TECHNICAL NOTE
No. 799

THE CORAL 66 COMPILER FOR FERRANTI ARGUS 500 COMPUTER

Author: E. Gorman

Royal Radar Establishment,
Procurement Executive,
Ministry of Defence,
Malvern, Worcs.

This document, unless subsequently declared to be unlimited, is issued for the information of such persons only as need to know its contents in the course of their official duties.

This document has been approved for public release and sale; its distribution is unlimited.

MAY 8, 1980
FOREWORD

Looking back now, the compiler described in this document can be seen as an important stepping stone in the history of computer languages in Britain. It was commissioned by Ferranti Ltd for use in industrial control and automation projects where there had been a strong resistance to high level programming languages. Much was going to depend on the quality of the first compilers, and in particular the first Coral 66 compiler for the Argus. In the event this compiler proved to be outstandingly efficient and robust, which happily prevented attention being diverted from the main objective - increased productivity. I am assured by Ferranti that this objective was realized very quickly.

That it was possible to produce the Argus 500 compiler so expeditiously was due in no small measure to the software tools available to compiler writers at RSRE - tools which were quickly exploited and sharpened by Mr Gorman. RSRE, or RRE as it was, had long learned the wisdom of applying computer methods to the production of computer software.

Languages change and Coral 66 must eventually be modified (and ultimately displaced) as users become more demanding and compiler writers discover new methods of dealing with language problems. But the techniques which were used in making the Argus 500 Coral compiler are not obsolete, and should prove of particular interest and value to the many implementers of Coral 66.

P M WOODWARD

February 1977
# CORAL 66 COMPILER FOR ARGUS 500

## Part 1 - Description

<table>
<thead>
<tr>
<th>Section</th>
<th>Description</th>
</tr>
</thead>
<tbody>
<tr>
<td>1</td>
<td>Introduction</td>
</tr>
<tr>
<td>2</td>
<td>Implementation Notes</td>
</tr>
<tr>
<td></td>
<td>Language Definition</td>
</tr>
<tr>
<td></td>
<td>Addresses</td>
</tr>
<tr>
<td></td>
<td>Data Allocation</td>
</tr>
<tr>
<td></td>
<td>Chaining and Jumps</td>
</tr>
<tr>
<td></td>
<td>Loader Interface</td>
</tr>
<tr>
<td></td>
<td>Compiler Data Structures</td>
</tr>
<tr>
<td></td>
<td>Library Procedures</td>
</tr>
<tr>
<td>3</td>
<td>Preprocessor and Macro Expander</td>
</tr>
<tr>
<td></td>
<td>Overall Description</td>
</tr>
<tr>
<td></td>
<td>Module Descriptions</td>
</tr>
<tr>
<td>4</td>
<td>Compiling Strategy</td>
</tr>
<tr>
<td>5</td>
<td>Syntax</td>
</tr>
<tr>
<td></td>
<td>Syntax Analyser</td>
</tr>
<tr>
<td></td>
<td>Syntax Rules</td>
</tr>
<tr>
<td>6</td>
<td>Module Descriptions - Procedures</td>
</tr>
<tr>
<td>7</td>
<td>Module Descriptions - Labelled Blocks</td>
</tr>
<tr>
<td>8</td>
<td>Errors</td>
</tr>
</tbody>
</table>

## Part 2 - Computer Printout

<table>
<thead>
<tr>
<th>Section</th>
<th>Description</th>
</tr>
</thead>
<tbody>
<tr>
<td>1</td>
<td>Compiler</td>
</tr>
<tr>
<td></td>
<td>Coral Program</td>
</tr>
<tr>
<td></td>
<td>Identifier Occurrences</td>
</tr>
<tr>
<td>2</td>
<td>Syntax</td>
</tr>
<tr>
<td></td>
<td>Syntax Input to SAG</td>
</tr>
<tr>
<td></td>
<td>Identifier Occurrences</td>
</tr>
</tbody>
</table>
Introduction

This documentation is a working description of the Coral 66 compiler for the ARGUS 500 produced by RRE under contract from Ferranti Ltd.

It serves not only as a guide to the actual operation of the compiler, but also as an example of the use of RRE compiler building tools and techniques. These represent the culmination of many man years of research, as also do some of the important Algorithms used within the compiler. These techniques and Algorithms have been used in other compilers produced by RRE, and details have been, or will be, made freely available.

This description is not intended to be a specification of the Compiler, nor is it intended as a guide to compiler writing for general publication. It is hoped however that it contains sufficient information for the necessary understanding of the compiler operation required for its future maintenance.

Where any doubt is raised by, or ambiguity discovered in, the description of any section of the compiler, reference should be made to the actual program of the compiler, as this uniquely determines its operation. As the compiler itself is written in Coral, this should not prove to be too difficult.

The use of the compiler to compile itself makes it to a large extent self checking. Its very modular nature enables testing to be simplified as many modules are non interacting. The central routine, about which the whole compiler is built, uses a transformed syntax which has been exhaustively checked by syntax manipulation programs. These features should enable a high degree of confidence to be placed in the correct operation of the Compiler.

Part 1 of this document contains the detailed description of the individual modules forming the Compiler. This is preceded by notes on particular aspects of the techniques and standards used, and by an overall description of the compiler strategy, quoting the syntax rules in which it is embedded.

Part 2 consists of printouts of the actual program of the compiler and of the final syntax used. Each of these is accompanied by a list of all the identifiers used, giving the names of the pages on which they occur, the number of occurrences on each page, and the total number of occurrences.
Language Definition

This implementation is based on the Official Definition of Coral 66, (May 1970) and with the exception of two features, conforms to this definition. (Type D without Recursion )

The first deviation concerns bit numbering: Ferranti standard bit numbering is used. The second of these concerns Tables. On the ARGUS 500, it is inefficient to multiply the index by the number of words per entry each time a field is referred to. Therefore only the total size of a Table is specified, all indices requiring to be implicitly multiplied by the appropriate factor. As a consequence of this, individual fields may not be preset.

In addition to the Official Definition, the following extra features are incorporated:

- Shift Operators
- Exponentiation
- VALUE Procedures
- Repeated Procedure Calls
- Overflow Tests
- Implied Tests against Zero
- LABEL addresses as operands
- SPECIAL arrays
- Tracing Facilities
- Test Compilation
- Integers of specified significance
- Alternative use of square brackets
- Page Titles
- Alternative Literal and Octal constants
- Hexadecimal constants
- Identifiers commencing with £ and %
- Macro Identifiers terminating with !
- Optional Items preceded with ?
**Addresses**

Within the compiler, and also within the loader, addresses are handled as a 24 bit word in a consistent manner. Certain types of address will however only occur in a limited context. In general an address consists of two components: a base and an index. The base is in the range 0-7 and is held in 'BITS' [3,6] of the word, and the index in 'BITS' [14,10]. The eight different values of base refer to eight conceptual areas of core as follows:

<table>
<thead>
<tr>
<th>Base</th>
<th>Area</th>
<th>Mnemonic</th>
</tr>
</thead>
<tbody>
<tr>
<td>0</td>
<td>Absolute</td>
<td>A</td>
</tr>
<tr>
<td>1</td>
<td>Program</td>
<td>P</td>
</tr>
<tr>
<td>2</td>
<td>Data</td>
<td>D</td>
</tr>
<tr>
<td>3</td>
<td>Special data</td>
<td>S</td>
</tr>
<tr>
<td>4</td>
<td>Common</td>
<td>C</td>
</tr>
<tr>
<td>5</td>
<td>Library</td>
<td>L</td>
</tr>
<tr>
<td>6</td>
<td>External</td>
<td>E</td>
</tr>
<tr>
<td>7</td>
<td>Working constants</td>
<td>W</td>
</tr>
</tbody>
</table>

These areas are referred to by the first letter of the area name, and addresses within these areas by the area letter followed by one or more octal digits for the index. In the case of bases 1-4 the actual address is given by the sum of the relevant base plus the index. In the case of base 5 (Library) the index is the reference number of the library item. The index in the case of base 6 (External) is made up of two components, 'BITS'[5,10] being the number of the external relocation base, and 'BITS'[9,15] being 'OCTAL'(400) plus the displacement about that base. Base 7 (Working constants) is used to enable the loader to optimise the allocation of storage for fixed constants at load time. Within the compiler the index is always zero.

Bases 1, 2 and 3 always refer to locations within the current segment, the other bases referring to locations outside.

The remaining spare bits are used in the compiler and the loader to carry auxiliary information in specialised cases.

The most frequent of these within the compiler is where 'BITS'[10,0] are all set to 1's. In this case the value of the index is the number of an accumulator. An address of this type is never passed to the loader as it is automatically converted to an absolute address in the procedure OUTI. The value used to set these bits is given by AHARK.

Within the loader 'BITS'[3,3] may be used to hold a relativiser setting, these bits being in the correct position for setting the NWREL and NRREL registers.
DATA ALLOCATION

All non-overlayed data declared within a program segment is allocated locations within the D area of the Segment, and is addressed relative to the start of this area. The variable used within the compiler for this allocation is DATAMAX.

Two other variables are used, in conjunction with DATAMAX, for the allocation of anonymous locations used as temporary workspace. These variables are DATASTART and DATAPTR.

A flag PRESETOK is used to indicate whether presetting is allowed, preset items being output using the transfer address DTA.

The actions taken within the compiler concerned with data allocation are as follows:

A Before the compilation of a Program segment, the flag PRESETOK is set. The variables DATASTART, DATAMAX, and DTA are all set to the starting "address" of the data area. (ie Tag=D, index=0)

B At the start of a block, the current values of DATASTART, DATAMAX and PRESETOK are placed on the stack. DATAMAX is then set to the value of DATASTART.

C As the declaration of each variable, table, and array, is processed, it is allocated the address given by the current value of DATAMAX, which is then incremented by the number of words required by the item being declared.

D After each declaration, DATASTART is set to the value of DATAMAX.

E Before the first statement of a block, the flag PRESETOK is cleared.

F Before each statement within the block, DATAPTR is set to the value of DATAMAX.

G Temporary locations, used within a statement, are allocated using DATAPTR, which is incremented by one each time. If its value exceeds that of DATAMAX, this is also incremented. Temporary locations allocated in this way may be re-used if it is known that they are vacant.

H At the end of a block, the values of DATASTART and PRESETOK are restored to their values held on the stack. DATAMAX is only reset if its stacked value exceeds its current value. As a result of this action, at the end of a Program segment, the value of DATAMAX will give the total data requirement of the segment. Similarly, after the declaration of a procedure, which is treated as a block even if it contains no declarations, the value of DATAMAX will have been incremented by the total data requirement of the procedure.

I Before a list of preset constants, the flag PRESETOK should be inspected. If it is clear, the presetting is illegal, and the list of constants must be ignored. If DTA is not equal to DATASTART, a directive is issued to the loader to skip the data transfer address forwards, and DTA is set to the value of DATASTART.

J After each preset constant, the number is rescaled to the scale required by the declaration. The number is only output if PRESETOK is set and DTA is less than DATAMAX. The value of DTA is incremented by one each time a constant is output.
Chaining and Jump Optimisation

In general, the destination of conditional and unconditional jump instructions generated by the compiler is not known at the time of their generation.

There are several solutions to this problem. One is to generate pseudo labels where the jump is not explicitly to a label (‘IF’ . . ‘THEN’ . . ‘ELSE’ and use an assembler to insert the correct addresses. This usually requires more than one assembler pass. Another method is to fill in the addresses by a further compiler pass. A third method is to use a "branch ahead" table, which records the addresses of incomplete jump instructions. When the destination is discovered, directives are output to fill in the required addresses. Unless this table is arbitrarily large it will become full at some time.

The chaining method used by the compiler has none of the above disadvantages. Consider the case of a simple forward jump to a label not yet set. The first jump instruction is output with a zero address field, and its address stored in the record for the label. If a second jump instruction to the same label is to be output, it is output with its address set to the address of the first jump, and the address of the first jump in the label record replaced with that of the second. This may be repeated an indefinite number of times without using any further storage space within the compiler or loader. This builds up a chain of jump instructions in the compiled program. When the destination is finally discovered, the compiler outputs a directive to set the addresses of the instructions on the chain to the required destination.

Further complexities arise in the single pass compilation of block structured languages. Consider the following example:

```
1 'BEGIN' 'FIXED' . .
2 L: . . . . . .
3 'BEGIN' 'FIXED' . .
4 'GOTO' L; . .
5 'BEGIN' 'FIXED' . .
6 'GOTO' L; . .
7 'GOTO' L; . .
8 'END' LEVEL 3;
9 L: . .
10 'END' LEVEL 2;
11 'GOTO' L; . .
12 'END' LEVEL 1;
```

When compiling the jump of line 4, it is not safe to assume that the jump is to the label L of line 2. In fact, although the information is not available at this point, it is a jump to the label L of line 9. Similarly, when compiling the jump of line 6, it is not safe to assume that the jump is to the label L of line 2, or even that it is the same label L referred to in line 4.

However, when compiling the jump of line 7, it is safe to assume that the reference to the label L is a reference to the same label used in line 6, because these both occur at the same block level. When the end of block level 3 is reached on line 8, it may be deduced that the label L referred to on line 4 is the same label L referred to on lines 6 & 7 because it has not been set at block level 3.
The setting of the label L on line 9 supplies the destination for the jumps of lines 4, 6, and 7. At the end of block level 2, line 10, this label goes out of scope, and any record of it may be safely discarded, because it has been set. The reference to the label L in the jump on line 11 may safely be taken to be a reference to the label of line 2 because these both occur at the same block level.

Thus, when searching for a label record, only the list pertaining to the current block level may be used.

These complexities introduced by block levels do not present any great problems when chaining is used. All cases may be handled with the following directives:

1. Set chain starting at N to current program address
2. Join chain M on to end of chain N
3. Set chain N to address A

The first directive is sufficient for non-block structured cases, and is used to set forward references within a block level. It may be seen that this directive is in fact a special case of the third.

The second directive would be used in such cases as given at line 8 of the above example. Here a chain exists for block level 3 and also for level 2 (in its simplest form). As it is desirable to discard the level 3 record, the level 3 chain is joined on to the end of the level 2 chain, using a type 2 directive, only one reference within the compiler being retained. When the label is subsequently set, in line 9, a type 1 directive is used to set this composite chain.

If however the label setting of line 9 had been omitted, when the end of block level 2 was reached in line 10, it would be discovered that the label was already set at block level 1. In this case, the third type of directive would be used to set the chain to the address of the label in line 2.

Apart from jumps generated as the result of explicit 'GOTO' statements, there will be a considerable number of jumps generated "anonymously", not to specific labels, but implied by certain language features. The most obvious are those concerned with conditional statements. A statement such as:

'IF' A=B 'THEN' C=D 'ELSE' E=F;

might generate:

- Compare A and B
- Jump if not equal
- Assign D to C
- Jump unconditionally
- Assign F to E

Here it may be seen that two forward jumps are involved, the first (conditional) jump skips over the consequence statement, and is due to be set at the commencement of the alternative statement. The second jump is used to skip over the alternative statement, and is due to be set at the commencement of the following statement.
Where a condition involves more than one comparison, the Official Definition states that the condition is to be evaluated from left to right only as far as is necessary to determine the overall truth or falsity. This implies that the condition is not to be evaluated as a Boolean expression, with only the final result being tested, but that each comparison must be tested individually, immediately after its evaluation.

Thus a statement such as:

'IF' A=B 'AND' C=D 'OR' E=F 'AND' G=H 'THEN' I=J 'ELSE' K=L;

might generate:

```
Compare A and B
Jump if not equal
Compare C and D
Jump if equal
Compare E and F
Jump if not equal
Compare G and H
Jump if not equal
Assign J to I
Jump unconditionally
Assign L to K
```

This example illustrates that, irrespective of the complexity of a conditional statement, only two chains are required. (per conditional depth). The first of these is the "skip" or "false" chain, which is used to skip over the consequence statement. The second being the "true" or "due" chain, which is used to jump directly to the consequence statement, and is due to be set at its commencement. The application of chaining to such cases should be obvious.

What is not so obvious however is how cases involving explicit conditional jumps may be compiled efficiently. As an example, a statement such as:

'IF' A=B 'THEN' 'GOTO' L 'ELSE' 'GOTO' M;

might, in the worst case, generate:

```
Compare A and B
Jump if not equal
Jump unconditionally
Jump unconditionally
Jump unconditionally
To M
```

whereas it would be preferable to generate:

```
Compare A and B
Jump if equal
Jump unconditionally
To M
```
and in a complicated case, such as:

'IF' A=B 'AND' C=D 'OR' E=F 'AND' G=H 'THEN' 'GOTO' L 'ELSE' 'GOTO' M;

to generate

<p>| | |</p>
<table>
<thead>
<tr>
<th></th>
<th></th>
</tr>
</thead>
<tbody>
<tr>
<td>Compare A and B</td>
<td></td>
</tr>
<tr>
<td>Jump if not equal</td>
<td></td>
</tr>
<tr>
<td></td>
<td></td>
</tr>
<tr>
<td>Compare C and D</td>
<td></td>
</tr>
<tr>
<td>Jump if equal</td>
<td></td>
</tr>
<tr>
<td></td>
<td></td>
</tr>
<tr>
<td>Compare E and F</td>
<td></td>
</tr>
<tr>
<td>Jump if not equal</td>
<td></td>
</tr>
<tr>
<td></td>
<td></td>
</tr>
<tr>
<td>Compare G and H</td>
<td></td>
</tr>
<tr>
<td>Jump if equal</td>
<td></td>
</tr>
<tr>
<td></td>
<td>Jump unconditionally</td>
</tr>
</tbody>
</table>

In the above example, the chains of jumps, which are set up before it is discovered that a jump to a label is required, may be set to, joined on to, or remembered as, the reference to the label at the appropriate block level.

This does not solve the problem of 'THEN' 'GOTO' optimisation. It may be inferred that it would be advantageous not to output the final test instruction on the occurrence of the symbol 'THEN', but to defer its generation. In the case of 'THEN' 'GOTO' LABEL its sense is required to be reversed, but not in the case of 'THEN' 'GOTO' SWITCH. Thus the syntactic detection of this case, although possible, becomes rather complicated.

Explicit conditions also occur following the symbol 'WHILE' in 'FOR' statements. It is interesting to note the similarity between the execution of the controlled statement if the condition is true, and the execution of the consequence statement in a conditional statement. Implicit jumps are also generated as a part of 'STEP' - 'UNTIL' elements, and also from the end of the controlled statement to the instructions to calculate and test a further value of the control variable.

Another case of implicit jumps is where procedure declarations occur at the head of a block, and require to be jumped around. Where however, this block is the outer block of a procedure, this jump may be omitted, as the entry point of the procedure has not yet been reached. In the Argus implementation, the location reserved to hold the pointer to the first executable instruction of a procedure needs to be set when this point is reached.

In the case of typed procedures, it is desirable to check that the procedure has at least one 'ANSWER' statement, and that a return to the point of call is only made by means of such a statement. The return to the point of call by obeying the implicit return instruction at the end of the procedure body is only applicable in the case of untyped procedures. In the Argus implementation, a single return instruction may be generated after each 'ANSWER' statement, except where 'PROCEDURE' 'TRACE' is required; in which case all 'ANSWER' statements except one occurring as the last statement of the body require to be followed by a jump to the tracing instructions generated at the end of the procedure.
The scheme used by this compiler requires two chains at each conditional level, referred to by the variables SKIPCHAIN and DUECHAIN within the compiler. It also makes use of a variable, STATUS, which indicates the context in which a statement occurs.

At the start of each statement, DUECHAIN holds a chain of jumps, which may in fact be null, due to be set before the start of the compilation of the statement; SKIPCHAIN holding a chain of jumps bypassing the statement. At this point, the values assigned to STATUS indicate the contexts set out below.

STATUS=0

Control is passed normally from the preceding statement. If DUECHAIN is non zero, the statement is also entered by means of a jump. (The contents of the accumulators being undefined)

STATUS=1

Control is not passed normally from the preceding statement. This statement can only be entered by a jump. If DUECHAIN is zero, the statement cannot be entered.

STATUS=2

The preceding statement was an 'ANSWER' statement, and requires either an exit instruction or jump to the end of the procedure (using EXIT) depending on whether trace is required. If DUECHAIN is zero, the statement cannot be entered.

STATUS=3

The preceding symbol was 'ELSE', and an unconditional jump around this statement is required before DUECHAIN (which will be non zero) can be set.

STATUS=4

The preceding symbol was 'THEN', (or 'THEN' 'ELSE' together) and a conditional jump around this statement is required before DUECHAIN (which may be zero) can be set.

At the beginning of the compilation of a block which does not form the body of a procedure, both STATUS and DUECHAIN will be zero, it being assumed that such blocks are actually entered at their head.

If a procedure declaration occurs whilst STATUS is zero, an unconditional jump, with an initially zero address part, is output, its location being recorded in DUECHAIN. This is the instruction to skip over procedure declarations. Its use could be avoided, but only at the cost of completely re-ordering a segment, which is an extremely complicated operation with a single pass compiler.

At the start of the procedure declaration, after the skip over instruction has been output if required, the current value of DUECHAIN is stacked. STATUS is then set to 1, and DUECHAIN set to the location of the pointer to the procedure. (In the case of Library procedures, to the Library "address", with a sign marker for the convenience of the loader.) After the declarations, if any, at the head of the procedure, DUECHAIN is set, STATUS is set to zero, and the instructions to store the link and parameters generated as required.
At the end of the procedure declaration, DUECHAIN is restored to its stacked value, and STATUS is set to 1. When the actual start of the block is found, DUECHAIN will be set, thus completing the skip over instruction.

At the start of each statement, the value of STATUS is required to be inspected, and the chain DUECHAIN is required to be set. Except in the case of statements of the form 'GOTO' Label, this is carried out by the procedure STATUS TEST, which is called by the compiling actions STATUSCHECK and STATUSCODE. This procedure outputs the required jump instruction if the value of STATUS is greater than one, and resets STATUS to zero. If DUECHAIN is non-zero, STATUSTEST sets the chain to the current program address, and resets DUECHAIN to zero.

Most statements will leave STATUS set to zero. Answer statements will however set it to 2, and most cases of 'FOR' and 'GOTO' statements will set it to 1.

A conditional statement will initially set the value of STATUS to 4. The current value of SKIPCHAIN is stacked, and SKIPCHAIN is set to zero in readiness for the following condition. (DUECHAIN will at this point be zero.) After each comparison (or overflow test), FUNCTION and ACCUMULATOR are set to the appropriate values for a conditional jump instruction, the jump occurring if the "result" of the comparison is false.

