3) If the operand is a nonempty vector the first element of which satisfies case (1) or case (2) above, then the action specified for that case is taken;
- (4) Else if the operand is an empty vector, the processor takes no special action.

The `BRANCH` operator does not return a value in the usual sense that a call to it could be nested within an expression (but of course in case (2) above the current function may return a value).

Traps:

(3,4)--Missing operand is `LTOP<LBASE+2`.

(3,5)--Unknown descriptor type if the operand is not a scalar or an array.



A is a byte address in the current function block of an instruction to transfer to.

GO, GO_TRUE, and GO_FALSE instructions

Figure 6.9

(5,0)--Rank error if the operand has rank greater than one.

(7,0)--Type error if the operand (or its first element if the operand is a vector) is a character.

(8,5)--Domain error if the operand (or its first element) is not integral.

(3,1)--Procedure segment address too large if PCTR is to be set to an address that lies past the end of the current procedure segment.

Also--in case (2) above, any trap listed under RETURN.

GO

This instruction does not correspond to any construct of APL\360. It has no operands or results, but sets the contents of the PCTR field in the current stack frame equal to the byte address contained in the second two bytes of the three-byte GO instruction (see Figure 6.9).

Traps:

(3,1)--Procedure segment address too large if the given byte address lies past the end of the current procedure segment.

GO_TRUE, GO_FALSE

These three-byte instructions each take a single operand, the value of which should be a numeric zero or one (either integer or floating-point). Execution of GO_TRUE or GO_FALSE is the same as for GO, except that setting PCTR is conditioned upon the value of the operand being one or zero, respectively. As a special case these instructions accept a suitable one-element array as operand; neither returns a result on the stack.

Traps:

(3,4)--Missing operand if $LTOP < LBASE + 2$.

(3,5)--Unknown descriptor type if the operand is not a scalar or an array.

(6,0)--Length error if the operand is an array with other than one element.

(7,0)--Type error if the operand (or its solitary element) is a character.

(8,5)--Domain error if the operand (or its element) is a number other than zero or one.

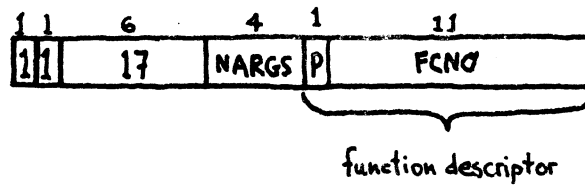
(3,1)--Procedure segment address too large if PCTR is to be set but the given byte address lies past the end of the current procedure segment.

RETURN

This instruction of no operands causes the processor to return from the current function, just as would be caused by executing a BRANCH to a nonexistent line number. The return is accomplished by resetting the state variables LBASE and LTOP to the values which were in effect when the current function was called, and additionally pushing a result value on the stack if the RESULT flag in the header of the current function is one (see Figure 4.2). The new value for LBASE is taken from the "last LBASE" field in the stack frame for the returning function activation (see Figure 3.4). LTOP is set equal to the old value of LBASE minus the quantity (NARGS+NLOCALS+1), where NARGS and NLOCALS are fields in the header of the current function. The result value, if any, is taken from the first local variable (the one with address -1) of the returning function.

Traps:

(2,3)--Return trace if both the global state flag FCNRTRC and the



NARGS is the number of arguments the caller has pushed onto the stack before executing this instruction.

P selects the procedure segment, regular or alternate, containing the function being called.

FCNO is the function number of the function being called.

The CALL instruction

Figure 6.10

RTNTRACE flag in the current function block's header are equal to one.

(3,12)--Attempt to return from bottom function activation in stack if value of "last LBASE" field in the current stack frame is not less than the current value of LBASE.

(10,0)--Value error if the function is to return a result but its result variable contains a descriptor for the "undefined value."

CALL

Execution of this instruction causes the processor to call a function, i.e., to establish a new frame on top of the stack and make it the current one, updating LBASE and LTOP. The three-byte CALL instruction takes a variable number (controlled by the NARGS field of the CALL instruction) of operands on the stack, the actual parameters to the function being called. Rather than being removed by CALL, these operands become the formal parameter variables for the new stack frame. The other local variables of this frame are automatically initialized to the "undefined value". The "function descriptor" field in the new frame is set equal to the field of the same name in the CALL instruction itself (see Figure 6.10); the PCTR field in the new frame is always set to 8, which is the byte address of the first instruction in any function block. After the call is complete, the state variable LTOP will be equal to (LBASE+1).

A function is in call trace mode if both the global state flag FCNRTRC and the CALLTRACE flag in the header of the function are equal to one. Execution of the first instruction after the call to a function in call trace mode always causes a call trace or (2,2) trap.

Traps:

(3,4)--Missing operand(s) if $(LBASE+NARGS+1) > LTOP$, where NARGS is a field of the CALL instruction.

(1,1)--Runtime stack overflow if $(SBASE+LTOP+NLOCALS+2) >$ the last address in the data segment (NLOCALS is a field in the header of the called function block).

NO_OPERATION

This instruction takes no operands, returns no results, and has no side effects. It could be used, for example, to patch over an instruction used for debugging that is no longer needed.

Traps: None.

BREAKPOINT

This instruction takes no operands, returns no results, and in fact never finishes execution. Its sole effect is to cause a "breakpoint encountered" (2,4) trap. Note that the PCTR field (in the current stack frame) is left pointing to the BREAKPOINT instruction; to enable the program to continue past the breakpoint, the PCTR must be increased by one (the length of this instruction) before the program is restarted.

Traps:

(2,4)--Breakpoint trap (unconditional).

ATTENTION

This instruction is similar to BREAKPOINT (q.v.): it takes no operands, returns no result, and always generates a trap. In this case the trap is an "attention" (4,0) trap, and is intended as part of an

escape mechanism so that APL programs can call functions running on the
SIMPLE processor.

Traps:

(4,0)--Attention trap (unconditional).

APPENDIX I

INDEX TO THE INSTRUCTIONS

<u>Instruction</u>	<u>Page Number</u>	<u>Instruction</u>	<u>Page Number</u>
ASSIGN	38	INDICATE\TRANSLATE	47
ASSIGN_INDEXED	36	INTERCHANGE	52
ASSIGN_NORESULT	38	LESS	46
ATTENTION	69	LOGICAL_PRODUCT	45
BRANCH	63	LOGICAL_SUM	45
BREAKPOINT	69	LONG_GLOBAL	33
CALL	68	LONG_LOCAL	33
CATENTATE	56	MAGNITUDE	43
CEILING	43	NEGATIVE	42
CHAR_VECTOR	62	NO_OPERATION	69
COEFBEXP	43	NOT	43
CONSTANT_SCALAR	59	PRODUCT	44
CONSTANT_VECTOR	62	QUOTIENT	45
DIFFERENCE	44	RAVEL	55
EAT1	52	REFERENCE	38
EQUAL	46	RESHAPE	54
FLOOR	42	RETURN	66
GETORIGIN	53	SETORIGIN	53
GO	65	SHAPE	54
GO_FALSE	65	SHORT_CONSTANT	50
GO_TRUE	65	SHORT_LOCAL	33
IDENTITY	42	SUM	44
INDEX	49	TEST_DEFINED	39
INDEX GENERATOR	57	UNDEFINED	52

APPENDIX II

SUMMARY OF TRAP CLASSES AND NUMBERS

<u>Class</u>	<u>Number</u>	<u>Description</u>
1		Storage overflow:
	1	Runtime stack overflow
	N	Array block storage overflow (N is the number of contiguous words (minus 1) that could not be allocated.)
2		Debug trap:
	1	Instruction completed in step mode
	2	Function call in call-trace mode
	3	Function return in return-trace mode
	4	Breakpoint encountered
3		System error (processor and/or compiler):
	1	Procedure segment address too large
	2	Stack address too large
	3	Global variable or array block address too large (or SBASE is less than 8)
	4	Missing operand(s) for instruction
	5	Unknown descriptor type for instruction operand
	6	Double indirection encountered
	7	Malformed MEMREF-family instruction
	8	Undefined MEMREF suboperation
	9	Bad freechain loop
	10	Processor state OPNO field too large on interrupt restart
	11	Undefined ETC suboperation
	12	Attempt to return from bottom function activation in stack
4		Attention trap
5		APL rank error
6		APL length error

<u>Class</u>	<u>Number</u>	<u>Description</u>
7		Type error--character operand(s) where numeric expected
8		Domain error:
	2	Floating-point exponent underflow
	3	Floating-point exponent overflow
	4	Division by zero
	5	Other domain error
9		Index error
	1	Wrong number of subscripts
	2	Subscript too large or small (APL index error)
10		APL value error

PCTR points to the next instruction for traps (2,1) and (2,2). For all other traps, PCTR points to the offending instruction.