If the comparison is followed by the symbol 'AND', this jump is output with its address on the SKIPCHAIN. If however the comparison is followed by the symbol 'OR', the type of the test is reversed, the instruction output with its address on the DUECHAIN, the SKIPCHAIN set to the following comparison, and SKIPCHAIN cleared.

Following the symbol 'THEN', the consequence statement will be processed. Provided that this statement is neither null, nor of the form 'GOTO' Label, this will present STATUSTEST with a value of STATUS of 4. In this case the conditional jump instruction will be output with its address on the SKIPCHAIN before DUECHAIN is inspected.

If however, the consequence statement is of the form 'GOTO' Label, the type of the test instruction is reversed. This is then output, using the chain for the label as its address, DUECHAIN being linked onto the end of this chain (or set, as appropriate). This case leaves DUECHAIN clear, and STATUS set to 5.

If the symbol 'ELSE' follows, preparation is made for the alternative statement. The values of SKIPCHAIN and DUECHAIN are exchanged, thus the skip over the consequence statement being due to be set at the start of the alternative statement. If STATUS is zero, it is set to 3 to indicate that a jump to skip over the alternative statement will be required. If however the value of STATUS is 5, this jump is not required, and STATUS is set to zero. If the value of STATUS is 4, the consequence statement was null, the type of the test instruction is reversed, and STATUS left set to 4.

The alternative statement will usually present STATUSTEST with a value of STATUS of 3. In this case, an unconditional jump instruction, with its address on the SKIPCHAIN, will be output. If however, the alternative statement is of the form 'GOTO' Label and STATUS is 3, no instructions will be output. In this case the DUECHAIN is set to the address of, or added onto the chain for, the label; STATUS being reset to zero.
At the end of the conditional statement, the value of STATUS will only be 4 if the consequence and alternative statements are both void. The test instruction is output only if it is an overflow test. If the value of STATUS is greater than 2, STATUS is set to zero. The SKIPCHAIN is then joined on to the end of the DUECHAIN, and SKIPCHAIN reset to its stacked value.

'FOR' statements also stack the value of SKIPCHAIN, this chain being used in conditions and tests within the for-list. Except in the case where the for-list is a single step-until element consisting only of three constants, the SKIPCHAIN is used to exit from each for element in turn, as it becomes exhausted, and finally from the for statement.

At the end of the for statement, the value of SKIPCHAIN is assigned to DUECHAIN, and SKIPCHAIN reset to its stacked value. Where the simple step-until case has been detected, STATUS will be left set to zero. Otherwise it will be set to 1, except in certain cases where it will be left set to 2.

( 'FOR' . . (,) . . 'WHILE' . . 'DO' . . 'ANSWER' . . )

The cases of a statement of the form 'GOTO' Label occurring as consequence and alternative statements has been mentioned above. Where a statement of this form occurs, and STATUS is 1 or 2, no instruction is output, the DUECHAIN being set to the address of, or joined on to the chain of, the label. Where STATUS is zero, an unconditional jump to the label is output, and STATUS is set to 1.

The compilation of a 'GOTO' Switch statement is always preceded by the execution of STATUSCHECK. After the required instructions have been output, STATUS is set to 1.

At the end of the body of an untyped procedure, a return instruction is output unless STATUS is 1 and DUECHAIN is zero. In this case the procedure has been left by means of a goto statement, and a return to the point of call is not required.

A typed procedure must have at least one 'ANSWER' statement. This will leave STATUS set to 2 and/or EXITCH set non-zero. If, at the end of the procedure, DUECHAIN is non-zero or STATUS is zero, an attempt is being made to return to the point of call other than by an 'ANSWER' statement.

At the end of a Program segment, a final stop instruction is required, unless DUECHAIN is zero and STATUS is 1. In this case the program is effectively terminated by a jump.

Conditional expressions require the values of both SKIPCHAIN and DUECHAIN to be stacked. STATUS optimisation is not applicable in this case, so the value of STATUS is left unchanged.
The Compiler-Loader Interface

The output of the compiler is a form of relocatable binary. In the basic form of the compiler, this is punched on paper tape. The unit of information output by the compiler ('PROCEDURE OUT5') is a five character group, each character requiring six bits. These characters are output as if they were legible characters, and so the standard internal-external code conversion routines may be used. The first character of the group is a directive or tag (T) character, the remaining four characters being treated as a 24 bit word, (w).

Where the tag indicates that the word (w) is to be stored as part of the current segment, the three most significant bits of the tag (T) specify with which transfer address the word is to be loaded, and the three least significant bits specify with which base the word is to be relocated. It should be noted that Program instructions are initially stored with the address in the least significant fourteen bits, and subsequently given ten places of left cyclic shift.

Most of the tag values used for directives require more than one word to specify the action required. Thus the requirement for a multi-length group arises. This is satisfied by allocating tag 0 as a (prefix) continuation tag. When a continuation tag is read the current word (w) is stored in an array, and the index used in referring to this array incremented, thus variable length directives may be used. This is required in the cases where a string is passed as part of a directive. The first location of this array is referred to as C, and the subsequent locations as C1, C2 etc.

Where an address is specified as part of a directive, the form of the address is exactly the same as that used within the compiler. Before use, the address is relocated with the base specified by bits 6-8 of the word.

List of Tags

0  Continuation, store W in Cn, m-n+1
1  Size Block, W=checksum, C1=Program, C2=Data, C3=Special
2  Display message, W=DISP, C=string
3  Start of common check, W=common checksum
4  Library Variable Block, W=checksum, C1=Lib no, C2=Value, C3=Identifier
5  Library Check, W=LCHQ
6  Entry Block, W=ENTRY
7  Stop Block, W=STOP
8  Store Program, relocate as Absolute
9  Store Program, relocate as Program
10 Store Program, relocate as Data
11 Store Program, relocate as Special data
12 Store Program, relocate as Common
13 Store Program, relocate as Library
14 Store Program, relocate as External
15 Store Program, relocate as Working constant, C=constant
16 Store Data, relocate as absolute
17 Set chain W to current program address unless W ≠ ve, in which case set entry point for library procedure W.
18 Set chain W to address C
19 Store Data, relocate as Special data
20 Join chain W on to end of chain C
21 Set common W to procedure C
22 Set common W to switch C
23 Set common W to label C
24 Store Special data, relocate as Absolute
25 Store Special data, relocate as Program
26  Store Special data, relocate as Data
27  Store Special data, relocate as Special data
28  Store Special data, relocate as Common
29  Store Special data, relocate as Library
30  Store Special data, relocate as External
31  Skip Data transfer address forward to W
32  spare
33  Print address W and Identifier string C
34  Print instruction W and comment string C
35  Change instruction W to 4 JCS
36-38 spare
39  End of segment, W=segment checksum. (C=Common checksum)
40-55 spare
56  Check Common location W, C=Identifier string
57-60 spare
61  Program Segment, W=Common checksum, C=Identifier string
62  Library Segment, W=Library number, C=Identifier string
63  Common Segment, W=0, C=(COMMON)

List of Compiler Sections Generating Tags

| 0  | OUTCONT  |
| 1  | FINISHSEG |
| 3  | COMOFF   |
| 4  | SETLIBVAR |
| 7  | TTAIL    |
| 8-16 | OUT24   |
| 17 | SETCHAINOPTA |
| 18 | JOINCHAINS & ENDPOR |
| 19 | OUT24    |
| 20 | JOINCHAINS |
| 21-23 | ENDPOR  |
| 24-30 | OUT24   |
| 31 | SKPDTA   |
| 33 | DIAG     |
| 34 | OUTI     |
| 35 | SETRT    |
| 39 | ENDS   |
| 56 | COMOFF   |
| 61-63 | STARTSEG |

Size Blocks (Tag 1)

The segment size block, which although only output after the segment has been compiled, must be read by the loader before the header of the segment. The format of a size block is as follows:

<table>
<thead>
<tr>
<th>W</th>
<th>Checksum for block</th>
</tr>
</thead>
<tbody>
<tr>
<td>C</td>
<td>Program Size</td>
</tr>
<tr>
<td>C₁</td>
<td>Data Size</td>
</tr>
<tr>
<td>C₂</td>
<td>Special Data Size</td>
</tr>
</tbody>
</table>

A size block must be followed by a segment header block (tags 61-63).
Program Header Block (Tag 61)

W  Common checksum
C  Segment Identifier String

Only if a segment has been compiled as a single segment program, without a Common communicator having been read, will the word W be zero. If the word W is non zero, it must be identical to the Common checksum of the previously loaded Common segment.

Library Procedure Header Block (Tag=62)

W  Library reference number of Procedure
C  Procedure Identifier String

If there is an outstanding requirement for the procedure, the segment should be loaded. Otherwise it should be ignored, by scanning forward until an end of segment tag (39) is read.

Common Segment Header Block (Tag=63)

W  Zero
C  "(COMMON)"

The Common checksum must be zero when the Common segment is loaded, ie only one Common segment is allowed to be accessed by a Program segment. The Common checksum is set by the C of the end of segment directive, and subsequently cleared if a successful Common check has been made.

End of Segment Directive (Tag=39)

W  Checksum for RLB tape
C  Only present at end of Common segment, and gives Common checksum

This directive indicates the end of a segment. In the case of a Common segment, the Common checksum is in C, and subsequent segments referring to Common must have the same checksum in their header. In the cases of Program and Library segments the Program area is to be given ten places of left cyclic shift.

Library Variable Block (Tag=4)

W  Checksum for block
C  Library reference number of Variable
C1  Preset value of variable
C2  Identifier String

If there is an outstanding requirement for the variable, this should be set to the initial preset value.
Common Check Header Block (Tag=5)

W Common Checksum

This block is the header of a Common Check tape. The Common checksum in W must match that set up by the previously loaded Common segment. This header is followed by directives (Tag=56) which check that the references within the Common area to Procedures, Switches, and Labels have been supplied. The format of these directives is as follows:

W Location to be checked
C Identifier string

Provided that all the checks have been successful, when the end of segment directive (Tag=39) is read, and its checksum for the tape found to be correct, the Common checksum is cleared. Unless this is cleared, the program cannot be entered. (or another Common loaded)

Instructions referring to Constants (Tag=15)

W Instruction
C Constant

It is left to the loader to optimise the following:

1. The use of the functions 04-07
2. Allocation of storage for constants

In the first case, if the constant is "short", after negation if negative, and the function is in the range 00-03, the function is changed, and the constant inserted into the address field.

In the second case, only one copy of each constant will be stored, irrespective of the number of references to it in different segments (including Library). The address of the constant is inserted into the instruction, and the function part unaltered.

Change Instruction Directive (Tag=35)

W Address of instruction

This directive is used in certain cases of 'FOR' statement. The instruction at location W is to be changed from an 0 JZE to a 4 JCS and the address set to the current Special Data transfer address. (See SETRT)

Print Address Directive (Tag=33)

W Address
C Identifier String

These directives are generated by DIAG if the 'LEVEL' is 2 or 3. These enable "milestones" to be printed at load time with their actual address.
Print Instruction Directive (Tag=34)

W Address of Instruction
C Comment string

These directives are generated by OUTI if the 'LEVEL' is 3. These enable instructions to be printed out in absolute form together with a commentary. As the address of the instruction may not be set until the end of the segment, this requires a second pass of the segment block.

References to Library (Tags=13,29)

W Least significant 14 bits contain reference number.

The Library reference number should be replaced by the address of the (pointer to) the Library item.

References to External Addresses (Tags=14,30)

W Least significant 14 bits contain external "address"

The address requires the addition of the specified external base to the displacement.

Paper Tape

The paper tapes punched by the compiler are made up from combinations of the items given below. All tapes are in ISO code with even parity, and except for blank and erase characters, contain only the 64 "printing" characters, corresponding to the internal code representations 0-63.

A Stop Block and Erases

The five characters 7STOP followed by 20 erases.

B Blank

Ten inches of blank tape.

C Size Block

Gives space requirements for following segment. Comprised of 20 characters, and includes checksum.

D Segment Block

Gives the relocatable binary form of a Program, Common, or Library segment; and includes checksum.

E Common Check Block

Enables a check to be made to ensure that all references within a common segment have been supplied.
F  Library Data Block
    Gives the Identifier, reference number, and preset value of
    a Library variable; and includes checksum.

In addition, the following items may be used.

G  Library Check Block
    The five characters 5LCHQ.

H  Entry Block
    The five characters 6ENTRY

Program Segments
The compiler punches the following items:

    B D B A B C B A B

The size block (C) should be spliced on to the head of the tape, and the
second stop block discarded; thus re-ordering the items as follows:

    B C B D B A B

The lengths of blank tape (B) between other items are ignored, and their
length is unimportant. Where several segments are to be loaded sequentially,
they may be spliced together, and all stop blocks (A) except the final
one, discarded.

Common Segments
The compiler punches the following items:

    B D B A B C B A B E B A B

As with Program segments, the size block (C) should be spliced on to the
head of the tape, thus giving:

    B C B D B A B

This tape should be loaded before any Program segments referring to it are
loaded. When all these Program segments have been loaded, the remainder
of the tape output for the common segment:

    B E B A B

should be loaded. This is the common-check tape, and unless this check is
satisfactorily completed, the loader will not allow the program to be
entered.
Library Segments

The compiler punches the following items:

\[ BDBA BCBAB \]

The size block (C) should be spliced on to the front of the segment block (D), both stop blocks being discarded. The resulting tape:

\[ BCBDDB \]

should be spliced on to the front of the library tape.

Library Variables

The compiler punches the following items:

\[ BFBA B \]

The stop block should be discarded, and the items:

\[ BFBB \]

should be spliced on to the front of the library tape.

Library Tape

The library tape should be built up incrementally, each Procedure referring only to previously compiled Procedures and Variables. These must occur on the relocatable binary library tape in the reverse order to that of their compilation. This is why new segments must be placed on the head of the tape. It is convenient for a library tape to end with the items:

\[ BGBHBB \]

This checks that the library is complete, and causes the program to be entered at the start of the (first) Program segment. (Where a complete Library is to be compiled, a utility program is available for processing the compiler output without the need for tedious splicing.)
Compiler Data Structures

The workspace required by the compiler consists of:

1. A large table, STACK
2. Global variables and arrays
3. Groups of variables specific to certain language features.
4. Local variables for procedures and labelled blocks.

The table STACK, is declared as having only 39 words of store. This is the fixed part which is directly referenced, usually by means of overlayed identifiers. The variable part, which should exceed 3K words if the compiler is to be compiled, is always referred to by pointers. These pointers usually pointing to chains of records. The actual area occupied by the variable part of the stack is set at load time, the three "area" pointers MCSTART, STACKSTART and STACKTOP being set up by the use of Externals. These three pointers refer to areas of core relative to the start of the stack.

The area from MCSTART to STACKSTART is used for macro expansion, its size depending upon the anticipated use of this facility. The area from STACKTOP downwards is used to hold records of labels, this area being shared with the main stack which starts at STACKSTART, and increases upwards. A check is made to ensure that these two stacks do not overlap.

With the exception of Arithmetic Operand records, all records held on the main stack are chained. The first word of each record holding a pointer, relative to the base of the stack, which refers to the succeeding record of the same type. At any one time there may be many chains of records on the stack, and extreme care is required in removing them in the reverse order to that in which they were placed on the stack. Fortunately, because of the nature of the language being compiled, this does not present much difficulty.

Chained records are always placed on the stack using the procedure ONSTACK. This has two parameters, the number of words of the record, and the location in core of the first word. This procedure uses GRABSTACK to allocate the required number of words, which uses STACKCHECK to ensure that this is available. When the record has been placed on the stack, by MOVE, the first location in core from which the record has been moved, is updated to point at the record. Records may be moved off the stack, back to fixed locations in core by the use of OFFSTACK. Note however that this procedure assumes that any space above the record on the stack is no longer required. There is automatically no stack compaction, nor is it needed to languages of the complexity of Coral.

Groups of variables associated with certain (recursive) language features are moved to and from the stack as the depth of (recursion of) the use of the feature increases and decreases. These groups commence with the identifiers:

```
BLOCKCHAIN, LABELCHAIN, PROCCHAIN, IFCHAIN, EXPCHAIN & FORCHAIN
```

The features with which these groups are associated may be deduced from their identifiers.

Not all groups on a given chain are always of the same length. In the case of EXPCHAIN for example: when the expression level is raised by one, at an opening round bracket, five words are placed on the stack, on the expression chain. When a procedure call is detected however, eight words are placed on the stack on the same chain. The extra expression level being required at the procedure call level for the evaluation of parameters.
A very important chain is the DECLIST chain of Identifier Specification records. This holds the specifications of all identifiers (except those of labels local to a segment) which are in scope. These records consist of five words plus the Identifier string. During the processing of declarations these records are built up incrementally in the fixed part of the stack. These are moved onto the main stack, usually in their completed form, by the compiling action NEWNAME. The details of this type of record are given in the declaration of STACK, this being the only type of record, (apart from Arithmetic Operand records, which share the same structure) which is accessed whilst being held on the stack.

Advantage is taken of the effect of OFFSTACK of reclaiming all space occupied above the record to which it is applied, in the actions at the end of a block. This automatically reclaims space occupied by Identifier Specification records of identifiers which go out of scope at that point. Because of the complexities introduced by labels, the identifier records of labels are kept in a separate stack, compaction being required at the end of a block. As no type information is required to be kept, these records are shorter.

Where an Identifier Specification record is that of a procedure, the PARAMSPEC word points at a Parameter Specification record, which usually follows the procedure's Identifier Specification record. This record is of between one and seven words in length. It holds a summary of the parameter requirements of a procedure, in the form of the TYPEBITS word of each parameter, and is terminated by a negative word. Thus if a procedure requires no parameters, its Parameter Specification record consists of a single negative word. As a procedure cannot have more than six parameters, this record cannot exceed seven words in length.

Arithmetic Operand records are created for each primary during the compilation of an expression. These records are five words long, the last four words having the same structure as Identifier Specification records. The first word of these records is not a chain pointer (CHAIN), but a pointer to a string (SPILL), which is used as the comment in a Level 3 listing.

Where a primary is an identifier, possibly subscripted or with parameters, the Arithmetic Operand record is created by copying the relevant four words from the Identifier Specification record, and setting the first word to point to the identifier string in the Identifier Specification record. (By LOOKUP)

Where a primary is not derived from an identifier, for example a constant, the Arithmetic Operand record is synthesised from the available information, and the string pointer set to point to an appropriate fixed string (eg "(CONST)").

Arithmetic Opernad records are placed on the top of the stack. Only the two records currently at the top of the stack are referred to during the generation of instructions for arithmetic operations. These two records are referred to by means of the pointers LH and RH, which point to the left hand and right hand operands of the current operator. When an Arithmetic Operand is added to, or deleted from, the stack, the pointers LH and RH are updated.

Whilst generating the instructions for an arithmetic operation, (eg *) the information held in the records referred to by LH and RH is utilised. At the completion of this, the top (RH) record is deleted, and the remaining (LH) record modified to become the record of the result of the operation. Thus, after the compilation of an arbitrarily complicated expression, only one Arithmetic Operand record remains, this giving the details of the result.
Procedure calls cause one record to be set up on the occurrence of the procedure identifier, and also one record for each parameter. When the call is completed, the records for the parameters are deleted, and the record for the procedure modified to give the details of the result.

Subscripts are processed term by term. The compiler actions operating on the record of the current term, and the record of the variable being subscripted. After each term has been processed, its record is deleted. This optimisation is also used for shift operations, in which case a temporary record, similar to that of an anonymous reference, is used for the operand giving the number of places of shift.
A For Trace and Diagnostics

'LIBRARY' 'PROCEDURE' %ENDOFPROG/1;

This is called at the end of every program segment which does not finish with a GOTO statement.

'LIBRARY' 'PROCEDURE' %INSETLABEL/2('VALUE' 'INTEGER' LABEL);

This is called when a jump is attempted to a label which has not been set. The name of the label is passed in the form of a string as the parameter.

'LIBRARY' 'PROCEDURE' %LABELTRACE/3('VALUE' 'INTEGER' LABEL);

This is called when a label in the scope of a 'LABEL' 'TRACE' directive is jumped to or passed through. The name of the label is passed in the form of a string as the parameter.

'LIBRARY' 'PROCEDURE' %ENTERPROC/4('VALUE' 'INTEGER' PROC);

This is called when a procedure in the scope of a 'PROCEDURE' 'TRACE' directive is entered. The name of the procedure is passed in the form of a string as the parameter.

'LIBRARY' 'INTEGER' 'PROCEDURE' %EXITPROC/5('VALUE' 'INTEGER' PROC);

This is called when a procedure in the scope of a 'PROCEDURE' 'TRACE' directive returns to the point of call. The name of the procedure is passed in the form of a string as the parameter. This procedure must 'ANSWER' %ANSRDUMP to restore the result in cases where the procedure being traced delivers a result.

'LIBRARY' 'PROCEDURE' %ANSWER/6('VALUE' ANSWER:SCALE);

This procedure is called when a procedure in the scope of a 'PROCEDURE' 'TRACE' directive delivers an answer. The first parameter is the result of the procedure and the second parameter its type-scale. The answer must be assigned to %ANSRDUMP so that it may be restored by %EXITPROC.

'LIBRARY' 'INTEGER' 'PROCEDURE' %ASSIGN/7('VALUE' RESULT:SCALE; 'VALUE' 'INTEGER' VAR)

This procedure is called after the value of an assignment statement in the scope of an 'ASSIGNMENT' 'TRACE' directive has been computed, and before the actual assignment takes place. The first parameter is the value calculated and the second parameter its type-scale. The third parameter is the name of the variable to which the assignment is to be made, this is passed in the form of a string. Where the assignment is to be made anonymously the string is "(ANON)".

This procedure must 'ANSWER' RESULT; to restore the value.

'LIBRARY' 'INTEGER' %ANSRDUMP/8

This variable is used to store the result of a procedure between calls of %ANSWER and %EXITPROC.
This procedure is called each time the controlled statement of a for-statement within the scope of a 'LOOP' 'TRACE' directive is entered. The first parameter is the value of the control variable, and the second its type-scale. The third parameter is the name of the variable in the form of a string.

B Floating Point Arithmetic

The following procedures preserve the setting of the overflow indicator on entry, and reset it to its previous setting on exit; unless the operands cause the procedure to deliver a result which lies outside the prescribed range, in which case the overflow indicator will be set on exit, and the result will be zero.

This procedure is used to convert 'INTEGER' and 'FIXED' values to 'FLOATING', in standardised form. The first parameter is the number to be converted, and the second its type-scale.

This procedure performs the floating point multiply operation, delivering as its result the floating point value of the first parameter times the second parameter. (A*B)

This procedure performs the floating point add operation. (A+B)

This procedure performs the floating point subtract operation. (A-B)

This procedure performs the floating point divide operation. (A/B)

This procedure performs the floating point exponentiate operation. (A*B)

This procedure performs the floating point exponentiate operation to an integer power. (A^I)

This procedure performs the floating point subtract operation with the operands interchanged. (B-A)

This procedure performs the floating point divide operation with the operands interchanged. (B/A)

This procedure performs the floating point exponentiate operation with the operands interchanged. (B*A)
This procedure performs the floating point exponentiate operation to an integer power, the operands being in the reverse order to those of %FPRAISEi. ( A11 )

This procedure converts the floating point number supplied as the first parameter to the scale specified by the second parameter.

This procedure normalises and packs floating point numbers. On exit the floating point number is in 7, and overflow is only set if the sign bit of 4 was 1 on entry. If however, the number is out of range, %FPERROR is called.

This procedure is entered by all floating point procedures where the operands or the results lie outside the specified range. This procedure sets overflow and clears 7; (This procedure could be revised to have a parameter specifying the type of error, this could be either a simple integer or a string).

This procedure evaluates the logarithm to base e of the floating point number supplied in 7 as a 'FIXED' (18,11) number in 7. If overflow is set on entry the sign bit of 4 is set to 1. This procedure does not use 6.

This procedure evaluates e to the number given in 7 after fixing the number, using the scale given in 3. This procedure exits to %FPNORM without altering 4.

This procedure evaluates Loge(A), where A>0.

This procedure evaluates e^A.

This procedure evaluates e^{-A}.

This procedure evaluates sine(A), where A is in radians.
This procedure evaluates cosine(A), where A is in radians.

This procedure evaluates \( \tan^{-1}(A/B) \). The result is in radians in the range \( \pm \pi \).

OTHER PROCEDURES DELIVERING A NUMERIC RESULT

This procedure evaluates the square root of \( X \), where \( X > 0 \). The result is scaled to the scale required by OUTSCALE. Overflow will only be set on exit if it was set on entry, if \( X \) is negative, or the result cannot be rescaled to the required scale.

This procedure evaluates \( \sin(\pi X) \). (ie \( X \) is scaled in half revs) Overflow is cleared on exit. As a result of \( +1.0 \) would cause overflow, \( \sin(\pi/2) \) is represented by \( 1-2^{-23} \). Similarly \( \sin(-\pi/2) \) is represented by \( -1+2^{-23} \).

This procedure evaluates cosine(\( X \)) using sine(\( (X+) \)).

This procedure delivers random numbers in the range 0 to 1.

This is used as workspace by RANDOM. Its initial setting determines the sequence of random numbers produced by RANDOM.
Introduction

The Reader (pre-processor) is the only section of the compiler concerned with the individual characters which form the source program. It processes these, grouping them as required, to produce a series of symbols representing the source program. This symbol sequence is then analysed by the syntax analyser, which in this implementation not only checks for the legality of their occurrence within the sequence, but also invokes the required compiling actions at the appropriate points.

Some symbols may have an arbitrarily large number of character combinations allowable as their representation. These differing representations although syntactically indistinguishable are semantically distinct. Thus a symbol group may not only have a syntactic "value" associated with it, but may also have a semantic "value" which is a function of the actual characters themselves. For example: in this implementation, a non-zero integer constant has a single syntactic "value", but has over sixteen million possible semantic values (which in this case is its numeric value), and several times this number of permitted representations.

The syntactic symbols are passed to the syntax analyser, one symbol per call of READER, using the variable T1. As semantic values are evaluated, they are passed to the compiling actions through global variables: NUMBER, NUMBERSCALE, RELOP, ACCUMULATOR, FUNCTION, HACC and NAME. In the last case NAME is the start of an area into which completed strings representing identifiers are placed. It would be possible to incorporate the syntax for identifiers into the table which is used to steer the syntax analyser. This would however prove to be rather slower in operation. As the macro expansion facility is required, the Reader must examine each identifier before it is "seen" by the syntax analyser, to determine whether it is the identifier of a macro. If macro expansion is not required, the string is moved to NAME. An alternative method could be used, where each identifier is represented by a unique integer value, new values being assigned to identifiers on their first occurrence. This technique is often used in conjunction with "Hash Tables".

The Reader removes from the source program all layout and comment, Macro definitions, deletions, and calls; and also 'HALT' directives and optional items.
READER: INFORMATION FLOW

READTAPE → READER → SYNTAX ANALYSER

Internal Code Characters → Terminal Symbols(T!)

Data to compiling actions

<table>
<thead>
<tr>
<th>T</th>
<th>Checked terminal symbols</th>
</tr>
</thead>
<tbody>
<tr>
<td>CSUM</td>
<td>Common checksum</td>
</tr>
<tr>
<td>NUMBER</td>
<td>}</td>
</tr>
<tr>
<td>NUMBERSCALE</td>
<td>}</td>
</tr>
<tr>
<td>RELOP</td>
<td>}</td>
</tr>
<tr>
<td>ACCUMULATOR</td>
<td>}</td>
</tr>
<tr>
<td>FUNCTION</td>
<td>}</td>
</tr>
<tr>
<td>NACC</td>
<td>}</td>
</tr>
<tr>
<td>NAME</td>
<td>}</td>
</tr>
</tbody>
</table>

Relational operators
Code instructions
Shift operators
Accumulator operands
Identifiers
and
strings
<table>
<thead>
<tr>
<th>Value</th>
<th>Symbol</th>
<th>Representation</th>
</tr>
</thead>
<tbody>
<tr>
<td>0</td>
<td>Zero Constant</td>
<td>Any zero constant</td>
</tr>
<tr>
<td>1</td>
<td>Integer Constant</td>
<td>Contains integer part only</td>
</tr>
<tr>
<td>2</td>
<td>Real Constant</td>
<td>Includes . and/or &amp;</td>
</tr>
<tr>
<td>3</td>
<td>Shift Operator</td>
<td>'SRA', 'SLA', 'SRL', 'SLC'</td>
</tr>
<tr>
<td>4</td>
<td>Relational Operator</td>
<td>'&gt; 'GT', '&lt; 'LE', '&lt; 'NE',</td>
</tr>
<tr>
<td></td>
<td></td>
<td>'=' 'EQ', '&lt; 'LT', '&gt;=' 'GE'</td>
</tr>
<tr>
<td>5</td>
<td>String</td>
<td>Enclosed in &quot; quotes</td>
</tr>
<tr>
<td>6</td>
<td>Code Instruction</td>
<td>Octal digit followed by 3 Letters</td>
</tr>
<tr>
<td>7</td>
<td>Differ</td>
<td>'DIFFER'</td>
</tr>
<tr>
<td>8</td>
<td>End</td>
<td>'END', 'E'</td>
</tr>
<tr>
<td>9</td>
<td>If</td>
<td>'IF'</td>
</tr>
<tr>
<td>10</td>
<td>Colon</td>
<td>:</td>
</tr>
<tr>
<td>11</td>
<td>Semicolon</td>
<td>;</td>
</tr>
<tr>
<td>12</td>
<td>Then</td>
<td>THEN</td>
</tr>
<tr>
<td>13</td>
<td>Overflow</td>
<td>'OVERFLOW', 'OVR'</td>
</tr>
<tr>
<td>14</td>
<td>No</td>
<td>'NO'</td>
</tr>
<tr>
<td>15</td>
<td>Trace</td>
<td>'TRACE'</td>
</tr>
<tr>
<td>16</td>
<td>Assignment</td>
<td>'ASSIGNMENT'</td>
</tr>
<tr>
<td>17</td>
<td>Loop</td>
<td>'LOOP'</td>
</tr>
<tr>
<td>18</td>
<td>And</td>
<td>'AND'</td>
</tr>
<tr>
<td>19</td>
<td>Or</td>
<td>'OR'</td>
</tr>
<tr>
<td>20</td>
<td>Union</td>
<td>UNION</td>
</tr>
<tr>
<td>21</td>
<td>Location</td>
<td>'LOCATION', 'L'</td>
</tr>
<tr>
<td>22</td>
<td>Mask</td>
<td>'MASK'</td>
</tr>
<tr>
<td>23</td>
<td>Special</td>
<td>'SPECIAL'</td>
</tr>
<tr>
<td>24</td>
<td>Open Round Bracket</td>
<td>( { ]</td>
</tr>
<tr>
<td>25</td>
<td>Close Round Bracket</td>
<td>) { ]</td>
</tr>
<tr>
<td>26</td>
<td>Multiply</td>
<td>*</td>
</tr>
<tr>
<td>27</td>
<td>Plus</td>
<td>+</td>
</tr>
<tr>
<td>28</td>
<td>Comma</td>
<td>, [ 'BIT' ]</td>
</tr>
<tr>
<td>29</td>
<td>Minus</td>
<td>-</td>
</tr>
<tr>
<td>30</td>
<td>Self</td>
<td>'SELF', *</td>
</tr>
<tr>
<td>31</td>
<td>Divide</td>
<td>/</td>
</tr>
<tr>
<td>32</td>
<td>Accumulator</td>
<td>@0, @1, @2, @3, @4, @5, @6, @7</td>
</tr>
<tr>
<td>33</td>
<td>Array</td>
<td>'ARRAY', 'A'</td>
</tr>
<tr>
<td>34</td>
<td>Begin</td>
<td>'BEGIN', 'B'</td>
</tr>
<tr>
<td>35</td>
<td>Code</td>
<td>'CODE'</td>
</tr>
<tr>
<td>36</td>
<td>Do</td>
<td>'DO'</td>
</tr>
<tr>
<td>37</td>
<td>Else</td>
<td>'ELSE'</td>
</tr>
<tr>
<td>38</td>
<td>For</td>
<td>'FOR'</td>
</tr>
<tr>
<td>39</td>
<td>Goto</td>
<td>'GOTO', 'G'</td>
</tr>
<tr>
<td>40</td>
<td>Common</td>
<td>'COMMON'</td>
</tr>
<tr>
<td>41</td>
<td>Integer</td>
<td>'INTEGER', 'I'</td>
</tr>
<tr>
<td>42</td>
<td>Fixed</td>
<td>'FIXED', 'F'</td>
</tr>
<tr>
<td>43</td>
<td>Floating</td>
<td>'FLOATING'</td>
</tr>
<tr>
<td>44</td>
<td>Label</td>
<td>'LABEL'</td>
</tr>
<tr>
<td>45</td>
<td>Procedure</td>
<td>'PROCEDURE', 'P'</td>
</tr>
<tr>
<td>46</td>
<td>Identifier</td>
<td>Starts with Letter, £, or %</td>
</tr>
<tr>
<td>47</td>
<td>Unsigned</td>
<td>'UNSIGNED', 'U'</td>
</tr>
<tr>
<td>48</td>
<td>Finish</td>
<td>'FINISH'</td>
</tr>
<tr>
<td>49</td>
<td>Switch</td>
<td>'SWITCH', 'S'</td>
</tr>
<tr>
<td>50</td>
<td>Library</td>
<td>'LIBRARY'</td>
</tr>
<tr>
<td>51</td>
<td>Step</td>
<td>'STEP'</td>
</tr>
<tr>
<td>52</td>
<td>Table</td>
<td>'TABLE', 'T'</td>
</tr>
</tbody>
</table>
[These alternatives may only be used in a limited number of cases]
On entry to the Reader, from the syntax analyser, the previous "terminal symbol" is moved from T1 to T, this symbol having been checked and found to be legal. If the one character lookahead facility has been used, an unprocessed character will be waiting in T2. This, for example, might be the character following an identifier, which will not be a letter or digit. If no character is awaiting processing, a call is made of the procedure READ to supply the next character. This may come directly from the program source, or from a macro expansion.

Having obtained the next character to be processed, it is used as an index of the switch CHARACTER. This causes the section appropriate to the character to be entered. This switch has 64 entries, but as many of the entries are repeated it has only 18 distinct ways. This could probably be programmed in rather less than 64 code instructions, but would be slower, and much less amenable to alteration.

**Simple Symbols**

The characters ( ) + - / [ ] ← are terminal symbols in their own right. The corresponding entries in CHARACTER cause a jump to be made to the label OK which is one of the exit points of the Reader. The character value has been assigned to T1, and T2 is cleared (set to -1) before a jump is made back to the syntax analyser at its label EXIT. This label is the common return point for all compiling actions, which although taking the form of a labelled block or section, are treated in the syntax as if they were procedures. This saves over four words per action.

The other exit point from the reader is EX. This is used where the character look ahead procedure READT2 has been used, and the character in T2 inspected but unprocessed. As the actual instruction labelled EX is a 'GOTO' EXIT all the jumps to EX will be set directly to EXIT, this being an effect of STATUS optimisation.

**Language Words**

A commonly occurring character is the ' character denoting the start of a "language word". As each of these words has a fixed meaning, they form in effect an extension to the character set. This is reflected in the operation of the section PR, which is entered in this case. After the word has been recognised by the use of RECOG, the value returned (range 7 - 90) is used as an index of the switch ULWORD, which is used in a manner similar to that of CHARACTER. As however values less than 72 represent language words which are terminal symbols in their own right, this case is treated separately by a test and jump to OK. This reduces the size of ULWORD to 19 entries, although it has only 11 distinct ways.

**List of Recognised Language Words**

(Lower case represents optional letters)

<table>
<thead>
<tr>
<th>Word</th>
<th>Value</th>
</tr>
</thead>
<tbody>
<tr>
<td>A (ARRAY)</td>
<td>33</td>
</tr>
<tr>
<td>Array</td>
<td>33</td>
</tr>
<tr>
<td>Absolute</td>
<td>57</td>
</tr>
<tr>
<td>AND</td>
<td>18</td>
</tr>
</tbody>
</table>

32
<table>
<thead>
<tr>
<th>Token</th>
<th>Page</th>
</tr>
</thead>
<tbody>
<tr>
<td>Answer</td>
<td>60</td>
</tr>
<tr>
<td>Assignment</td>
<td>16</td>
</tr>
<tr>
<td>B (BEGIN)</td>
<td>34</td>
</tr>
<tr>
<td>Bits</td>
<td>66</td>
</tr>
<tr>
<td>BEGIN</td>
<td>34</td>
</tr>
<tr>
<td>C (COMMENT)</td>
<td>82+</td>
</tr>
<tr>
<td>CODE</td>
<td>35</td>
</tr>
<tr>
<td>COMPile</td>
<td>65?</td>
</tr>
<tr>
<td>COMMON</td>
<td>40</td>
</tr>
<tr>
<td>COMMENT</td>
<td>82+</td>
</tr>
<tr>
<td>DO</td>
<td>36</td>
</tr>
<tr>
<td>DUMP</td>
<td>68?</td>
</tr>
<tr>
<td>DIFFer</td>
<td>7</td>
</tr>
<tr>
<td>DEFINE</td>
<td>85+</td>
</tr>
<tr>
<td>DELETE</td>
<td>86+</td>
</tr>
<tr>
<td>E (END)</td>
<td>8</td>
</tr>
<tr>
<td>EQ</td>
<td>75*</td>
</tr>
<tr>
<td>END</td>
<td>8</td>
</tr>
<tr>
<td>ELSE</td>
<td>37</td>
</tr>
<tr>
<td>EXTERNAL</td>
<td>56</td>
</tr>
<tr>
<td>F (FIXED)</td>
<td>42</td>
</tr>
<tr>
<td>FOR</td>
<td>38</td>
</tr>
<tr>
<td>FLOATing</td>
<td>43</td>
</tr>
<tr>
<td>FIXED</td>
<td>42</td>
</tr>
<tr>
<td>FINISH</td>
<td>87*</td>
</tr>
<tr>
<td>GE</td>
<td>77*</td>
</tr>
<tr>
<td>GT</td>
<td>72*</td>
</tr>
<tr>
<td>GOTO</td>
<td>39</td>
</tr>
<tr>
<td>HEX</td>
<td>90*</td>
</tr>
<tr>
<td>HALT</td>
<td>84+</td>
</tr>
<tr>
<td>I (INTEGER)</td>
<td>41</td>
</tr>
<tr>
<td>IF</td>
<td>9</td>
</tr>
<tr>
<td>INTEGRER</td>
<td>41</td>
</tr>
<tr>
<td>L (LOCATION)</td>
<td>21</td>
</tr>
<tr>
<td>LT</td>
<td>76*</td>
</tr>
<tr>
<td>LABEL</td>
<td>44</td>
</tr>
<tr>
<td>LE</td>
<td>73*</td>
</tr>
<tr>
<td>LEVEL</td>
<td>70</td>
</tr>
<tr>
<td>LIBRARY</td>
<td>50</td>
</tr>
<tr>
<td>LITERAL</td>
<td>89*</td>
</tr>
<tr>
<td>LOAD</td>
<td>67?</td>
</tr>
<tr>
<td>LOOP</td>
<td>17</td>
</tr>
<tr>
<td>LOCATION</td>
<td>21</td>
</tr>
<tr>
<td>MASK</td>
<td>22</td>
</tr>
<tr>
<td>NE</td>
<td>74*</td>
</tr>
<tr>
<td>NO</td>
<td>14</td>
</tr>
<tr>
<td>OR</td>
<td>19</td>
</tr>
<tr>
<td>OCTAL</td>
<td>88*</td>
</tr>
<tr>
<td>OVR (OVERFLOW)</td>
<td>13</td>
</tr>
<tr>
<td>OVERLAY</td>
<td>58</td>
</tr>
<tr>
<td>OVERFLOW</td>
<td>13</td>
</tr>
<tr>
<td>P (PROCEDURE)</td>
<td>45</td>
</tr>
<tr>
<td>PAGE</td>
<td>83+</td>
</tr>
<tr>
<td>PROCEDURE</td>
<td>45</td>
</tr>
<tr>
<td>RECOVER</td>
<td>69?</td>
</tr>
<tr>
<td>S (SWITCH)</td>
<td>49</td>
</tr>
<tr>
<td>SELF</td>
<td>30</td>
</tr>
<tr>
<td>STEP</td>
<td>51</td>
</tr>
</tbody>
</table>
Certain characters may be terminal symbols in their own right, or they may be the first of a pair of characters representing a terminal symbol. Each of these characters is allocated a section which tests the following character to determine if the two characters form a pair. These pairs are tabulated below:

<table>
<thead>
<tr>
<th>Character</th>
<th>Section</th>
<th>Possible Pairs</th>
<th>Alternative</th>
</tr>
</thead>
<tbody>
<tr>
<td>:</td>
<td>CN</td>
<td>:=</td>
<td>←</td>
</tr>
<tr>
<td>&lt;</td>
<td>LS</td>
<td>≤</td>
<td>'LE'</td>
</tr>
<tr>
<td>&gt;</td>
<td>GS</td>
<td>≥</td>
<td>'GE'</td>
</tr>
<tr>
<td>*</td>
<td>AS</td>
<td>**</td>
<td>↑</td>
</tr>
</tbody>
</table>

If a pair is detected, a jump is made to OK to delete the second character, the appropriate value having been assigned to T1. If a pair is not detected, a jump is made to EX leaving the second character set up in T2 until the next entry to the Reader. If there were more pairs than tabulated above, it would be preferable to use a single section to process character pairs, using a preset table of combinations.

Comparators

The six possible comparators are treated as identical symbols syntactically. As it does not affect the syntax for comparison which comparator is used, they are all assigned the same terminal symbol value of 4. The variable RELOP is however set to one of six values depending on which comparator has been used.
There are two alternative representations for each comparator, either a single character or character pair (see above), or a two letter "language word". In the first case the sections GS, ES, & LS are entered via CHARACTER, and in the second case the section RO is entered via ULWORD.

The representations of the comparators, and the associated value of RELOP, are as follows:

<table>
<thead>
<tr>
<th>Representation</th>
<th>Alternative</th>
<th>RELOP</th>
</tr>
</thead>
<tbody>
<tr>
<td>&gt;</td>
<td>'GT'</td>
<td>-2</td>
</tr>
<tr>
<td>&lt;=</td>
<td>'LE'</td>
<td>-1</td>
</tr>
<tr>
<td>&lt;&gt;</td>
<td>'NE'</td>
<td>0</td>
</tr>
<tr>
<td>=</td>
<td>'EQ'</td>
<td>1</td>
</tr>
<tr>
<td>&lt;</td>
<td>'LT'</td>
<td>2</td>
</tr>
<tr>
<td>&gt;=</td>
<td>'GE'</td>
<td>3</td>
</tr>
</tbody>
</table>

These values have been carefully chosen. The two negative values represent comparators for which there is no equivalent machine instruction, but by reversing the order of the expressions to be compared (and adding 4 to RELOP) use may be made of available machine instructions. In the case of the four non-negative values, the appropriate machine function code is obtained by adding octal twenty. This gives an if-false-jump instruction, whose sense may be reversed as required.

Shift Operators

The four shift operators are treated as identical symbols syntactically, and are assigned the terminal symbol value of 3. These are recognised by RECOG and the section SH entered via ULWORD. This calculates the function code number from the value assigned to the symbol, which is stored in $T_1$. This number is then stored in FUNCTION.

<table>
<thead>
<tr>
<th>Shift Operator</th>
<th>FUNCTION (octal)</th>
</tr>
</thead>
<tbody>
<tr>
<td>'SRA'</td>
<td>30</td>
</tr>
<tr>
<td>'SLA'</td>
<td>31</td>
</tr>
<tr>
<td>'SRL'</td>
<td>32</td>
</tr>
<tr>
<td>'SLC'</td>
<td>33</td>
</tr>
</tbody>
</table>

Spaces and Newlines

Spaces are ignored in two ways. If a space is read and used as an index to CHARACTER, the entry is SP which causes the next character to be read. Where READT2 is used, this procedure specifically tests for the space character, and if it is found, reads the next character. Newlines are removed by READ.

Comments

Three types of comment are allowed, end comment, bracketed comment, and explicit comment.

After reading and packing an identifier, a check is made to see if the last terminal symbol was 'END'. This is performed by testing whether $T$ contains the terminal symbol number for 'END', which is 8. If it does, the identifier is discarded and a jump back made to process the following character, which will be held in $T_2$. 
After each unquoted semicolon, occurring as a terminal symbol, or as the terminator of a 'DEFINE', 'DELETE', 'COMMENT' or 'PAGE' item, a call is made of SEMICOM.

This procedure checks for, and reads if present, bracketed comment. It keeps a count of the bracket level, so that matched round brackets are allowed within this type of comment.

Explicit comment, prefixed by 'COMMENT' (or 'C') is read in and ignored, up to and including the terminating semicolon. 'PAGE' comment is a special case of this type of comment. It is treated in the same manner, but where LEVEL is non zero, it is printed on the monitor printer.