APPENDIX III
ARRAY BLOCK STORAGE ALLOCATION ALGORITHMS

```

FUNCTION MAKE( SCALAR N )
*
* THIS FUNCTION RETURNS THE ADDRESS OF A NEW RESERVED BLOCK WITH SIZE=N.
*

    SCALAR P
    P ← ROVER
    WHILE P.SIZE < N DO
        P ← P.NEXT
        GOTO GLOBALOVERFLOW IF P=ROVER
    ENDWHILE
    (P+P.SIZE).PREVFREE ← FALSE
    ROVER ← P.NEXT
    IF P.SIZE-N ≤ MAXSLOP THEN
        P.LAST.NEXT ← P.NEXT
        P.NEXT.LAST ← P.LAST
        P.SLOP ← P.SIZE-N
        P.SIZE ← N
        P.THISFREE ← FALSE
        P.PREVFREE ← FALSE
    ELSE
        P.SIZE ← P.SIZE-N
        (P+P.SIZE-1).SIZE ← P.SIZE
        P ← P+P.SIZE
        P.SLOP ← 0
        P.SIZE ← N
        P.THISFREE ← FALSE
        P.PREVFREE ← TRUE
    ENDIF
    RETURN P
ENDFUNCTION

```

```

FUNCTION FREE( SCALAR P )
*
* FREE THE BLOCK POINTED TO BY P.
*
    SCALAR Q
    P.SIZE ← P.SIZE+P.SLOP
    Q ← P+P.SIZE
    IF P.PREVFREE=FALSE THEN
        IF Q.THISFREE=FALSE THEN
            Q.PREVFREE ← TRUE
            P.NEXT ← ROVER.NEXT
            P.LAST ← ROVER
            P.LAST.NEXT ← P
            P.NEXT.LAST ← P
        ELSE
            P.SIZE ← P.SIZE+Q.SIZE
            P.NEXT ← Q.NEXT
            P.LAST ← Q.LAST
            P.NEXT.LAST ← P
            P.LAST.NEXT ← P
            ROVER ← P
        ENDIF
        P.THISFREE ← TRUE
    ELSE
        IF Q.THISFREE=FALSE THEN
            Q.PREVFREE ← TRUE
        ELSE
            P.SIZE ← P.SIZE+Q.SIZE
            Q.LAST.NEXT ← Q.NEXT
            Q.NEXT.LAST ← Q.LAST
        ENDIF
        P ← P-(P-1).SIZE
        P.SIZE ← P.SIZE+(P+P.SIZE).SIZE
        ROVER ← P
    ENDIF
    (P+P.SIZE-1).SIZE ← P.SIZE
    RETURN
ENDFUNCTION

```

APPENDIX IV

The Handling of Exponent Underflow

The underflow condition is signalled when the true result of SUM, DIFFERENCE[†], PRODUCT or QUOTIENT has magnitude greater than zero but less than 2^{-128} . The APL machine cannot represent the true result in this case. Instead, the underflow condition causes the APL machine to consult the UFMODE bit in the FLAGS part of the APL processor state area.

When UFMODE is turned on, the APL machine immediately halts with trap (8,4), the exponent underflow trap. Otherwise, the value returned by the APL machine for SUM, DIFFERENCE, PRODUCT, QUOTIENT, LESS and EQUAL is zero. While zero is the true result for EQUAL, it is not the true result for LESS or, naturally, for SUM, DIFFERENCE, PRODUCT or QUOTIENT.

Programmer's Note: It may be desirable to know whether the underflow condition was signalled during DIFFERENCE when UFMODE is turned off. Here is one way to find out: Underflow is signalled by the DIFFERENCE $A - B$ iff $(A = B) = A - B$ is true.

[†] EQUAL and LESS signal the underflow condition iff DIFFERENCE applied to their operands would signal underflow.