Optional Items

The use of the unquoted ? character allows a limited form of compile time selection from the source program. If handswitch 1 is up, the character is ignored, and the following characters processed. If handswitch 1 is down, the ? is treated as if it were 'COMMENT', and the following characters up to and including the following semicolon ignored.

Code Instructions

Except in identifiers, and hexadecimal constants, a letter following a digit can only occur in code. This allows the fields of a code instruction to be used without intervening separators such as commas. Where an octal digit is followed by a letter, the digit is stored in ACCUMULATOR and a call made to RECOG to process the following three letter function code, the value returned being stored in FUNCTION. The flexibility of RECOG would allow function mnemonics to be of varying length, but the three letter groups are used to ensure as near as possible conformity with Astral. An accumulator-function pair has a terminal symbol value of 6.

Where an accumulator is to be used as an operand or modifier it is represented by the symbol @ followed by a single digit accumulator number. This has AMARK added and is stored in NACC. This representation has a terminal symbol value of 32.

Strings

Strings are processed by the section ST which is entered when the opening " is read. The string is assembled in MCID and subsequently moved to NAME, from where it is accessed by the compiling actions. As the string is terminated by another " it is necessary to have an alternative representation for this character within a string. This is achieved by representing the symbol as ". Thus to represent a string containing "A"B" it is necessary to write "A""B". A string is packed in standard Argus form and may contain up to 63 characters.

A string has a Terminal Symbol value of 5.
Constants

The Reader processes all constants completely, before passing a Terminal Symbol code representing the type of constant to the syntax analyser. Constants may be in one of four forms; Character, Hexadecimal, Octal and Decimal. In the first two cases the constant is an Integer, but in the last two it may be either Integer or Real. A zero constant, whatever its form, is treated as a special case. The constant is assembled in NUMBER and its scale factor is placed in NUMBERSCALE for subsequent use by the compiling actions. The types of constant and their Terminal Symbol values are tabulated below.

<table>
<thead>
<tr>
<th>Type</th>
<th>Terminal Symbol value (T1)</th>
</tr>
</thead>
<tbody>
<tr>
<td>Zero</td>
<td>0</td>
</tr>
<tr>
<td>Integer</td>
<td>1</td>
</tr>
<tr>
<td>Real</td>
<td>2</td>
</tr>
</tbody>
</table>

Character Constants

The official definition only specifies one form of character constant. This takes the form:

'\text{LITERAL}'(c)

where c is any printing character. In this implementation this may be abbreviated to $c$. In either of these cases the standard internal character value is stored in NUMBER and the constant treated as an integer constant.

As a further extension, up to four characters may be used. These are packed right justified in a similar manner to Alpha constants in Astral. In this case the constant takes the form:

'\text{LITERAL}'n(ccc)

where n is in the range 1 - 4 and specifies the number of characters ccc. 'LITERAL' may be abbreviated to 'LIT'. These constants are processed by the sections CH ($\$ $), and LI ("LIT").

Hexadecimal Constants

These constants have only one form of representation, the symbol 'HEX' followed by one or more hexadecimal digits (0-9, A-F). These are assembled by the section HX as a right justified integer constant in NUMBER. A hexadecimal constant is treated as an Integer constant.

Octal Constants

These constants may be represented in a number of ways.

1. # ddd
2. 'OCTAL' ddd
3. 'OCTAL'(ddd)
4. # J ddd
5. 'OCTAL' J ddd
6. # J ddd.ddd
7. 'OCTAL' ddd.ddd
8. 'OCTAL'(ddd.ddd)
where ddd represents from one to eight octal digits, and 'OCTAL' may be abbreviated to 'OCT'. Forms 1, 2 & 3 are treated as right justified Integer constants, and forms 4 & 5 as left justified Integer constants. The remaining forms 6, 7 & 8 are treated as Real constants, and in this case the integer part must be less than #40000000. The Official Definition only permits forms 3 & 8.

**Decimal Constants**

These constants may be represented in a number of ways.

1. ddd
2. ddd.dd
3. &sd
4. ddd&sd
5. ddd.dd&sd

Where:

- ddd represents 1 - 7 decimal digits whose value is less than 8388608
- dd represents 1 - 6 decimal digits
- sd represents a single decimal digit, optionally preceded by a sign (+,-).

The first form is treated as an Integer, the remaining forms as Real. The character & replaces the symbol 10 and is used to indicate powers of ten. The real and fractional parts of a Real constant are stored as a fixed point mixed number in NUMBER.FRAC prior to normalising to discover a suitable scale factor. Thus if a small number is represented in form (2), as a zero integer part and a fraction with leading zeros, accuracy will be lost. In this case it is better to use form (5), with a negative decimal exponent. eg 1.2&-3 will give a more accurate constant than 0.001200.

**Macro Definition**

Macro definitions are processed by the section DF which is entered when the symbol 'DEFINE' is recognised. This reads the macro definition and stores it in the area of the compiler's stack reserved for macros. The record of the macro definition consists of the following information stored in sequential locations.

1. A chain pointer to the previously defined macros.
2. The number of parameters required.
3. The macro name, in standard string form.
4. The parameter identifiers (if any), in standard string form.
5. The body of the macro, stored as packed characters, four to a word. The character " is represented by ", and the body is terminated by ";
It should be noted that no attempt is made at this point to recognise the occurrence of the identifiers of the parameters within the macro body. Although the body of a macro is enclosed in quotes and is similar in appearance to a string, it is not stored as a standard string as this would restrict the number of characters in the body to 63. Advantage is taken of the fact that the representation of the quote symbol within a macro body, as in a string, is by a pair of quotes. A macro body being terminated by a single unpaired quote, followed by an arbitrary character, semicolon being used for convenience. Unlike a string, where paired quotes are reduced to a single quote whilst packing, paired quotes are reduced to a single quote whilst unpacking a macro body.

**Macro Deletion**

Macro deletions are processed by the section DL which is entered when the symbol 'DELETE' is recognised. If the macro to be deleted is on the top of the macro stack, and macro expansion is not in progress, the macro is deleted by unchaining its record and recovering the space occupied. Its record is thereby lost.

Otherwise the macro is deleted by setting the first word of its identifier string to zero, thereby rendering its record inaccessible. In this case the space occupied is not recovered, as a "garbage collector" is not incorporated. Thus to ensure that the macro stack does not overflow, the following rule should be observed:

> Delete all macros in the reverse order to their definition.

It should be noted that the scope of macros does not follow the block structure of the language, all macros being effectively global irrespective of the block level at which they are defined. It is therefore suggested that the following rule be observed:

> All macros defined within a program segment should be deleted at the end of the segment.

**Identifiers and Macro Expansion**

An identifier commences with a letter A-Z, a $, or a %; and is followed by letters and digits. The identifier is assembled in MCID by the section ID calling IDENT.

If the identifier follows the symbol 'END' it is ignored as comment. If the identifier is not the name of a currently defined macro, or of a current macro parameter, it is moved to NAME and a Terminal Symbol value of 46 placed in T1.

Otherwise a new level of macro expansion is set up. If the macro requires parameters, the identifier must be followed by the correct number of parameters, separated by commas and enclosed in round brackets. Each of these parameters is used to form the body of a temporary macro which is given the name of the appropriate parameter, which is obtained from the record of the macro being invoked. Thus no special test is required for the occurrence of a macro parameter within a macro body. If a macro requires no parameters (and this applies also to parameters within macros) it may be optionally followed by a ' character, which is ignored.
Once a macro expansion level has been set up, characters are read sequentially from the macro body until the terminating unpaired quote is read. This sequence can only be interrupted by the setting up of a further level of macro expansion. When the terminator of a macro body is read, the macro expansion level is reduced, which deletes any macros declared within that level. The input is resumed from the next lower level, at the point at which it was interrupted, after restoring T2, if relevant. This case will only occur where a parameterless macro, or parameter, is not followed by the \texttt{I} character.

Example of Macro Expansion

Consider the following portion of a program:

\begin{verbatim}
MACK
 'BEGIN' 'INTEGER' ALPHA;
 'DEFINE' LDX(ACC,ADD)"ACCI
LDXADD!";
 'CODE' 'BEGIN' LDX(7,ALPHA);
\end{verbatim}

After storing the definition of LDX and reading the following \texttt{CODE} \texttt{BEGIN}, the identifier LDX would be recognised as being that of a macro. After increasing the level of macro expansion by stacking the current values of MCCHAIN, T2, MCSOURCE, MCBODY and MC LIST, temporary macros would be set up for the two parameters. This would be equivalent to reading:

\begin{verbatim}
'DEFINE' ACC"7";
'DEFINE' ADD"ALPHA";
\end{verbatim}

The body of LDX would then be read as the current input, and when the identifier \texttt{ACC} was recognised as being that of a macro a further expansion level would be set up, and the \texttt{I} discarded. At this point, the next character to be read would be 7. The state of the macro stack at this point is given in the diagram below.

After reading the 7 the level would be reduced and the LDX would be read. As this occurs after an octal digit it would be treated as a function mnemonic and not an identifier. The letters ADD which follow would be treated as an identifier, which would be discovered to be a macro, and replaced by \texttt{ALPHA}. The expansion level would then be reduced twice, and the semicolon read. This would terminate the identifier \texttt{ALPHA}, which not being a macro would be passed on.

It should be noted that the occurrence of second parameter ADD in the body of LDX is followed by an \texttt{I} character. This may at first sight seem to be superfluous. If it were not present however, the recognition of the termination of the occurrence of ADD as an identifier would only be achieved after the semicolon had been read; by which time the macro expansion level would have been reduced to zero, and the temporary macro for ADD deleted!

In fact, had the semicolon not been present, and the next characters were say \texttt{X+}, the effective identifier would have been \texttt{ADDX}, for which no definition or declaration exists.

A general rule for macro parameters, which is an oversimplification of the facilities available, is:

\begin{verbatim}
Name all macro parameters \%1, \%2, \%3 etc, and in the body of the macro refer to them as \%1, \%2, \%3 etc.
\end{verbatim}

This rule is guaranteed to avoid confusion, and render macro definitions more legible.
Macro Stack state during expansion of macro

Input Tape:  
LDX(7,ALPHA);

Level 1:  
ACC!LDXADD!

Level 2:  
7

↑ next character to be read

Increasing addresses

Fixed Workspace

MCLIST  
MCSOURCE

MCSOURCE  
T2

MCMEM  
MCTOP

Increasing addresses

MCLIST  
MCBODY  
MCSOURCE  
T2  
MCMEM  
MCTOP

Stacked workspace during replacement of ACC

Macro body

Definition of parameter

Macro name

Number of parameters

Pointer to next macro

Macro body

Definition of parameter

Macro name

Number of parameters

Pointer to next macro

Stacked workspace during expansion of ACC

Macro body

Definition of macro

Second parameter

First parameter strings

Macro name

Number of parameters

Pointer to next macro

41
Label AC:

This section is entered, via the switch CHARACTER, when the character 0 is encountered. The following character must be an octal digit specifying an accumulator number for use in code. This number has AMARK added and is stored in NACC. A jump is then made to OK, leaving the value of T1 (32) unaltered.

'PROCEDURE' ADDCHAR;

This procedure is used to append the current character held in T1 to the standard string being formed in the MCID area at the base of the stack. The count field, in the first character position, is first incremented, and if it overflows (sets carry) a jump is made to FT. All identifiers are assembled in MCID in order that they may be tested to determine whether they are the identifiers of macros, or current macro parameters. If they are not, the string will subsequently be moved to NAME. For convenience, this process is also used for quoted strings, but in this case the completed string is not treated as a potential macro.

Label AM:

This section is entered, via the switch CHARACTER, when the character & is encountered. This denotes a Real constant which is the specified power of ten. NUMBER is set to 1024 and NUMBERSCALE to -10. This is equivalent to the constant 1, but guards against loss of significant accuracy. A jump is then made to EXPT to read the required power of ten.

Label AS:

This section is entered, via the switch CHARACTER, when the character * is encountered. The following character is read, using README2, and if it is a second *, the * in T1 is replaced by t, and a jump made to OK. Otherwise a jump is made to EX.

Label CH:

This section is entered, via the switch CHARACTER, when the character $ is encountered. The following printing character is treated as an integer constant, and the value of its internal representation placed in NUMBER. A jump is then made to INTX via INTY, which reads the character following, into T2.

Label CM:

This section is entered, via the switch ULWORD, when the symbol 'COMMENT' (or 'C') is recognised. The following characters, up to and including the terminating semicolon, are read and ignored. A call is then made of SEMICOM to test for following bracketed comments, and a jump made to SP, to process the character left in T2 by SEMICOM.

This section is also used by PG (T1=83) to process page titles, and is slightly more complex than would otherwise be the case.
Label CN:

This section is entered, via the switch CHARACTER, when the character : is encountered. The following character is read, using READE2, and if it is the = character, the : in T1 is replaced by ← and a jump made to OK. Otherwise a jump is made to EX.

Label DD:

This section is entered, via the switch CHARACTER, when a digit is encountered. If this digit is followed by a letter, a call is made of RECOG, to process the three letter function code mnemonic, and return its number. The initial digit is stored in ACCUMULATOR, and the function number in FUNCTION. T1 is set to 6, the terminal symbol value for a code instruction, and a jump made to EX.

Otherwise a call is made of NOASSY to assemble the integer part of a decimal constant. If this is followed by the symbol for a decimal exponent (\&), a jump is made to EXPT. Otherwise if the following symbol is not a decimal point, a jump is made to INTX. If a decimal point is present, a call is made of NOASSY to assemble the following fractional part, as an integer. This is then divided by the appropriate power of ten, to obtain an unsigned fraction which is stored in FRACTION. If the following symbol is not a decimal exponent symbol (\&) a jump is made to NORM, otherwise the section EXPT is entered.

A flowchart covering the whole of the assembly of decimal numbers is given.

Label DF:

This section is entered, via the switch ULWORD, when the symbol 'DEFINE' is recognised. This section sets up the definition record of a new macro.

MCFLAG is set to indicate that a macro definition is in progress, and MCMAKE, MCIDENT and MCMOVE called to read the name of the macro and initiate its record. If this is followed by an open round bracket, the macro has parameters, and the names of these are read and added to the record. The parameter identifiers are separated by commas, and the list is terminated by a close round bracket.

The macro body, which must start with a quote symbol, is read and added to the record, character by character, up to and including the last unpaired quote and the following mandatory semicolon. The macro definition is now complete, and MCFLAG is cleared. Calls are made of MCTIDY and SEMICOM. In the latter case this causes any following bracketed comment to be ignored. A jump is then made to SP to process the character left in T2 by SEMICOM.

A flowchart is given for this section.

Label DL:

This section is entered, via the switch ULWORD, when the symbol 'DELETE' is recognised. This must be followed by the identifier of a current macro, which is followed by a semicolon.
The reference to the record of the macro is obtained using MCLOOK. If the macro does not exist, the reference will be zero, and this is treated as an error. If there are no macros currently being expanded and the macro to be deleted is the top macro on the stack, the macro is deleted by unchaining its record, and recovering the space occupied. Otherwise the macro is rendered anonymous by clearing the first word of its identifier string. The space occupied by its record is not recovered in this case. If however the macro is defined within a macro expansion, or is a temporary macro set up for a parameter, this space will subsequently be recovered. Thus as a general rule: delete macros in the reverse order to their definition.

Label ES:

This section is entered, via the switch CHARACTER, when the character = is encountered. T1 is set to 4, to indicate a relational operator, and RELOP set to 1. A jump is then made to OK.

Label EX:

This is the exit point from the reader in those cases where the character look ahead facility READT2 has been used, and an unprocessed character remains in T2. This will be processed on the next entry to reader.

Label EXPT:

This section is entered from AM and DD when the & character is encountered. This may optionally be followed by a sign (+,-), and then must be followed by a single decimal digit. The double length number NUMBER.PRAC is then repeatedly multiplied (+) or divided (-) by ten, the number of times being specified by the digit.

To improve the numerical accuracy, by preserving as many significant bits as possible, the calculation is carried out in a degenerate form of floating point arithmetic. To multiply by ten, the number is actually multiplied by 0.625 and 4 added to its exponent held in NUMBERSCALE. To divide by ten, the number is multiplied by 0.8 and 3 is subtracted from its exponent. Because of the use of multiplication in both cases, there is no possibility of overflow, and intermediate normalisation is not required.

The section NORM is then entered.

Label FI:

This section is entered, via the switch ULWORD, when the symbol 'FINISH' is recognised. T1 is set to 48, the value representing the symbol, and T2 is set to an arbitrary value selected from the set of characters which are terminal symbols in their own right. This ensures that no more tape is read after the syntax analyser has accepted the 'FINISH'. A jump is then made to EX which leaves T2 set.

Label FT:

This section is entered on the discovery of an illegal character, or combination of characters. Compilation is terminated with the message "CHARACTER FAULT".
'PROCEDURE' GENOCT;

This procedure processes the general case of octal numbers. These may have simply an integer part, in which case they are treated as Integer constants. Or they may have both an integer and a fraction part, in which case they are treated as real constants.

The integer part is first assembled by OCTALNO, and assigned to NUMBER. If this is not followed by an octal point, T1 is set to 1 to denote an Integer constant. Otherwise a check is made to ensure that the integer part appears positive, and T1 set to 2 to denote a Real constant. The fraction part is assembled by OCTALFRAC, the result assigned to FRAC, and NUMBERSCALE cleared.

Label GS:

This section is entered, via the switch CHARACTER, when the character > is encountered. T1 is set to 4, to indicate a relational operator, and the following character read, using READT2. If this is =, RELOP is set to 3 and a jump made to OK. Otherwise RELOP is set to -2, and a jump made to EX.

Label HA:

This section is entered, via the switch ULWORD, when a 'HALT' directive is recognised. A call is made of HALT with "DIRECTIVE" as the parameter. On return, READT2 is called to read the next character, and a jump made to SP to process it. This causes the 'HALT' directive to have no syntactic effect.

Label HX:

This section is entered, via the switch ULWORD, when the symbol 'HEX' is recognised. It assembles hexadecimal integers in NUMBER. After the last character, a jump is made to INTX.

Label ID:

This section is entered, via the switch CHARACTER, when a letter, or £, or %, occurring as the first character of an identifier is read. The remainder of the identifier is then read by IDENT.

If this identifier follows the symbol 'END', it is ignored as this is an allowable form of comment. Otherwise McLook is used to test whether the identifier is that of a current macro. If it is not, the identifier is moved to NAME, the value (46) for an identifier assigned to T1, and a jump made to EX.

If the macro is parameterless, it may optionally be followed by I, which is then ignored. (If this facility were not available it would be impossible to specify the accumulator field of a code instruction as a macro parameter.) If the macro requires parameters, they must be supplied, and in this case the macro identifier must be followed by an open round bracket.

After this check, a new macro expansion level is set up, by MOVEing the variables of the previous level onto the macro stack. A temporary macro is then set up for each parameter. These are parameterless macros having the parameter identifier, and their body consists of the actual parameter supplied by the call. These macros will be automatically deleted when the macro generation level is subsequently reduced at the end of the expansion.

The starting address of the macro body is set into MCSOURCE, and a jump made to SP to process the first character.

A flowchart of this section is given.

47
'PROCEDURE' IDENT;

This procedure is used to assemble identifiers, after the first character has been checked. The identifiers are assembled into the area at the base of the stack, starting at MCID. This area is used as each identifier (except those following 'END', 'DEFINE' and 'DELETE') have to be treated as potential macro calls.

The area is first initialised as an empty string, and the first (checked) character inserted. Subsequent characters are read into T1 and also T2, and if the character is a letter or digit, it is added to the string by ADDCHAR. When a character is found which is not a letter or digit, this procedure is terminated, leaving the character in T2 for subsequent examination.

Label INTX:

This section is entered from DD,OC and INTY. In all these cases an integer, but not necessarily numeric, constant has been read. If this is zero, a jump is made to ZERO. Otherwise the integer is shifted left to discover the number of significant bits required to represent it without loss of accuracy. This number is required in certain cases of scaled arithmetic. T1 is set to 1 to indicate a non-zero integer constant, and NUMBERSCALE is set to indicate an integer constant requiring the number of significant bits determined above. A jump is then made to EX.

Label LI:

This section is entered, via the switch UL,ORD, when the symbol 'LITERAL' is recognised. This is then followed by a single printing character, enclosed in round brackets. As an extension in this implementation, more than one character, but less than five, may be included between the brackets, provided that their number is specified by a single digit placed before the opening bracket. The characters are assembled as a right justified integer to base 64 in NUMBER, and a jump made to INTX via INTY which reads the following character into T2.

Label LS:

This section is entered, via the switch CHARACTER, when the character < is encountered. T1 is set to 4, to indicate a relational operator, and the following character read, using READT2. If this is =, RELOP is set to 01, or if it is >, RELOP is set to zero; in either case a jump is then made to OK. Otherwise RELOP is set to 2, and a jump made to EX.

'INTEGER' 'PROCEDURE' MCCHAR

('VALUE' 'INTEGER' CHAR);

This procedure is used to append the character CHAR to the body of the macro currently being defined. It uses MCTOP as a character modifier, (m.s. 2 bits give character position within word) which is incremented by one character position each time. The character CHAR is returned as the result in case subsequent examination is required.

This procedure is functionally the inverse of NEXTCHAR, but whereas the parameter of NEXTCHAR refers to an absolute core location, the implied parameter MCTOP, of MCCHAR refers to a core location relative to the start of the stack.
'PROCEDURE' MCIDENT;

This procedure is used to read and assemble the identifier of a macro which must follow 'DEFINE' and 'DELETE', and also the parameter identifiers in a macro definition. The first character is read into Ti. If it is not a letter (A-Z), nor either of the characters £ or %, a jump is made to ME. Otherwise the identifier is read in using IDENT.

'INTEGER' 'PROCEDURE' MCLOOK;

This procedure is called to test whether the identifier held in the locations starting with MCID, is the identifier of a currently defined macro. If it is, a pointer to the macro definition record is returned, this value also being assigned to MAC. Otherwise a zero result is returned.

Using MAC as the control variable, with an initial value of MCLIST, a scan is made down the list of macros, comparing the macro identifier string with that at MCID, using TESTSTRING. The scan terminates either when equality is found, or the list is exhausted. The current value of MAC serves as the result in either case.

'PROCEDURE' MCMAKE;

This procedure is called to prepare for the definition of an additional macro. This may be a macro explicitly defined by the 'DEFINE' symbol, or a temporary macro defining the parameter of a macro as the actual parameter supplied.

Two locations on the top of the macro stack are reserved and initialised. The first is used as the macro chain, and is set to the current contents of MCLIST, which is then set to point at the new macro. The second word will be used to contain the number of parameters required by the macro, and is at this point set to -1. This two word record will be followed by the string giving the macro identifier, the parameter identifiers (if any), and finally the macro body.

'PROCEDURE' MCMOVE

('LOCATION' 'INTEGER' FROM);

This procedure is used during the definition of a macro, to allocate space for and to insert identifier strings into the head of the macro record. For each record, the first call is used to insert the macro name, and any subsequent calls insert the names of the parameters. The second word of the macro record has one added to it each time by this procedure, and as its initial value is -1, it may be seen that this location will contain a count of the number of parameters.

When setting up a record for an explicitly defined macro, the parameter FROM will be the location of MCID, which contains the current identifier. When setting up temporary macros defining the parameters of a macro, the parameter FROM will point at the identifier string of the parameter in the record of the called macro.
'PROCEDURE' MCTIDY;

This procedure is called to ensure that the macro stack pointer NCTOP, starts on a word boundary, (it is used as a character modifier during macro definition) and that its value does not exceed the size of the macro stack. If its value exceeds STACKSTART, the value for the starting point for declarations on the stack, compilation is terminated by a call of GIVEUP with "MACRO STACK OVERFLOW" as the parameter.

Label ME:

This section is entered when an error is discovered in the definition, deletion, or expansion of a macro. Compilation is terminated with the message "MACRO . . . OR".

'PROCEDURE' MOVESTRING

('LOCATION' 'INTEGER' FROM, TO);

This procedure moves the string whose first location is given by the parameter FROM, to the area of core whose location is given by the parameter TO. The number of words to be moved is calculated from the count character of the string, this number being used as the first parameter of the call of MOVE which performs the operation.

'INTEGER' 'PROCEDURE' NOASSY

('VALUE' 'INTEGER' BASE, LIM);

This procedure is a general number assembly routine. It assembles up to LIM digits of a number to base BASE, where BASE may be in the range 2 to 10. On entry the first digit must be in T1, and the second character (which may not be a digit) must be in T2. The assembled number is returned as the result, and the number of digits processed is placed in NODS. If the first character is not a valid digit, or the number of digits exceeds the specified limit, a jump is made to FT. This procedure is used to assemble both the integer and the fractional parts (as integers) of octal and decimal numbers.

Label NORM:

This section is entered from DD, EXPT, and OC. A double length number is supplied in NUMBER.FRAC with a binary scale factor in NUMBERSCALE. The number is shifted left as far as possible without overflow, NUMBERSCALE being adjusted according to the number of places of shift. Provided that the number is non zero, (if it is a jump is made to ZERO) the appropriate 'FIXED' scale is calculated for the number. This scale may subsequently be inserted in the TYPEBITS field of an Arithmetic Operand. The CON marker is set, the TYP field set to 1 (fixed), and the SGB field set to 24. The PVL field is calculated from the value of NUMBERSCALE. This gives a scale with the minimum number of integer bits required to hold the number. This is then stored in NUMBERSCALE.

Provided that this scale lies within the range allowed by the implementation, T1 is set to 2, to indicate a Real constant, and a jump made to EX.
Label OC:

This section is entered, via the switch ULWORD, when the character # is encountered. This denotes an octal constant. It is also entered at OQ and OX from OI.

If the following character is J, the number is a left justified integer. This is read using OCTALFRAC, which left justifies it, and a jump is made to INTX.

Otherwise a call is made to GENOCT to read an Integer or Real octal constant. If the constant is an integer, a jump is then made to INTX, otherwise a jump is made to NORM.

'INTEGER' 'PROCEDURE' OCTALFRAC;

This procedure returns a 24 bit (unsigned), left justified octal number. It is called after the octal point to assemble the following octal fraction, and after the symbols #J to assemble a left justified octal integer.

The first digit is read by a call of READT1, and the number assembled, right justified, by OCTALNO. This is then shifted left, the number of places of shift being calculated from the number of digits NODS, and returned as the result.

'OCTALNO';

This procedure is used to assemble octal numbers of up to eight digits. On entry the first digit must be in TI, and the next character is obtained by means of a call of READT2. NOASSY is then used to assemble the number.

Label OI:

This section is entered, via the switch ULWORD, when the symbol 'OCTAL' is recognised. In this implementation the following octal number may be unbracketed, if required, and in this case a jump is made to OQ. Otherwise, after the opening round bracket has been read, a call is made to GENOCT to read the number, and the closing round bracket checked. After reading the character following the closing brackets, a jump is made to OX. This redundant character read operation is required as an exit will subsequently be made via EX, which does not clear T2.

Label OK:

This is the exit point from the reader in those cases where the character held in T2 has been processed. T2 is cleared, by setting it negative, so that on the next entry to the reader a further character will be read.

Label PG:

This section is entered, via the switch ULWORD, when the symbol 'PAGE' is recognised. If LEVEL is non-zero, "PAGE:" is printed on a new line, and followed by the characters supplied by READ, up to but not including the terminating semicolon. This is followed by two newlines. If LEVEL is zero, 'PAGE' is equivalent to 'COMMENT'.

52
Label PR:

This section is entered, via the switch CHARACTER, when the starting character of a "language word" is read. A call is made of READT2 to read the first letter, and then RECOG called, with a parameter of 1, to identify the symbol. The result of RECOG is assigned to T1, and if this is less than 72, it represents the value allocated to a terminal symbol, and a jump is made to OK. Otherwise the value is used to select an entry in the switch ULWORD.

Label QM:

This section is entered, via the switch CHARACTER, when the character ? is read. This is the starting symbol of an optional item. If Handswitch 1 is up (0), the ? symbol is ignored, and the following item processed in the normal manner. Otherwise the ? symbol is treated as the 'COMMENT' symbol, and the following characters, up to and including the terminating semicolon, ignored.

'INTEGER' 'PROCEDURE' READ;

This procedure is the source of characters for the pre-processor. They are obtained either from the input tape or from the expansion of a macro.

If there are no macros currently being expanded, READTAPE is called to read the next character from the input tape. If this character is the newline character, it is ignored, and the call repeated.

Otherwise, the next character is unpacked from the body of the current macro definition, using NEXTCHAR. If this character is the quotes character ("), the next character is also read. If this second character is also a quote, this is allowed to stand. Otherwise the macro expansion level is reduced by the use of MOVE to restore the "MC" variables of the next lower level, and to restore T2. Note that although these variables are only used in READER, they are declared in the global data, as their values are required to be preserved from one call of READER to the next. If a macro is currently being defined, MACFLAG will be set, and it is not possible to reduce the macro expansion level at this point. If the restored value of T2 is a valid character, this is allowed to stand as the next character. Otherwise a jump is made to the start of this procedure to obtain the next character.

After the character has been obtained, it is used to form part of the common checksum CSUM if a 'COMMON' segment is currently being compiled.

A flowchart for this procedure is given.

'INTEGER' 'PROCEDURE' READT1;

This procedure is used as a one character look-ahead facility in cases where it is known that the following non-ignored character, if legal, must form part of a composite symbol. (e.g. a decimal point must be followed by at least one decimal digit.)

This procedure calls READT2 to obtain a character, which it stores in T1, and also returns as the result.
'INTEGER' 'PROCEDURE' READT2;

This procedure is chiefly used to give a one character look-ahead facility in cases where the next non-ignored character may, or may not, form part of composite symbol with the previous characters. (e.g. the character following two integer decimal digits may be another digit, a decimal point (.), or an exponent symbol (^); in which case it forms part of the number. Any other character does not.)

This procedure repeatedly calls READ until a non-space character is obtained. This is stored in T2 and also returned as the result. Unless overwritten, the character stored in T2 will be processed at the next call of READER.

'INTEGER' 'PROCEDURE' RECOG

('VALUE' 'INTEGER' ULW);

This procedure is a recogniser for alphabetic symbols which are members of two fixed groups. The first of these groups (ULW=1) being the larger group consisting of the "language words". (e.g. 'BEGIN', 'END'.) In this second case the use of unambiguous abbreviations is allowed in all cases, and the use of a few specified ambiguous abbreviations is also permitted. (e.g. 'E' for 'END' (not 'ELSE')). Certain of these language words have other representations, which are not the concern of this procedure. (e.g. 'NE' and <>, 'STEP' and :).

This procedure operates by interpreting one of the two fixed arrays FUNCTIONS and ULWRDS. These consist of six bit characters, packed four to a word; and have been automatically generated from lists of the allowable symbols. Each letter in a symbol is represented by its internal code value (range 33 to 58, A to Z). Values in the range 60 to 63 represent terminating characters, and together with the following character give the number allotted to the symbol (0-255). Values less than 33 are "jump" characters, and are used in the selection of alternative routes.

These arrays are a one dimensional representation of a two dimensional list structure. This procedure moves through the structure, comparing the stored letters with the letters read in, and taking the appropriate action. Where differing symbols have a common starting letter or letters, no backtracking or scanning is required. In order to speed up the recognition process, and because of the limit on the "jump" size, each structure is entered via a small "hash" table stored at the start of the array. This is indexed by the least significant one (FUNCTIONS) or two (ULWRDS) bits of the first letter. The structure of FUNCTIONS is illustrated as follows:

```
Hash 0
D N L J
I E L A D O B C G L N Z
V Q X B X C S S E T Z E

Hash 1
E M A O S
X P N D R U V S U R T L
C Y D D S F T R B B A L N O A C L V
```
The procedure moves across the diagram to select alternatives, and downwards to select successors. The diagram for ULWRDS is too large for simple representation. In this case the incoming symbol is terminated by a character, and if it is an unambiguous abbreviation, the rest of the route through the structure must not contain any branches. Part of the structure, of the entries for C, of ULWRDS is illustrated below:

```
C
  O
  D
   M
    P
     M
      I
       O
        E
         L
          N
           N
            E
             40
                T
                 62
                 62
```

On entry to the procedure, the array CHOOSE is used to select the array to be used, according to the value (0 or 1) of ULW. CHOOSE also contains the masks required to select the hashed entry into the selected table. The character address of the current character is stored in SI, which is used by the procedure READS, which uses NEXTCHAR to unpack successive characters.

On successful recognition of an alphabetic symbol, a result in the range 0 to 255 is returned. This is formed by the two least significant bits of the terminator (60 to 63) and the six bits of the following character. In all cases of unsuccessful recognition, a jump is made to the label FT.

A detailed flowchart is given for this procedure.

Label RO:

This section is entered, via the switch ULWORD, when one of the relational operator symbols ('GT','LE','NE','EQ','LT' or 'GE') is recognised. The value assigned to RELOP is calculated from the symbol number, held in Ti, which is set to 4. A jump is then made to OK.

Label SC:

This section is entered, via the switch CHARACTER, when a semicolon, occurring as a terminal symbol, is read. SEMICOM is called to process any following comment, and a jump is made to EX.
'PROCEDURE' SEMICOM;

This procedure is called to check for, and to skip over, bracketed comments following the semicolon symbol. One or more of these comments may follow the symbol, and these are detected by the outer 'FOR' loop, which reads the opening bracket, if present. The inner 'FOR' loop counts the bracket level \( B \), as a side effect of the 'WHILE' element, and terminates when the matching closing bracket is found. On return from the procedure, \( T_2 \) will contain the character following the last closing bracket.

Label SH:

This section is entered, via the switch ULWORD, when one of the shift operator symbols ('SRA', 'SLA', 'SRL' or 'SLC') is recognised. The required function code number is calculated, using the value associated with the symbol which is held in \( T_1 \), and stored in FUNCTION. \( T_1 \) is then set to 3, to indicate a shift operator, and the section OK entered.

Label SP:

This section is the starting point of the reader. If a valid character has been left in \( T_2 \) by a previous call, then this character is copied into \( T_1 \), otherwise READ is called and its result stored in \( T_1 \). This character is then used to select an entry in the switch CHARACTER. This jumps to the section appropriate to the character. The characters may be categorised as follows:

1) Characters which are terminal symbols (e.g. +)
2) Characters which start terminal symbols (e.g. &)
3) Characters which may alone be terminal symbols, or may be part of a composite symbol (e.g. <)
4) Illegal Characters in this context (! and .)
5) Ignored characters (Space)

As the entry for Space causes this section to be re-entered, all Spaces supplied by READ are effectively ignored. It should be noted that \( T_2 \) must never contain the code for Space.

Label ST:

This section is entered, via the switch CHARACTER, when an (opening) quote is read. The string is assembled, character by character, in the MCID area at the base of the stack using ADDCHAR, after this area has been initialised as an empty string. Paired quotes cause a single quote to be stored, and the string is terminated by a single unpaired quote. The string is then moved from the MCID area to the NAME area by MOVESTRING. The terminal symbol value of 5, the value for a string, is set into \( T_1 \); and a jump made to EX, which does not delete the character in \( T_2 \).
Label ZERO:

This section is entered from NORM and INTX when it is discovered that NUMBER is zero. For optimisation purposes a zero constant, whatever its form, is treated as a special case syntactically. T1 is set to zero to indicate this, and NUMBERSCALE is set to indicate an integer constant requiring one significant bit only (sign bit). A jump is then made to EX.
COMPILER STRATEGY

Introduction

The strategy of the compiler is intimately tied up with the "working syntax" used to drive the syntax analyser. This syntax is based on the syntax of the Official Definition after it has been processed by Foster's SID (syntax improving device) program to reduce it to a "one track" form. This process introduces a number of auxiliary rules into the syntax, and these are named by appending one or more digits after the letter Q.

Into this syntax are embedded the names of the compiling actions required to be carried out at that point. The syntax has also been expanded in places to enable optimisations to be carried out. Some use has also been made of Simpson's "Semantically Selected" rules. In this case the alternative chosen at the start of the rule depends not on the incoming terminal symbol, but upon the result returned by a previous compiling action.

In each section of the strategic description the relevant part of the syntax will be quoted. The basic notation used is that of SAG, but with the following identifier conventions.

1 Rule Names
   First letter in upper case, subsequent letters in lower case. The rule name is underlined where semantic selection is carried out.

2 Terminal Symbols
   Lower case letters.

3 Compiling Actions
   Upper case letters. The action name is underlined where it delivers a result to be used as a semantic selector.

Recapitulations will be enclosed in square brackets and dots used to denote alternatives irrelevant in the context.

Compiler Activation

Run=(Compileitem,Q70)
Q70=(finish)
    (semi,Run)

When the compiler is first loaded and entered, it initialises certain words of workspace, and then enters the syntax analyser. The first rule is named Run, and represents a list of items separated by semicolons and terminated by the symbol 'FINISH'. These are processed one by one, the compiler terminating awaiting re-activation after the last.

Segments

Compileitem=()
   (Program)
   (Commondec)
   (library,Q69)
   (external,orb,Extlist,crb)
   (absolute,orb,Abelists,crb)
   (Leveldec)
   (Yesno,Trace,trace)
   (test,SETTEST,Q68)
Thus the list of items to be compiled may be split into three main categories:

1. Segments to be compiled
2. Communicators specifying identifiers external to a segment
3. Trace and diagnostic directives

Segments may be processed in a test mode by prefixing them with the symbol 'TEST'. This causes the action SETTEST to be executed which inhibits the paper tape output for the duration of the following segment.

Program Segments

Program=(id,BEGINPROG,begin,Body,ENDPROG,end)

Body=(D1,ENDDECS,S1)

D1=(STARTDEC,D,semi,Q50)

Q50=()

D1)

After the initial identifier giving the name of the segment, the action BEGINPROG is executed. This uses the identifier as part of the segment header output on tape. This must be followed by the symbol 'BEGIN', one or more declarations (each being terminated by a semicolon), one or more statements (separated by semicolons), and finally the symbol 'END'. Immediately before this final symbol the action ENDPROG is executed.

Before each declaration is processed the action STARTDEC is executed, and after the final declaration the action ENDDECS is executed. These are both involved with the housekeeping of storage allocation.

Declarations

D=(Type,Q47)

(Tabled,Presetlist)

(switch,ADDRSW,TYPE SWITCH,Newid,becomes,Swlist)

(Specialdec)

(Overlaydec)

(Valproc,Newid,BEGINPROC,colon,Newid,NEXTPSET,Procrest)

(SETYES,TYPEPROC,Q48)

(no,SETNO,Tracetype,trace)

(Levldec)

Thus a declaration list may consist of:

1. Data declarations, which may be preset in certain cases.
2. Overlayed data declarations which may not be preset.
3. Switch declarations.
4. Procedure declarations.
5. Trace and diagnostic directives.
It should be noted that at this point in the analysis, a declaration of a typed procedure cannot be distinguished from a simple data declaration, this being resolved in rule Q47. Similarly a declaration of an untyped procedure cannot be distinguished from a 'PROCEDURE' 'TRACE' directive, this being resolved in rule Q49 which is entered via rule Q48.

The declaration of each identifier causes an Identifier Specification Record to be generated, and placed on the compiler's main stack. This record is generated at the base of the stack, and consists of five words plus the string of the identifier. This information is built up incrementally using the overlayed names of the locations. Thus the information which will occupy the TYPEBITS word of the record on the stack is built up in the location named IDTYPE, chiefly by the actions named TYPE----. This information is usually complete by the time the identifier is read, and can be moved bodily onto the stack.

Data Types

[D1=(STARTDEC,D, . . .
    D=(Type,Q47)
    . . . . . . ]

Type=(integer,Q59)
    (fixed,orbosb,int,NOBITS,comma,Sgint,NOAFTER,crbcosb)
    (floating,TYPEFLOAT)

Q59=(osb,int,NOBITS,csb,PARTINT)
    (TYPEINT)

Sgint=(int)
    (plus,int)
    (minus,int,NEGNUM)

The action STARTDEC clears IDTYPE prior to a number type being read. If the following symbol is 'FLOATING' the action TYPEFLOAT is executed, which sets the TYP field of IDTYPE to 2, the PVL field being left clear.

Alternatively if the following symbol is 'FIXED' this must then be followed by the specification of the number of significant bits, and the number of fractional bits. At the request of Ferranti this specification may be enclosed in either round or square brackets. ( orbosb = (or [ , crbcosb = ] or ] ). After the integer constant specifying the number of significant bits, the action NOBITS is executed. This inserts the number of significant bits into the SGB field of IDTYPE. This is followed by a comma and a (signed) integer giving the number of fractional bits, which may be zero or even negative. (If the number is preceded by a - sign it is negated by NEGNUM). The action NOAFTER is then executed, which sets the TTP field of IDTYPE to 1, and calculates the scaling factor PVL.

The third possibility is the symbol 'INTEGER'. This may optionally be followed by a specification of the number of significant bits. This is only required where an integer is to be used in mixed arithmetic. In this case NOBITS is executed. This integer may only be enclosed in square brackets, as an ambiguity may arise concerning typed primaries if round brackets were allowed. After the closing bracket the action PARTINT is executed. This is in fact the same action as TYPEINT, which inserts the scale for integer into the PVL field of IDTYPE, (#67, P23) and leaves the TTP field set to zero.
The rule Type is used in many places in the syntax, and in all cases the relevant fields of IDTYPE will have been cleared before its use, often by the use of CLEARTYPE.

Simple Data Declarations

\[ D = (\text{Type}, Q47) \]

\[ Q47 = (Q21, \text{Presetlist}) \]

\[ (\text{TYPETPROC}, \text{procedure}, \text{Newid}, \text{BEGINPROC}, \text{Procrest}) \]

\[ Q21 = (\text{Idlist}, \text{DECSIZE}) \]

\[ (\text{array}, \text{TYPEARRAY}, \text{Arraylist}) \]

\[ \text{Idlist} = (\text{Newid}, Q52) \]

\[ Q52 = () \]

\[ (\text{comma}, \text{Idlist}) \]

\[ \text{Newid} = (\text{id}, \text{NEWNAME}) \]

This set of rules states that a declaration of one or more simple variables consists of a data type followed by a list of identifiers, separated by commas.

The rule Type sets up IDTYPE and the action STARTDEC has previously inserted the current data address into DIRADD[0]. After each identifier, the action NEWNAME is executed. This checks that the identifier is unique at this block level, and moves the Identifier Specification Record onto the main stack. After this, it then increments DIRADD[0].

At the end of the list of identifiers, DECSIZE is executed, which copies the current value of DIRADD[0] back into DATAMAX, and hence allocates the storage for the variables.

Presetting Data Declarations

\[ Q47 = (Q21, \text{Presetlist}) \]

\[ (\text{becomes}, \text{SKIPDTA}, \text{Presets}) \]

\[ \text{Presets} = (\text{Presetnum}, Q40) \]

\[ Q40 = () \]

\[ (\text{comma}, \text{Presets}) \]

\[ \text{Presetnum} = (\text{Pnumber}, \text{OUTPRESET}) \]

\[ (\text{orb}, \text{Presets}, \text{orb}) \]

\[ (\text{string}, \text{PRESETSTRING}) \]

\[ \text{Pnumber} = (\text{pconst}) \]

\[ (\text{plus}, \text{pconst}) \]

\[ (\text{minus}, \text{pconst}, \text{NEGNUM}) \]

\[ (\text{ZERONUM}) \]
Data declarations occurring at the head of a program may be preset. In this case the declaration is followed by a becomes (←, :=) symbol and a list of preset values. These are separated by commas and may be arbitrarily grouped within matched round brackets, to any depth. The number of items in the list must not exceed the number of items declared in the declaration.

When the becomes symbol is read, the action SKIPDATA is executed. This ensures that the loader will skip over any unpreset locations in the data area. The preset values are output by OUTPRESET after rescaling to the required scale. If this is not possible an error message is output. (Integers cannot be preset with real constants.) With the above restriction, any type of number, signed or unsigned, may be included in a list of presets. If a number is omitted (e.g. two commas occur in succession) ZERONUM is executed to give the effect of a zero constant.

As an extension, quoted strings may occur in the preset list. In this case the variable is preset with the address of the string which is output by PRESETSTRING. This should only be used to preset integers. This gives the same effect as the assignment of a string.

Array Declarations

\[ D = (\text{Type}, Q_{47}) \]

\[ Q_{47} = (Q_{21}, \text{Presetlist}) \]

\[ Q_{21} = (\text{array}, \text{TYPEARRAY}, \text{Arraylist}) \]

\[ \text{Arraylist} = (\text{ZEROARRAYS}, \text{Idlist}, \text{osb}, \text{Boundpair}, \text{Morebounds}, \text{csb}, \text{ENDARRAY}, Q_{42}) \]

\[ Q_{42} = () \]

\[ (\text{comma}, \text{Arraylist}) \]

\[ \text{Boundpair} = (\text{Sgint}, \text{SETLBcolon}, \text{Sgint}, \text{SETUB}) \]

\[ \text{Morebounds} = (\text{ONEDIM}) \]

\[ (\text{FIRSTDIM}, \text{comma}, \text{Boundpair}, \text{Arraytail}) \]

\[ \text{Arraytail} = (\text{LASTDIM}) \]

\[ (\text{MIDDIM}, \text{comma}, \text{Boundpair}, \text{Arraytail}) \]

When the symbol 'ARRAY' is read following a data type, the action TYPEARRAY is executed, which sets the AYN marker of IDTYPE. This ensures that the Identifier Specification Records of the following identifiers are flagged as being those of arrays, and will require subscripting in normal usage.

As a number of groups of arrays of the same data type but with differing dimensions and bounds may be declared within one declaration, each group is treated as virtually being a separate declaration by the rule Arraylist.

ZEROARRAYS is first executed, which clears the location ARRAYS used for counting numbers of arrays. (This is overlayed on DECLIST[0] and is incremented by NEWNAME.) The list of array identifiers is read by Idlist, and the Identifier Specification Records created, which do not at this stage have their address field correctly set, placed on the stack. The identifier(s) are then followed by a list, enclosed in square brackets, of pairs of integers
giving the bounds of each dimension of the array(s). These integers may be
signed, and are separated by a colon. The second must not be less than the
first. After the first integer SETLB is executed, and after the second of the
pair, SETUB. These calculate the size and offset of the dimension.

If the array has only one dimension, ONEDIM is then executed. This inserts
the addresses of the arrays into the Identifier Specification Records on
the stack.

If the array has more than one dimension, the action FIRSTDIM is executed
after the first pair of bounds. This inserts the addresses of the first
level Iliffe vectors into the Identifier Specification Records. After each
subsequent pair of bounds except the last, MIDDIM is executed. This outputs
(to Special Data) the Iliffe vector referred to by the previous level, and
referring to the next level. After the final pair of bounds LASTDIM is
executed. This outputs the final Iliffe vector, which refers to the actual
data addresses of the arrays.

At the end of a group of arrays ENDARRAY is executed. This uses the total
data requirement (stored in ARrys) to allocate data space.

Expressions

The compilation of expressions forms a very important part of the operation
of the compiler, and probably accounts for the vast majority of the code
generated. A very loose description of an expression is a collection of
operands separated by operators. The syntax rules which follow express this
in more detailed and rigorous terms. The priority of the operators expressed
by these rules is as follows:

<table>
<thead>
<tr>
<th></th>
<th>1</th>
<th>2</th>
<th>3</th>
<th>4</th>
<th>5</th>
<th>6</th>
<th>7</th>
</tr>
</thead>
<tbody>
<tr>
<td>[Highest]</td>
<td>Shifts</td>
<td>'SRA', 'SLA', 'SRL', 'SLC'</td>
<td>Logical AND</td>
<td>'MASK'</td>
<td>Logical OR</td>
<td>'UNION'</td>
<td>Non-Equivalence</td>
</tr>
<tr>
<td></td>
<td>5</td>
<td>Exponentiation</td>
<td>'!', '**'</td>
<td>7</td>
<td>Multiply/Divide</td>
<td>'*'</td>
<td>6</td>
</tr>
<tr>
<td>[Lowest]</td>
<td>7</td>
<td>6</td>
<td>5</td>
<td>4</td>
<td>3</td>
<td>2</td>
<td>1</td>
</tr>
</tbody>
</table>

Within the compiler each of these operators has its associated compiling
action. These operate on the two top Arithmetic Operand records on the stack,
generate the appropriate instructions, and leave one Arithmetic Operand record
on the Stack which gives the details of the result. These actions are
invoked as soon as possible (bearing in mind the priorities of the operators),
thus there is no reordering or regrouping of expressions. This is not very
disadvantageous on a multi accumulator machine, and allows the programmer
control over the order of evaluation.

Each expression is assigned an accumulator (PREFACC) in which the result is
to be evaluated if possible. Further, if the scale to which the expression
is to be evaluated is known, a marker (SCALEFIRM) is set, and the required
scale placed in EXPSCALE. Otherwise the scale to which the expression will
be evaluated depends only on the operands.
1 Simple Expressions

Expr=(Cexpr)
   (Axpr)
      (zero,CONSTANT,SCALETERM)
      (string,STRINGEX,SCALETERM)

Axpr=(Term,Q6)
   (Axpra)

Axpra=(plus,Term,Q6)
   (minus,Term,UNARYMINUS,Q6)

Q6=()
   (Q6a)

Q6a=(plus,Term,ADD,Q6)
   (minus,Term,SUB,Q6)

Ignoring, for the moment, the special cases, an expression consists of at least one Term, optionally preceded by a unary + or -, possibly followed by one or more terms, each being preceded by a + or - sign.

If the first term is preceded by a plus sign, this may be ignored. If however it is preceded by a minus sign, the Term requires negation, after evaluation. The action UNARYMINUS is executed in this case.

After each subsequent Term, the appropriate action ADD or SUB is executed. These output instructions for addition and subtraction.

2 Terms

Term=(Factor,Q5,SCALETERM)

Q5()
   (mult,Factor,MPY,Q5)
   (div,Factor,DIVIDE,Q5)

These rules give the next level of expression. It should be noted that the action SCALTERM is executed after each Term. In cases where no required scale is demanded, this action determines the scale to which the expression will be evaluated, after examining each term in turn. In cases where a scale is mandatory, its action is limited to the output of instructions for fixed to floating, and floating to fixed conversions. It is preferable that these conversions are performed as soon as possible, as they involve calls on Library procedures, and may require the temporary storage of intermediate results.
3 **Factors**

Factor = (Difference, Q24)

Q24 = ()
    (power, Difference, RAISE, Q24)

This rule introduces the exponentiation operation into this implementation. Care should be taken however, as unless the right hand operand is the integer constant 2, the result will be in floating point form, and may cause the rest of the expression to be evaluated by floating point arithmetic. The special case is treated as squaring, and is performed by multiplication.

4 **Word Logic**

Difference = (United, Q4)

Q4 = ()
    (differ, United, NEQ, Q4)

United = (Collation, Q3)

Q3 = ()
    (union, Collation, ORF, Q3)

Collations = (Tert, Q2)

Q2 = (mask, Tert, MSK, Q2)

These rules govern the priority of word-logic operators. It should be noted that the results of these operations are regarded as being of type integer (of unspecified significance) regardless of the types of the operands, which strictly should be typed primaries.

5 **Shifts**

Tert = (Sec, Q1)
     (Pri, Q1)

Q1 = ()
    (shift, SETSHIFT, Shifts, DOSHIFT, Q1)

Shifts = (Sec, PLUSSUB)
     (Tpri, SCALTERM, PLUSSUB)
     (orb, Subex, orb)

Because of the nature of the shift functions, the right hand operand must be of type integer. Furthermore, if it is a variable or an expression, it is required to be held in a modifier before use. Thus it may be seen that the operand giving the number of places of shift has a similarity to a subscript, and in this implementation uses many compiling actions in common. The left hand operand may be of any type, and the result is regarded as being of type integer.
6 **Word Slicing**

Sec=(Bitset,BITSIN,Pri,RHSBITS)

Bitset=(bits,CLEARTYPE,osb,int,Oneormore,FIELDPOSN,UNSFIELD,csb)

Oneormore=(NOBITS,NOSIG,PARINT,comma,int)

(ONEBIT)

The compiling actions in the rules Bitset and Oneormore cause a data specification to be calculated in IDTYPE, and a part word specification to be calculated in BITSPEC, in a similar manner to those required for an unsigned integer table field. These values are stored by BITSIN for subsequent use by RHSBITS. This latter action does not usually involve the generation of any instructions at this point. Unless the PARTWORD field of the Arithmetic Operand of the Primary is non-zero, the two values are simply inserted into the TYPEBITS and PARTWORD fields of the record.

If instructions have to be generated at this point, it usually indicates bad programming practice, e.g. taking a part of a table field which is itself a part of a word. (This should be avoided by specifying a field consisting of the relevant bits.)

The result of this "operation" is of type integer, having the number of significant bits (plus one for the (zero) sign bit) given in the bit specification. Where the number of bits is one, this (and the following comma) may be omitted, thus allowing 'BIT' [9] etc.

7 **Untyped Primaries**

Primaries are the building blocks from which expressions are constructed, operators being the mortar. These may be divided into two main categories, typed and untyped. The untyped primaries do not have an intrinsic scale, this having to be determined by the compiler.

Within the compiler, the net result of processing a primary is to leave on the stack an Arithmetic Operand record giving type and address information for the primary. (This is usually also the case when processing the special cases of expression.)

Pri=(Tpri)

(realcon,CONSTANT)

(orb,STACKEXPR,SETPREF,Expr,crb,OFFSTEXPR)

In the case of a Real constant (ie containing . or & ) there is no definite scale associated with the constant. A scale may however be chosen which has the minimum number of integer bits required to hold the number. This scale is supplied by READER, and is used by CONSTANT when setting up the Arithmetic Operand record. As constants are a special case of this type of record, the marker CON (m.s. bit) of TYPEBITS of the record is set to denote that the DIRADD field contains the actual constant and not an address. After any subsequent rescaling, when this record is used as an operand record for the output of an instruction, this constant is converted to an "address" by OUTWCONST.
In the case of an untyped bracketed expression, the scale to which it is evaluated is determined by the terms of the expression itself, following the rules for expressions occurring in an undefined context, irrespective of the scale of the expression of which it forms a part. (It should however be noted that a bracketed expression following a shift operator is implicitly typed Integer.) The action STACKEXPR is executed to set up a new "level" of expression evaluation. As the scale to which this level is to be evaluated is undefined, SCALEFIRM is cleared, and EXPSSCALE initially set to zero. The preferred accumulator PREFERENCE, which will be used if possible for the evaluation of the expression, is set by SETPREF. After processing the following expression, the action OFFSTEXPR is executed, which reduces the expression level by restoring the variables concerned with the evaluation of expressions to their previous values, and also "floats" down the stack the Arithmetic Operand record left by the expression.

8 Typed Primaries

\[ \text{Pri}(\text{Tpri}) \]

\[
\text{Tpri}=(\text{intcon}, \text{CONSTANT})
\]

\[
(\text{id}, \text{LOOKUP}, \text{Q28})
\]

\[
(\text{NONAME}, \text{Singlesubs})
\]

\[
(\text{CLEARTYPE}, \text{Type}, \text{TYPEEXPR}, \text{orb}, \text{Expr}, \text{crb}, \text{EXPRTYPE})
\]

\[
(\text{location}, \text{orb}, \text{Locvar}, \text{LOCAT}, \text{crb})
\]

\[
(\text{label}, \text{orb}, \text{id}, \text{Q32}, \text{crb})
\]

Q28= (Subscript, VARCHCHECK)

(RHSPTEST, Rhapsel)

Rhapsel= (VARCHCHECK)

(Proccall)

The simplest case of a typed primary is a (non zero) integer constant. In this case CONSTANT is executed to set up an Arithmetic Operand record, as it is for real constants in untyped primaries, but in this case however, the scale supplied by READER is of integer typed, but with the number of significant bits specified in SGR.

An identifier occurring as a primary may be the identifier of either a simple variable, an array or table (field) name, or the name of a procedure. Before this is resolved, LOOKUP is executed. This action searches through the whole of the list of Identifier Specification records until one is found with a matching identifier string. The first five words of this record are then copied onto the top of the stack to form the Arithmetic Operand record, and the first word of this (SPIEL) is set to point at the identifier string on the stack.

If the identifier is followed by a subscript (starting with [ ) rule Q28 selects the first alternative. The actions associated with subscripts check that the identifier may validly be subscripted, and VARCHCHECK check that its type is Arithmetic. (It could have been a switch identifier). Otherwise the action RHSPTEST is executed. This returns a result of one if the identifier is that of a typed procedure, and zero otherwise. This value is used to select the appropriate alternative of Rhapsel. If the result supplied by RHSPTEST is zero, VARCHCHECK is executed to check the type of the identifier further (switch and untyped procedure illegal) and that no subscript is required (ATM clear). Otherwise, if the result is one, the rule Proccall is selected.
An anonymous reference consists of a single expression enclosed in square brackets. In this case NONAME is executed, which sets up an Arithmetic Operand record, with no name or address information, but of type integer array. The actions associated with subscripts will insert the address information, and delete the array marker, leaving the type set to integer.

Where a bracketed expression is preceded by a data type, the expression will be evaluated to the type specified. CLEARTYPE is executed before the type information is processed by Type, this clearing IDTYPE in readiness. The action TYPEXPR is then executed, which sets up a further expression level, in a similar manner to STACKEXPR, but setting SCALEFIRM to denote that EXPSCALE (which is set from IDTYPE) contains a mandatory scale. At the end of the expression, EXPRTYPE reduces the expression level in a similar manner to OFFSTEXPR, but only after ensuring that the expression has been evaluated to the required scale.

9 Address Primaries

[Tpri= . . .
 (location,orb,Locvar,LOCACT,crb)
 (label,orb,id,Q32,crb)]

Locvar=(id,LOOKUP,Q29)
(NONAME,Singlesubs)

Q29=(Subscription)
(AYMCHECK)

Q32=(LOOKUP,Singlesubs,LABSK)
(LABL)

The address of any item of data may be obtained by the use of 'LOCATION'. (In this implementation this includes special data, and in particular the locations of pointers to procedures, and of entries in switch lists.)

Although the type of the identifier is immaterial, it must be subscripted if a subscript is normally required, and unsubscripted otherwise. The actions associated with subscripts check the first case, and AYMCHECK the second. After this, the action LOCACT is executed. This changes the type of the Arithmetic Operand record to Integer, and except in certain optimised cases, sets the marker LCM to denote that the address itself is to be used as the operand instead of a reference to an operand.

As an extension, the address of a labelled statement may be obtained by the use of 'LABEL'. This must be followed by an opening round bracket and an identifier. At this point it is not yet determined whether the identifier is that of a label (which requires looking up in the label list) or that of a switch (which requires looking up in the main declaration list). This is resolved in rule Q32. In the case of a label, no subscript follows, and the action LABEL is executed. This looks up the label in the (local) label list, and forms an Arithmetic Operand of type Integer for it. This has the marker LBM set to denote that the direct address contains a pointer to the label record (at this point the label may not be set), and the marker LCM also set. This reference to a label record will subsequently be converted to an "address" by USELAB.
If the identifier is subscripted, LOOPUP is executed, and the subscript processed by Singlesubs. The action LABSK is then executed, which checks that the Arithmetic Operand record is that of a switch, and changes its type to Integer. Note that LCM is not set, as the label address will be copied from the switch list.

\[\text{('LABEL'(S[K])} \equiv \text{['LOCATION'(S[K])])} \]

**Special Cases of Expressions**

\[\text{Expr=} \ldots \]
\[\text{(zero,CONSTANT,SCALETERM)} \]
\[\text{(string,STRINGEX,SCALETERM)}\]

These two cases are chosen because, although they are similar to typed primaries, their explicit use in conjunction with an arithmetic operator is extremely obscure.

The detection of a zero constant as a special case enables useful optimisations to be performed in Assignments and Comparisons. In cases where these optimisations are not required zero is treated as an integer constant by CONSTANT, and its type (trivially) changed to that required by the context by SCALETERM.

The inclusion of strings is principally to allow them to be used as parameters of procedures, there being no operators which act on the individual characters of a string. The action STRINGEX outputs the string and sets up an Arithmetic Operand record of type integer, with the LCM set so that the address of the string is obtained.

**Subscripts**

\[\text{Singlesubs}=(\text{Substart,csb,KILLEXPR})\]
\[\text{Subscript}=(\text{Substart,Moresubs,csb,KILLEXPR})\]
\[\text{Substart}=(\text{osb#,SETUPSUB,Subex})\]
\[\text{Moresubs=}()\]
\[\text{(comma,SUBCOMMA,Subex,Moresubs)}\]

Arrays may have more than one dimension, but switches and anonymous references are strictly one-dimensional. This is reflected in the first two rules above.

After the opening square bracket has been read, the action SETUPSUB is executed. This checks that the Arithmetic Operand record on top of the stack is flagged as requiring a subscript. A new level for expression evaluation is set up, requiring a mandatory integer scale, the preferred accumulator being a modifier.

After each comma separating subscript expressions, the action SUBCOMMA is executed. This outputs an instruction to copy an entry in an Iliffe vector into a modifier. In this way, any number of dimensions may be processed.
At the end of a subscript, the action KILLEXPR is executed to reduce the expression level. (Note that this discards any result of the expression, this will be explained below.)

\[
\text{Subex} = (\text{Subterms})
\]

\[
\quad= (\text{Cexpr,PLUSSUB})
\]

\[
\quad= (\text{Term,PLUSSUB,Subexcont})
\]

\[
\quad= (\text{zero})
\]

\[
\text{Subexcont} = ()
\]

\[
\quad= (\text{Subterms})
\]

\[
\text{Subterms} = (\text{plus,Term,PLUSSUB,Subexcont})
\]

\[
\quad= (\text{minus,Term,MINUSSUB,Subexcont})
\]

A subscript expression is not evaluated as a normal expression, but is processed Term by Term by the actions PLUSSUB and MINUSSUB, which both use the procedure SUBTERM. This optimises subscript expressions containing only one integer variable and constants for a one dimension non-parametric array, or only integer constants for a parametric or two dimensional array. As each Term is processed, its record is discarded, as the information resulting from the Term is stored in the address fields of the record of the operand being subscripted. Thus there is no Arithmetic Operand record left on the stack as the result of a subscript expression. Note that in the case of a subscript expression consisting of a zero constant, no action need be executed.

**Conditional Expressions**

\[
\text{Expr} = (\text{Cexpr})
\]

\[
\quad= (\text{if,IFEX,Conditionlist,then,THENEX,Expr,else,ELSEFIFIEX,ELSEXFIFIEX})
\]

As a conditional expression is an alternative case of expression it does not need to set up a further level for expressions, even where a conditional expression occurs within a conditional expression. A further level will however be temporarily set up for the evaluation of the conditions(s).

If the expression level at which the conditional expression is to be evaluated has SCALEFIRM set, i.e. a definite scale is required, both the consequence and alternative expressions will be evaluated to this scale. If however, it is not initially set, it will be set after the consequence expression has been evaluated, to ensure that the alternative expression is evaluated to the same scale. In either case however, the result of each expression will be evaluated in the accumulator specified by PREFACC.

After the symbol 'IF' has been read, the action IFEX is executed. This sets up a level for conditionals, and dumps any Arithmetic Operands existing only in accumulators. The following condition(list) is then processed, and the symbol 'THEN' read. The action THENEX is then executed, this being concerned with the output of a jump over the consequence expression, and setting the DUECHAIN if the symbol 'OR' occurred in the conditionlist.
The consequence expression is then processed, and ELSEFI executed after the symbol 'ELSE' is read. This action ensures that the result of the expression is in the accumulator specified by PREFACC. The action ELSEX is then executed, which outputs a jump over the alternative expression, sets the chain for the jump to the alternative expression, and sets SCALEFIRM. After the alternative expression has been processed, ELSEFI is again executed. The final action is FIEX. This reduces the conditional level, and creates an Arithmetic Operand record for the result of the conditional expression. This record will have its TYPEBITS set from EXPScale, and its direct address will be that of the accumulator specified by PREFACC.

**Conditions**

Conditionlist=(Condition,Q31)

Q31=()
   (and,OUTCJ,Conditionlist)
   (or,ORACT,Conditionlist)

Condition=(Yesno,overflow,OVRTEST)
   (STACKEXPR,ANYPREP,Locnd,Rcond,Relation)

Yesno=(no,SETNO)
   (SETYES,TYPEPROC)

In the interests of efficiency, it is required that conditions are evaluated only as far as is necessary to determine their truth or falsity. Thus, following the occurrence of the symbols 'AND' and 'OR', the associated actions output the appropriate jump instruction to test the preceding condition. This enables inconsequential tests to be skipped.

The above syntax does not need to express the relative priority of these two "Boolean operators", as this is obtained as a consequence of the operation of the tests output by the compiling actions.

Following the occurrence of the symbol 'AND', the action OUTCJ is executed. This outputs an if-false-jump instruction to skip over the following condition, or conditions if more 'AND's follow. Unless the symbol 'OR' intervenes, the jump will be made to the alternative, if any.

Following the occurrence of the symbol 'OR', the action ORACT is executed. This outputs an if-true-jump, directly to the consequence (statement or expression as appropriate). The destination of any preceding 'AND' jumps is then set to the following condition.

In this implementation an extra type of condition is introduced, this being a (destructive) overflow test. The symbol 'OVERFLOW' ('OVR') may be preceded by 'NO', in which case SETNO is executed setting ACCUMULATOR to 0, or 'NO' may be absent in which case SETYES is executed setting ACCUMULATOR to 1. The execution of TYPEPROC is irrelevant in this context.
Following the occurrence of the symbol 'OVERFLOW', the action OVRTEST is executed, which sets FUNCTION to #24. Thus it may be seen that tests are prepared as if-false-jumps, this being the test normally required following 'THEN', and are reversed as required (by REV CJ).

**Comparisons**

```plaintext
[Condition= . . .
 (STACKEXPR, ANYPREP, Lcond, Rcond, RELATION)]
```

Lcond=(Aexpr)
  (zero, CONSTANT)

Rcond=(r0, Q30)
  (SETNEZ)

Q30=(Condterm, Condterms)
  (zero, ZEROCOMP)

Condterm=(Term, CONDPLUS)
  (plus, Term, CONDPLUS)
  (minus, Term, CONDMINUS)

Condterms=()
  (plus, Term, CONDADD, Condterms)
  (minus, Term, CONDSUB, Condterms)

Instead of evaluating the expressions on either side of the comparator, and the performing the comparison by means of a subtraction, the two expressions are evaluated as one expression. In a simple case this is performed by effectively reversing the signs of all the terms of the right hand expression. This results in less instructions being generated, and usually one less accumulator being used. As the scale to which the expression is to be evaluated is not specified, it is determined by the compiler, using the normal method (see SCALETERM).

The expression to the left of the comparator is processed by the rule Lcond, and that on the right hand side by Q30. The first Term of this second expression causes either CONDPLUS or CONDMINUS to be executed, and subsequent Terms cause either CONDADD or CONDSUB to be executed. If however this second expression is a zero constant, the action ZERO COMP is executed instead. This optimises cases of comparison against zero.

In this implementation, the comparator and second expression may be omitted. This allows the use of "pseudo" Boolean expressions, zero being false, and non-zero true. Thus the symbols <>O are effectively assumed. This allows such conditions as:

'IF' 'BIT'[9]X 'THEN' . .

In this case the action SETNEZ is executed.

When the comparison is complete the action RELATION is executed. This sets up ACCUMULATOR and FUNCTION in preparation for a subsequent call of OUT CJ, and optimises tests of partwords against zero.
After processing each statement in the body of a program, the action ENDST is executed. This frees any store used temporarily during the statement. Statements are separated by semicolons (Q39), and may be void (S).

Where a statement commences with an identifier this may be a label, or the start of an assignment or procedure call. The first case is resolved in rule Q38. If the identifier is followed by a colon (:= having been converted to ←) the action SETLAB is executed before the colon is read. This action looks up the record of the label in the label list, and inserts the current address in this record, setting a marker to denote that this label is now set.

### Assignment Statements

[S= . .
 (id,Q38)
 (STATUSCHECK,Q26,A1)
 . . . .

Q38= . . .
 (LOOKUP,STATUSCHECK,Q33)]

Q33=(Q25,A1)
 (LHSPROC,Proccall,BEHEAD)

Q25=()
 (Subscript)

Q26=(Bitset,Lhsb)
 (NONAME,Singlesubs)

Lhsb=(id,LOOKUP,LHSBITS,Q25)
 (NONAME,LHSBITS,Singlesubs)

A1=(becomes,VARCHCK,SETASS,A2,STORE,BEHEAD)
The rules for the left hand side of an assignment are complicated by the necessity to remove the cases of label settings and procedure calls, and the number of possibilities available.

The action STATUSCHECK is executed once, irrespective of which route is chosen, this completes any deferred actions associated with conditions and Answer statements, and checks that the statement can be entered.

An arithmetic operand record will have been set up, either by LOOPUP if an identifier has been processed, or by NONAME if an anonymous reference has been used. If the identifier was subscripted, or an anonymous reference has been used, the address fields of this record will have been processed by the actions associated with subscripts, and in a non-optimum case code may have been generated to evaluate the subscript expression(s).

If 'BITS' has been used the PARTWORD field of the record will have been set to the partword specification generated by Bitset, and after checking, its type will have been changed to integer, by LESBITS. (Note that the assignment to a part of a partword table field is illegal.)

\[ A1 = (\text{becomes}, \text{VARCHECK}, \text{SETASS}, A2, \text{STORE}, \text{BEHEAD}) ] \]

\[ A2 = (A3) \]

(Expr)

A3 = (Cexpr)

(string, STRINGEX)

(zero, STOREZERO)

(Term, ADDA, A4)

(minus, Term, SUBA, A4)

A4 = (SIMPLEASS)

(SELOPTA, A5)

A5 = (Q6a)

(Axpra)

After the left hand side of an assignment statement has been processed, the symbol becomes is read and the action VARCHECK executed. This ensures that the type of the left hand side is a data type, and that a subscript has been supplied where required.

The action SETASS is then executed. This sets up the variables required for an assignment, and also sets up the lowest expression level to the type of the left hand side. If 'ASSIGNMENT' 'TRACE' is required, the right hand expression must be evaluated in accumulator 7 and assignment optimisation inhibited. In this case SETASS returns a result of 1 to be used as a selector by A2.

If trace is not required, SETASS returns a result of zero, which causes rule A2 to select rule A3.

The assignment of zero is a special case which causes STOREZERO to be executed. Strings and conditional expressions are processed in the normal manner. The remaining cases are equivalent to those represented by the rule Axpra.
After the first Term has been processed its Arithmetic Operand record is compared with that of the left hand side, by ADDA or SUBA as appropriate, to detect whether an add to store type of assignment may be used. If no more Terms follow, the action SIMPLEASS is then executed to optimise certain cases of simple assignments. Otherwise the action SELOPTA is executed. This returns a zero result if an add to store assignment is not appropriate, and the rest of the expression is processed using rule Q6a. If however, an add to store assignment is to be used, a result of 1 is returned. This causes rule A5 to select the rule Axpra, which processes the rest of the expression, without using the previously processed first Term. (The difference between these two alternatives is vital to this optimisation.)

After the right hand side of the assignment has been processed the action STORE is executed. This outputs the instructions to perform the assignment, using the instruction number set up in ASSFUN, after generating trace instructions if required. The final action BEHEAD, removes the Arithmetic Operand record of the left hand side from the stack.

**Conditional Statements**

```
ifstat = (if, IFS, Conditionalist, then, ENDST, St, else, ELSES, S, FI)
```

In this implementation, the syntax of a conditional statement differs from the Official Definition in that a conditional (or for) statement is allowed to follow 'THEN'. This avoids the necessity to use a 'BEGIN' - 'END' pair to surround the consequence statement in these cases. As processed by the syntax analyser generator, only the first alternative of the above rule is used. After checking the syntax, and generating the SYNTAX array, a one word overwrite introduces the second alternative.

```
(Overwrite #70450000 with #70450000+('SELF'+3-SYNTAX[0]))
```

This would seem to allow:

```
'IF' . . 'THEN' 'IF' . . 'THEN' . . 'ELSE' . . ;
```

to be interpreted as either:

```
'IF' . . 'THEN' 'BEGIN' 'IF' . . 'THEN' . . 'ELSE' . . 'END';
```
or:

```
'IF' . . 'THEN' 'BEGIN' 'IF' . . 'THEN' . . 'END' 'ELSE' . . ;
```

This is the pathological "dangling else" from Algol 60. As implemented, the first interpretation is invariably chosen by the compiler. In cases of doubt, the user is advised to use 'BEGIN' - 'END' to remove apparent ambiguity.

After reading the symbol 'IF' the action IFS is executed. This prepares for the processing of the following condition, setting STATUS to 4 for jump optimisation. The action ENDST, which is executed after the symbol 'THEN' is read, serves to recover any workspace used temporarily during the evaluation of the conditions. The consequence statement is then processed. If this is followed by the symbol 'ELSE', the action ELSES is executed, and the alternative statement processed. At the end of the conditional statement, the action FY is executed. This ensures that any jump around the alternative statement will subsequently have its address set.
Goto Statements

Gotostat=(goto,id,Q27)

Q27=(GOTOL)
    (LOOKUP,STATUSCHECK,Singlesubs,LABSK,GOTOSK)

The destination of a Goto may be either a label or a switch entry. This is resolved by rule Q27.

If the destination is a label, the action GOTOL is executed. This looks up the label identifier in the local label list, and optimises such cases as 'THEN' 'GOTO', as the final test instruction of the condition has not yet been output.

The second case, that of a switch entry, is treated in a manner analogous to that of an expression. An Arithmetic Operand record is created by LOOKUP, and has subscript optimisation applied to it by the actions invoked by the processing of the subscript by Singlesubs. This record is then checked for type by LABSK, and finally the action GOTOSK is executed. This action outputs the indirect jump instruction, using the output procedure normally used for Arithmetic instructions, INST. This is not surprising, as the address used for this instruction will normally lie in the (special) data area of the segment.

For Statements

Forstat=(for,STARTFOR,Locvar,CHECKCV,becomes,F1)

F1=(A3,ASSCV,F2)

F2=(F3)
    (while,Conditionlist,WHILEL,F3)
    (step,Expr,STEPUNT,until,Expr,STEPUNT,F3)

F3=(do,DOCS,S,ENDFOR)
    (comma,MOREFOR,F1)

In regard to the structure of the code generated, for statements may be split into three categories, depending upon the complexity of the for-list. The simplest, and very common case, is where there is a single step-until element with three constants. In this case the initial value is first assigned to the control variable, the controlled statement executed, the control variable incremented, and finally its value tested against the limit value. A conditional jump back to repeat the controlled statement is made if the limit is not exceeded. This is the only case where the instructions for incrementation and testing of the control variable follow those of the controlled statement in core. In all other cases they precede those of the controlled statement, thus involving the requirement for a jump round the controlled statement when the for list is exhausted, and a jump back at the end of the controlled statement.

In the second case, the for list consists of one element of any type, or two elements where the first is simply an expression and the second a while element. In this case there is only one destination required for the jump back which is made at the end of the controlled statement.
In the last case, there is more than one destination required for the jump back at the end of the controlled statement. This statement is compiled as an anonymous procedure, which is called from the code generated for the for-list, and returns using the link supplied by the call. The detection and selection of these three cases is made using the variables FORSTATE (complexity) and CCC (detection of three constants).

After the symbol 'FOR' has been read the action STARTFOR is executed, which sets up and initialises a new for level. It also sets up a block level for labels, to ensure that the controlled statement cannot be illegally entered. The control variable is then processed by Locvar, which checks that it is correctly subscripted, if required. The control variable is further checked by CHECKCV which ensures that it is either a 'FIXED' or 'INTEGER' and is not a partword. This action also sets up the variables used by assignment optimisation, and expressions.

The first expression of each for element is processed by the rule A3, which optimises the assignment, where possible, which is output by the action ASSCV.

Where the element is of the form step-until, the two expressions are processed by the rule Expr, and after each, the action STEPUNT is executed. This ensures that these expressions are converted to the type of the control variable, and are stored in temporary locations if required.

Where the for element is of the type while, the following condition is processed by the rule Conditionlist, the final jump being output by WHILEL.

Where there is more than one for element, the action MOREFOR is executed following the comma. This releases any temporary locations used by step-until elements, and resets the variables for the following expression. This action also outputs the instructions required for the call of, or jump to, the controlled statement. In the case of a step-until element, this is preceded by a test of the control variable, and followed by its incrementation.

Following the symbol 'DO' the action DOCS is executed. In cases where the controlled statement is an anonymous procedure, this outputs the instruction to store the link. If 'LOOP' 'TRACE' is required, the instructions for the printing of the control variable are output at this point.

After the controlled statement has been processed the action ENDFOR is executed. This outputs a jump back as required, or the incrementation and testing of the control variable in the single element three constant case. All store used temporarily during the execution of the for statement is now released, the label block level reduced, and the for level reduced.

Blocks and Compound Statements

Compound=(STATUSCHECK,begiQ37)

Q37=(sl,end)
(BEGINBLOCK,Body,end,ENDBLOCK)

The distinction between blocks and compound statements is that, although they both start with 'BEGIN', this is only followed by declarations in the case of a block. This is resolved by rule Q37. There are no specific compiling actions associated with a compound statement, the 'BEGIN' and 'END' merely acting as "statement brackets".
In the case of a block however, the action \texttt{BEGINBLOCK} is executed after the symbol 'BEGIN' has been read. This increases the block level, and prepares for the following declarations. At the end of the block the action \texttt{ENDBLOCK} is executed. This reduces the block level, removes any records of identifiers declared within the block, and releases any data space allocated. Note however that the value of \textt{DATAMAX} is not reset to its value prior to the block, if this value was smaller than its value at the end of the block.

Code Statements

\begin{verbatim}
[S= . .
 (code,STATUSCODE,begin,Q35)
    . . . . . . . . . . . .]
Q35=(Cs1,end)
   (BEGINBLOCK,D1,ENDDECS,Cs1,end,ENDBLOCK)
Cs1=(Cs,Q34)
Q34=()
    (semi,Cs1)
Cs=()
    (Compound,STATUSCODE)
    (inst,INSTTYPE,Npart,OUTCODE)
    (id,CODELAB,colon,Cs)
\end{verbatim}

A code statement is bracketed by 'CODE' 'BEGIN' . . 'END'. In this implementation the 'BEGIN' may be followed by declarations, causing the code statement to be treated as a block, following the usual rules. This enables local workspace to be obtained without having to resort to the use of labelled locations, this being forbidden.

After any declarations, an item in a code statement can only be a code instruction, a compound statement, or a block. These may be labelled in the usual manner. It is not possible to include constants as code items. The 'SPECIAL' 'ARRAY' facility is provided to give an alternative method of including preset constants, possibly containing address information. The advantage of this is that not only is it more efficient, (no jump over being required) but it allows programs to be run with different O and N relativiser settings.

The Npart of a code instruction is not treated as an address but as an operand, in much the same manner as a primary. The actions which process the Npart leave on the stack an Arithmetic Operand record which is used by \texttt{OUTCODE} to output the instruction. This is provided for the convenience of the Coral programmer who wishes to use code (perhaps innocently by macro) rather than for the Astral programmer who wishes to use Coral.
The syntax of the allowable Npart of an instruction depends on the type of the instruction. The class to which an instruction belongs is returned as the result of INSTTYPE and is used as the selector for the rule Npart.

Where a constant is allowed as an operand, this is treated as a reference to a constant and not as a numeric address. The constants are optimised by the loader, which uses the functions 04-07 where possible. Where a constant is prefixed with a data type, the constant is converted to this type by SCALECON before being inserted in the Arithmetic Operand record set up by CONSTANT.

Where an identifier is used as the operand of a non-jump instruction, this is used by LOOKUPD in the usual manner. This may be followed by a modifier and/or a displacement, enclosed in square brackets, this construction also being allowed as the equivalent of an anonymous reference. Thus to obtain a numeric address, this must be enclosed in square brackets in exactly the same way in which it would have to be if it were used in a normal Coral expression.

The destination of a (direct) jump instruction can only be a label identifier, or a relative jump.
Table Declarations

Tabledec=(table,TYPEINT,TYPEARRAY,Newid,osb,int,TABLESIZE,csb, Tabledetail,CLEARTYPE,TYPEINT)

Tabledetail=(osb,TABADD,Fieldlist)

Fieldlist=(CLEARTYPE,id,Fieldtype,TYPEARRAY,NEWNAME,Q60)

Q60=(csb)
    (semi,Q60a)

Q60a=(Q60)
    (Fieldlist)

After reading the symbol 'TABLE' the actions TYPEINT and TYPEARRAY are executed to set the type of the following table identifier, which is processed by Newid, to that of an integer array.

In the Official Definition both the number of words per entry, and the number of entries has to be specified. This implies that all subscripts must be multiplied by the width of the entry each time a field is referred to. This was felt to be an unjustifiable overhead for the Argus. Instead, only the total space required for the table is specified, and subscripts must be arranged to have the appropriate factor applied. This may often be done with no overhead. If for example, a for statement is used to scan the table, the step would be the entry width, instead of one. If entries are referred to by pointers, variable length and mixed record types may be kept in the same table. (As in the Compiler)

After the single bracketed integer specifying the size of the table has been read the action TABLESIZE is executed, this action reserving the required amount of data space. This is followed by the description of the fields, enclosed in square brackets. At the end of the declaration the actions CLEARTYPE and TYPEINT set IDTYPE to integer in case the table is to be preset.

Fieldtype=(Type,Sgint,FIELDDISP)
    (Partfield)
        (unsigned,Partfield,UNSFIELD)

Partfield=(orbosb,int,NOBITS,NOSIG,Intfixfield,orbsb,
    Sgint,FIELDDISP,commab,int,FIELDPOSN)

Intfixfield=(comma,Sgint,NOAFTER)
    (PARTINT)

A table field may occupy a whole word. In this case the field identifier is followed by a data type and a signed integer giving its word position in the entry. The action FIELDDISP calculates the address required for the field using the address TABLED and TABLEI stored by TABADD, and inserts it into the appropriate locations at the base of the stack in readiness for the subsequent call of NEWNAME.
A table field which occupies part of a word may be stored either with a sign bit, or without, in which case 'UNSIGNED' is specified. The type of the field is either integer or fixed, this being resolved by rule Intfixfield. It should be noted, particularly in the case of an integer field, that the number of significant bits is taken to be the number of bits occupied, plus one in the case of an unsigned field. The action FIELDPOSN is executed after the integer specifying the starting bit position has been read, this completing the part word specification. It should be noted that Ferranti bit numbering is used, and the starting bit position specifies the most significant bit of the field. At the request of Ferranti, either square or round brackets may be used in the specification of the field type.

**Switch Declarations**

[D= . . . . . . .
 (switch,ADDRSW,TYPESWITCH,Newid,becomes,Swlist)
  . . . . . . .]

Swlist=(id,SPECLAB,Q41)
Q41=()
 (comma,Swlist)

A switch declaration sets up, in the segment's Special Data area, a set of sequential locations containing the addresses of the labels, one per entry. Where a label is external to a segment, the address will be that of the indirect jump to the label, set up at the end of the segment. The action ADDRSW sets the direct address of the record being formed, to the current special data transfer address minus one. Its type is set by TYPESWITCH, and is effectively Label Array.

After each label identifier in the switch list, the action SPECLAB is executed. This looks up the label in the (local) label list, and outputs either zero or its chain address to special data. (The label cannot yet have been set at this block level, except in the special case of label parameters.)

**Overlay Declarations**

Overlaydec=(overlay,Base,with,OVERON,Datadec,OVEROFF)

Base=(id,LOOKUPD,Q43)
 (NONAME,osb,int,INCTOS,csb)
Q43=()
 (osb,Sgint,Inctos,csb)
Datadec=(Type,Q2i)
 (Tabledec)

Any data location may be overlaid with a data declaration. The base defines the starting point of the overlaying data. As an extension, this may be an (absolute) anonymous reference, thus giving a limited local equivalent of an 'ABSOLUTE' communicator.

An arithmetic Operand record is created by either NONAME or LOOKUPD. In the second case, if the direct address is zero, the indirect address replaces it. The action INCTOS increments the direct address by the integer constant.
The data declaration is preceded by the execution of the action OVERON, which deletes the Arithmetic Operand record stores its direct address in OVERBASE, and sets the OVERLAY flag. This flag is cleared later by OVEROFF.

The following points should be noted:

1. Where a multi-dimensional array is overlayed, the actual area overlayed is its first level Iliffe vector. It is usually better to overlay a single dimensional array with multi-dimensional array(s).

2. There is no check to ensure that the space required by the overlaying data declaration does not exceed that of the overlayed data area.

3. Where a procedure parameter is overlayed, the base is taken as the location occupied by the parameter within the data space local to the procedure, even if the parameter is a 'LOCATION' or 'ARRAY' pointer.

4. Where the base is a field of a table passed as a parameter, the effective address is the displacement of the field relative to the start of the table.

Special Declarations

Specialdec=(special,array,TYPEINT,TYPEARRAY,ADDRSPEC,Newid,becomes,Speclist)

Speclist=(Specitem,Q45)

Q46=()  
  (comma,Speclist)

The Special Array facility is introduced as an aid to the production of software and not as a facility for everyday use. Its introduction compensates for the restriction that constants may not be embedded within code. A special array is classed as an Integer Array, with a zero lower bound, and must be completely preset. The normal restrictions on presetting do not apply. The preset values may contain address and numeric information packed into one word (10/14 bit packing). It is usually bad practice to alter the contents of a special array during the execution of a program.

Specitem=(Zint,Q45)  
  (string,SPECSTRING)  
  (CLEARTYPE,Type,orb,Pnumber,SPECNUM,Morespec)

Morespec=(orb)  
  (comma,Pnumber,SPECNUM,Morespec)

Q45=(SEXTEN,Specadd)  
  (CLEARTYPE,TYPEINT,SPECNUM)

Specadd=(colon,id,SPECLAB)  
  (self,Sgint,SPECREL)  
  (div,Q44)

Q44=(Sgint,NONAME,INCTOS,SPECCON)  
  (Base,SPECCON)
A special item may be either a packed word, a string, or a numeric constant. If the constant is not to be stored as an integer, one or more constants, enclosed in round brackets may be preceded by a data type. These constants will be rescaled by SPECNUM before being output. In the case of a string, the actual standard string, occupying one or more words, is stored in the special array at the point of its occurrence by SPECSTRING. This differs from presetting an integer with a string, where the preset value is its address.

Where a packed word occurs, the action SETTEN is executed to store the ten bit constant temporarily in TOPTEN. This constant, which may be void, is then followed by either a colon, the symbol 'SELF' (*), or an oblique stroke. The colon must be followed by an identifier, which is taken as that of a label by the action SPECLAB, and the label address subsequently stored.

In the case of a relative address ('SELF' or *), the symbol is followed by a signed integer, the action SPECREL treating the address as relative to the current special data transfer address. (NB not the current program transfer address.) In the third case, the fourteen bit field may be either a second integer constant, or a data address, as processed by the rule Base. The Arithmetic operand record set up in this case is used by the action SPECCON.

**Level Directives**

```plaintext
[Compileitem= . . .
 (Leveldec)
 . . . . .
]
```

```plaintext
Dm = . . . .
 (Leveldec)]
```

```
Leveldec=(level,int,SETLEVEL)
```

A level directive sets the level of diagnostic information required at compile time, and required to be output to the loader. After the unsigned integer giving the level required, the action SETLEVEL is executed.

**Trace Directives**

```plaintext
[Compileitem= . . .
 (Yesno,Tracetype,trace)
 . . . . .
]
```

```plaintext
Dm = . . . .
 (SETYES,TYPEPROC,Q48)
 (no,SETNO,Tracetype,trace)
 . . . . . . . . . . . . . .
```

```
Tracetype=(assignment,ASSTRACE)
 (loop,FORTRACE)
 (label,LABTRACE)
 (procedure,PROCTRACE)
```

```
Q48=(assignment,ASSTRACE,trace)
 (loop,FORTRACE,trace)
 (label,LABTRACE,trace)
 (procedure,Q49)
```

```
Q49=(Newid,BEGINPROC,Procrest)
 (PROCTRACE,trace)
```

85
Trace directives govern the amount of instructions specifically generated to enable the action of a program to be traced. Trace is split into four categories: assignment, loop, label and procedure. Each type may be turned on or off as required, the trace state following the block level. At the end of a block the state is reset to that of the outer block, but at the start of a block it is initially that of the outer block. Normally it is off.

The actions SETYES and SETNO set the value of ACCUMULATOR to 1 and 0, respectively. (These actions also being used for overflow tests.) The actions ASSTRACE, FORTRACE, LABELTRACE and PROCTRACE setting the appropriate bit of TRACE from the value of ACCUMULATOR.

Because of the clash (starting with same symbol) between 'PROCEDURE' 'TRACE' directives and the declaration of untyped procedures, the rules have been expanded by SID, and the action TYPEPROC is always executed in conjunction with the action SETYES. A manual transformation of the above rules could reduce their size.

Trace instructions are generated by the actions STOREAWAY, DOCS, SETLABEL, and PROCENTRY & EXITCHECK.

Procedure Type Declarations

\[(D=(Type, Q47))
\quad(Valproc, Newid, BEGINPROC, colon, Newid, NEXTPSET, Procrest)
\quad(SETYESTYPEPROC, Q48)
\]

Q47 =

\[(procedure, Q49)
\]

Q49 =

(Valproc, value, TYPEVPROC, procedure)

The above rules "untangle" the start of a procedure declaration. Before the procedure identifier is read, the type information in IDTYPE is complete, so that NEWNAME can move the Identifier Specification Record onto the stack. At this point however, the PARAMSPEC pointer has not been set, and the DIRADD field contains the address of the location which will subsequently be used to hold the link. DIRADD[0] is set correctly however for allocating a location for the first parameter.

After the procedure identifier has been read, the action BEGINPROC is executed, this completing the procedure's Identifier Specification record, and initiating its Parameter Specification record. A location in special data is allocated as a pointer to the procedure, and the address of this inserted into the DIRADD field of its record. In preparation for the following parameters, if any, the PAC field of IDTYPE is set to 7.

If the procedure is a 'VALUE' procedure, its name is followed by a colon, and the identifier of the "context" parameter, which is processed by Newid. This is followed by the execution of NEXTPSET to clear the types and special markers from IDTYPE, leaving the PAC field set to 6.
Procedure Parameter Declarations

Procrest=(Parameterpart,semi,Procbody,ENDPROC)

Parameterpart=(orb,Parameters,orb)

Parameters=(Parameterset,Q65)

Q65=()

(Parameters)

Parameterset=(Type,Q63)

(location,TYPELOC,Q62)

(label,TYPELABEL,Idlist)

(switch,TYPEISWITCH,Idlist)

(table,TYPEINT,TYPEIARRAY,Newid,csb,int,csb,PARAMTAB,Tabledetail,PARAMTAB)

(value,Q64)

(SETYPES,TYPEPROC,procedure,Procparamlist)

Q64=(Type,Idlist)

(TYPESPEC,Pairlist)

(TYPEVPROC,procedure,Procparamlist)

Pairlist=(Newid,colon,Newid,Q61)

Q61=()

(comma,Pairlist)

Q62=(Type,Idlist)

(TYPESPEC,Pairlist)

Q63=(array,TYPEIARRAY,Idlist)

(TYPESPEC,Pairlist)

Procparamlist=(Newid,Paramlev2,Q58)

Q58=()

(comma,Procparamlist)

The type of a set of parameters is built up incrementally in IDTYPE, the PAC field having been previously set. In particular, where the location of an object is to be passed as a parameter the marker LCM will be set. Each parameter identifier will cause the action NEWNAME to be executed. As well as adding the Identifier Specification record to the DECLIST, in the case of parameters NEWNAME will call NEXTPARAM to add the specification of the parameter to the procedure's Parameter Specification record.

When the procedure has been compiled, the Identifier Specification records of the individual parameters will be deleted, as they are then out of scope. The Parameter Specification record of the procedure will however remain until the procedure itself goes out of scope. This record, and not the records of the individual parameters, is required for the generation of calls of the procedure.
In the cases of typed 'VALUE' & 'LOCATION', 'ARRAY', 'SWITCH' and 'LABEL' parameters the specification is followed by a list of parameter identifiers, separated by commas. In the case of untyped 'VALUE' & 'LOCATION' parameters, the identifiers occur in pairs, the second of which is used to pass the type/scale of the first. 'TABLE' parameters require the size of the table to be specified, even though this is ignored. This is followed by the specification of the individual fields.

In the case of typed or untyped 'PROCEDURE' parameters, the parameter identifier is followed by a specification of its parameter requirements, if any. This enables a Parameter Specification record to be formed for each 'PROCEDURE' parameter, this being required when a call of the parametric procedure is to be compiled.

After the semicolon separating sets of parameters has been read, the action NEXTSET is executed, clearing the type information from IDTPE in readiness for the next set.

Procedure Parameter Specifications

Paramlev2=(BEGINPSPEC,Q57)
Q57=(orb,Paramtypelist,crb,ENDPSPEC)
(ENPSPEC)
Paramtypelist=(Paramtype,NEXTPARAN,Q56)
Q56=()
(comma,NEXTPSET,Paramtypelist)
Paramtype=(Type,Q54)
(location,TYPELOC,Q53)
(label,TYPELAB)
(switch,TYPEISWITCH)
(table,TYPEINT,TYPEIARRAY)
(value,Q55)
(SETYES,TYPEPROC,procedure)

Q55=(Type)
(TYPESPEC,NEXTPARAN)
(TYPEPROC,procedure)
Q54=(array,TYPEIARRAT)
(TYPEPROC,procedure)

Where a procedure is specified without being declared, this occurring in the case of 'PROCEDURE' parameters in a procedure declaration, and also where procedures are specified in communicators, its parameter requirements must also be specified if the procedure requires parameters. This enables a Parameter Specification record to be prepared, and a reference to it inserted in the PARAMSPEC field of the Identifier Specification record of the procedure. This record is required even if the procedure (apparently) has no parameters. (NB A "parameterless" 'VALUE' procedure has the "context" parameter.)
In the rule Paramlev2 the action BEGINSPEC is executed to initiate the Parameter Specification record, tentatively reserving eight words for this. The rule Q57 detecting the case where the procedure requires no parameters. In this case the action ENDFSPEC being immediately executed. These two actions setting up a null record.

Where parameters are required, an opening round bracket will be read. This is followed by a list of parameter types, separated by commas, and terminated by a closing round bracket; after which the action ENDFSPEC is executed.

The type of each parameter is built up incrementally in IDTPE, and added to the record by the action NEXTPARAM. After each comma the type information is cleared from IDTPE by NEXTPSET. In the case of untyped 'VALUE' & 'LOCATION' parameters the action NEXTPARAM is effectively executed twice, to insert the specification of the context parameter. Where procedures are specified at this level, a specification of their parameters is not required.

The Procedure Body

Procbody=(code,begin,KILLPARAMS,Od1,ENTERPROC,Cs1,end)  
           (SETPARAMS,Q51)

Q51=(begin,Od1,PROCENTRY,S1,end,EXITCHECK)  
    (ENDDECS,PROCENTRY,SS,EXITCHECK)

Od1=(D1,ENDDECS)  
    (ENDDECS)

SS=()  
    (Ifstat)  
    (Forstat)  
    (Answerstat)  
    (Gotostat)  
    (Id,LOOKUP,STATUSCHECK,Q33)  
    (STATUSCHECK,Q26,A1)

Where the body of a procedure is a code statement, the normal parameter mechanism is inhibited by the action KILLPARAMS. This cancels the allocation of data space for the parameters, and deletes their Identifier Records.

Otherwise the action SETPARAMS is executed, which confirms the allocation of the data space required to hold the parameters. If the procedure body is a block, a new block level is not set up, as one has already been set up for the procedure. This prevents the redeclaration of the identifiers used for parameters, at the outer level, thus rendering them inaccessible.

When the first statement of the procedure body is reached, the action PROCENTRY is executed. This outputs instructions to store the link and the parameters. At the end of the procedure the action EXITCHECK is executed. This ensures that a result is returned if a result is required.
**Answer Statements**

Answerstat=(STATUSCHECK,answer,SETANS,Expr,ANSCHECK)

This statement is only allowed within the body of a typed procedure. After the symbol 'ANSWER' has been read, the action SETANS is executed. This sets EXPSCALE to the type required for the result, and sets SCALEFIRM. If however the procedure is a 'VALUE' procedure, these are both cleared.

At the end of the expression delivering the result, the action ANSCHECK is executed. This ensures that the result has been evaluated in accumulator 7, to the scale given in EXPSCALE. The variable STATUS is set to 2 to denote that an exit instruction is required.

**Procedure Calls**

[Tpri= . . . . . .
  (id,LOOKUP,Q28)
  . . . . . .
Q28= . . . . .
  (RHSPTEST,Rhspsel)

Rhspsel= . . . .
  (Procall)

S= . . . . .
  (id,Q38)
  . . . .
Q38= . . . .
  (LOOKUP,STATUSCHECK,Q33)

Q33= . . . .
  (LHSPROC,Procall,BEHEAD)]

Procall=(SETUPPROC,Selparms,CALLPROC,FINISHPROC)

Selparms=()
  (orb,Paramloop)

Paramloop=(NEXTTYPE,Selptype,ANYMORE,Selmore)

Selptype=(Expr)
  (Locvar)
  (id,LOOKUP)
  (id,Q32)

Selmore=(Multitest)
  (comma,Paramloop)

Multitest=(orb)
  (comma,CALLPROC,MULTICALL,Paramloop)
When the requirement for a procedure call is detected, the rule **Procall** is entered and the action **SETUPPROC** executed. This prepares for the call of the procedure, and raises the expression level in readiness for the evaluation of parameters. If the procedure requires parameters a result of one is returned, otherwise zero. This value is used as the selector for the rule **Selparams**.

After the call of the procedure has been generated the action **FINISHPROC** reduces the expression level, and alters the type of the Arithmetic Operand record set up for the procedure identifier, to the type of the result of the procedure.

Before each parameter is processed, the action **NEXTPTYPE** is executed. This prepares for the evaluation of the parameter to the required scale, in the required accumulator. It returns a result, given by **PARAMCLASS**, which is used as a selector by the rule **Selptype**. This selects the syntax for the parameter, choosing the rule containing the actions appropriate to the type required.

After each parameter has been processed, the action **ANYMORE** is executed. This checks that the parameter is of exactly the required type. It returns a result of zero if no more parameters are required, and a result of one otherwise. This is used as a selector by the rule **Selmore**.

If more parameters are supplied than are required, this case being detected by the rule **Multitest**, it is assumed that a multiple call of the procedure is required. The call is output by the action **CALLPROC**, which in this case is followed by the action **MULTICALL**, which prepares for the second and subsequent calls.

**Common Communicators**

```plaintext
[Compileitem= . . .  
  (Commondec) 
  . . . . . . . . . .] 

Commondec=(common,orb,COMON,Commonlist,COMOFF,crb) 

Commonlist=(STARTDEC,Commonitem,Q23) 

Q23=()  
  (semi,Commonlist) 
```

A common communicator is compiled to produce a loadable segment. This consists of two main parts, a common data part providing data space accessible to all the segments, and a reference part containing references to switches, labels, and procedures accessible to all the segments, each being provided by only one segment.

In this implementation, a means of checking that all these references have been set before the programs are entered is provided by the "common check tape". Within the common segment all internal references are made by means of D and S address tags, but all references by segments to the common segment are made by means of C tags.

A checksum is formed from the characters of the common communicator to ensure that a segment is loaded with the correct common.
Data declarations within a common communicator are handled in exactly the same manner as data declarations in a program segment. Where multi-dimensional arrays are used, the Iliffe vectors are set up in the special data area in the usual way. This also applies to strings and Specialdecs. The identifiers of these items of data are originally addressed relative to the D and S areas of the common area, these addresses being subsequently changed to addresses relative to the C area. It is important to note that this re-addressing assumes that the S area of the C area immediately follows the D area.

Each identifier referring to an item to be provided by a segment, is processed by the rule Comspecid. This addresses the identifier with the current value of the special data transfer address, and after the identifier has been processed by Newname, outputs a preset value of zero to the allocated location by the execution of the action SPECONF.

Where an identifier is used in a label context in a special array in a common communicator, this reference is chained up to the previous specification or reference if one exists. In the case where one does not, an implicit specification of the identifier as a common label is made.
The purpose of the absolute communicator is to communicate to the compiler the specifications and addresses of items which may be regarded as having fixed core store locations. Each identifier is followed by its absolute address, this being processed by the rule Absid. This communicator may be used to enable peripherals to be referred to by an appropriate identifier.

The interpretation of the absolute address in the case of single word variables, and single dimensional arrays and tables is straightforward. In the case of multi-dimensional arrays, the absolute address is taken to be that of the zeroth element of the first level Iliffe vector. In the case of switches the address is taken to be that of the zeroth, not first, entry in the switchlist, and should therefore be set one back if the switchlist is regarded as having a lower bound of one.

In the cases of labels and procedures the absolute location specified is regarded as containing a pointer to the appropriate point in the program.
The purpose of a library communicator is to communicate to the compiler the names, specifications, and reference numbers of procedures and preset variables available in the library. Thus library procedures explicitly referred to do not have to be built-in to the compiler, necessitating its alteration each time the library is extended. Private libraries may be built up as required without difficulty. Treating library procedures in exactly the same manner as any other procedure avoids the necessity for placing restrictions on the use of library procedures as parameters of procedure calls. The special library procedures for tracing and floating point arithmetic are never explicitly referred to by the user, and do not need specification in a communicator.

Library Compilation

```
Library Communicators

Q69=(orb,Liblist,crb)
   (CLEARTYPE,Q66)

Liblist=(CLEARTYPE,Libitem,Q19)
Q19=()
   (semi,Liblist)

Libitem=()
   (Type,Q18)
   (Valproc,Libprocs)
   (SETYES,TYPEPROC,procedure,Libprocs)

Q18=(Libidlist)
   (TYPETPROC,procedure,Libprocs)

Libidlist=(Libid,Q16)
Q16=()
   (comma,Libidlist)

Libid=(id,dir,int,LIBADD,NEWNAME)
Libprocs=(Libid,Paramlev2,Q17)
Q17=()
   (comma,Libprocs)
```

The special library procedures for tracing and floating point arithmetic are never explicitly referred to by the user, and do not need specification in a communicator.
A library procedure referring only to previously compiled procedures and variables, may be compiled when required. The RLB output tape in this case should be spliced on to the front of the library tape. After the procedure or variable has been compiled its specification remains set up for the duration of the run of the compiler. For subsequent runs however, its specification should be added to the library communicator.

Each new item added to the library is required to be allocated a reference number in the range 0 to 16383. This is the number by which it will be referred to by the loader, and is required by the rule Libid to follow the identifier of the library item in communicators and when compiling library items. The user however refers to the library item always by its identifier only.

The only two compiling actions specific to library compilation are SETLIBSEG for procedures, and SETLIBVAR for variables. The action LIBADD is used in both library compilation and library communicators, and is the only action specific to the second case.

External Communicators

[Compileitem=...
  (external,orb,Extlist,crb)]

Extlist=(CLEARTYPE,Extitem,QI5)

QI5=()
  (semi,Extlist)

Extitem=()
  (Type,Q14)
    (Valproc,Extprocs)
    (label,TYPELAB,Extidlist)
    (switch,TYPESWITCH,Extidlist)
    (table,TYPEINT,TYPEARRAY,Extid,osb,int,osb,Tabledetail)
    (SETYES,TYPEPROC,procedure,Extprocs)

Q14=(Q9,Extidlist)
    (TYPEPROC,procedure,Extprocs)

Extidlist=(Extid,Q12)

Q12=()
  (comma,Extidlist)

Extid=(id,div,SETTEN,int,Zint,EXTADD,NEWNAME)

Extprocs=(Extid,Paramslev2,Q13)

Q13=()
  (comma,Extprocs)
An external communicator is similar in function and form to an absolute communicator. The difference being that whilst the address of an item must be specified before compilation in the case of an absolute communicator, it need only be specified before load time in the case of an external communicator.

This is performed by allocating 32 "external bases" (numbered 0-31) and allowing an item to be addressed up to $255$ from the specified base. This is specified in the rule Extid.

The first integer specifies the base, and the second optional integer the displacement. Care should be taken that a reference to an item such as an external array does not exceed the displacement as there will be overflow into the field used for the base number. If this case occurs, it may be necessary to set sufficient adjacent bases at intervals of 512.
THE SYNTAX ANALYSER

The syntax analyser is the main controlling routine of the compiler. By interpreting the processed syntax stored in SYNTAX, together with the auxiliary information stored in BOOLWORD, it not only checks the incoming source program for validity but invokes the required compiling actions at the requisite points.

These two tables have been produced automatically by the use of the SID and SAG syntax processing programs. These both check that there is no possible ambiguity in the definition of the syntax. The original syntax rules have been reduced to "one track form" by the use of SID, this obviates the requirement for backtracking, which can be time consuming. This produces a legible transformed syntax, which is used as the actual working syntax by the compiler writer, in this case, and has been edited as the compiler developed. The working syntax is checked and transformed into the two 'SPECIAL' 'ARRAY'S by the use of an Argus version of SAG (#BSAG) which runs on an ICL 1900.

The syntax table produced can be regarded as the machine code of a special purpose machine. In this interpretation, the syntax analyser is the machine, the syntax its "high level" language, SAG its compiler, and the Coral compiler its assembler. The importance of this approach lies in the fact that the syntax rules input to SAG are not merely used as the bare bones of the language definition, but are used as a high level language to express the overall compiling strategy.

A flowchart is given for the syntax analyser. This expresses unambiguously what its operations are, but should be used in conjunction with the following definition of the interpretive code in order to gain a fuller understanding of its operation.
THE INTERPRETIVE CODE

The interpretive instructions fall into two groups, symbol tests (m.s. bit=1), and unconditional operations (m.s. bit=0).

Unconditional Operations

In these operations the most significant two octal digits specify the operation, and the least significant five octal digits the "address". This may either be an address relative to the start of the syntax (s), or a core location (n).

<table>
<thead>
<tr>
<th>Operation Code</th>
<th>Action</th>
</tr>
</thead>
<tbody>
<tr>
<td>00</td>
<td>Goto syntax at s</td>
</tr>
<tr>
<td>04</td>
<td>Select one of the following (s) alternatives depending on the value stored in SSEL</td>
</tr>
<tr>
<td>10</td>
<td>Goto the labelled action n, taking the stacked interpretive link on return</td>
</tr>
<tr>
<td>14</td>
<td>Call the procedure whose address is stored at n, storing its result in SSEL and taking the stacked interpretive link</td>
</tr>
<tr>
<td>20</td>
<td>Call the rule s, stacking the interpretive link for return</td>
</tr>
<tr>
<td>30</td>
<td>Goto the labelled action n, continuing with the following interpretive instruction on return</td>
</tr>
<tr>
<td>34</td>
<td>Call the procedure whose address is stored at n, storing its result in SSEL and continuing with the next interpretive instruction</td>
</tr>
</tbody>
</table>

The actual representation of these operations in the output of SAG is as follows, where dd represents a decimal integer and id an identifier.

<table>
<thead>
<tr>
<th>Operation Code</th>
<th>Representation</th>
</tr>
</thead>
<tbody>
<tr>
<td>00</td>
<td>#000/dd</td>
</tr>
<tr>
<td>04</td>
<td>#100/dd</td>
</tr>
<tr>
<td>10</td>
<td>#200:id</td>
</tr>
<tr>
<td>14</td>
<td>#300:id</td>
</tr>
<tr>
<td>20</td>
<td>#400/dd</td>
</tr>
<tr>
<td>30</td>
<td>#600:id</td>
</tr>
<tr>
<td>34</td>
<td>#700/id</td>
</tr>
</tbody>
</table>

There are three special cases generated:

1  #600:READER to call the reader, at the start of the syntax to initialise T1, and subsequently to read further symbols.

2  #200:EXIT to exit from the end of an alternative of a rule. (Uses stacked link)

3  #200:FAIL to indicate syntax failure.
CONDITIONAL OPERATIONS

In these operations the most significant octal digit specifies the operation, the next three digits a terminal symbol or symbol group (b), and the least significant four octal digits an address relative to the start of the syntax (s). A special case occurs where the s field is zero, a successful jump in this case indicating the discovery of a syntactic error in the source program.

Where b is less than 72 a symbol match is defined to occur if the value of b equals that of the terminal symbol stored in T1. Where b is 72 or greater (in steps of three) it refers to a three word group in BOOLWORD. A symbol match is defined to occur if the T1 th bit of this group is set.

As a side effect of a conditional operation, a terminal symbol may be "accepted". This causes the next symbol to be read, by means of a jump to READER. This is not done directly, but by interpreting a type 30 instruction stored at the start of the syntax. This is done in order to reduce the size of the syntax analyser, although it is slightly slower.

<table>
<thead>
<tr>
<th>Operation Code</th>
<th>Action</th>
</tr>
</thead>
<tbody>
<tr>
<td>4</td>
<td>If symbol match jump to s.</td>
</tr>
<tr>
<td>5</td>
<td>If symbol match read next symbol and jump to s.</td>
</tr>
<tr>
<td>6</td>
<td>If no symbol match jump to s</td>
</tr>
<tr>
<td>7</td>
<td>If symbol match read next symbol, otherwise jump to s.</td>
</tr>
</tbody>
</table>

All operations of this type are represented by eight digit octal numbers. (#cbbbssss)
The Working Syntax  Introduction

The syntax which follows is the actual syntax used by the compiler. It has been transformed from the original syntax by SID. This has introduced the auxiliary rules Q--. The following convention has been used for the identifiers, to aid the reader:

1  Rule Names First letter in upper case, subsequent letters in lower case. The rule name is underlined where semantic selection is carried out.

2  Terminal Symbols Lower case letters

3  Compiling Actions Upper case letters. The action name is underlined where it delivers a result to be used as a semantic selector.

In the notation used for terminal symbols, lower case letters are used. Most of the symbol names are self explanatory, but a list of the more cryptic and composite symbols is given below.

<table>
<thead>
<tr>
<th>Symbol</th>
<th>Description</th>
</tr>
</thead>
<tbody>
<tr>
<td>zero</td>
<td>zero constant (of any form)</td>
</tr>
<tr>
<td>int</td>
<td>integer constant, including zero</td>
</tr>
<tr>
<td>intcon</td>
<td>integer constant, excluding zero</td>
</tr>
<tr>
<td>realcon</td>
<td>real constant, excluding zero and integer constants</td>
</tr>
<tr>
<td>pconst</td>
<td>zero, integer, or real constant</td>
</tr>
<tr>
<td>shift</td>
<td>shift operator (eg 'SLA')</td>
</tr>
<tr>
<td>ro</td>
<td>comparator (eg 'EQ')</td>
</tr>
<tr>
<td>inst</td>
<td>accumulator and function of code (eg 7 ADD)</td>
</tr>
<tr>
<td>colon</td>
<td>;</td>
</tr>
<tr>
<td>semi</td>
<td>;</td>
</tr>
<tr>
<td>crb</td>
<td>, [</td>
</tr>
<tr>
<td>crbs</td>
<td>] , ]</td>
</tr>
<tr>
<td>mult</td>
<td>*</td>
</tr>
<tr>
<td>self</td>
<td>'SELF', *</td>
</tr>
<tr>
<td>div</td>
<td>/</td>
</tr>
<tr>
<td>step</td>
<td>'STEP', ;</td>
</tr>
<tr>
<td>until</td>
<td>'UNTIL', ;</td>
</tr>
<tr>
<td>osb</td>
<td>[</td>
</tr>
<tr>
<td>osb</td>
<td>]</td>
</tr>
<tr>
<td>power</td>
<td>!, **</td>
</tr>
<tr>
<td>becomes</td>
<td>←, :=</td>
</tr>
<tr>
<td>commab</td>
<td>'BIT' or ,</td>
</tr>
<tr>
<td>id</td>
<td>identifier</td>
</tr>
</tbody>
</table>
Syntax 1 Segments

Run=(Compileitem,Q70)

Q70=(finish)
  (semi,Run)

Compileitem()  
  (Program)
    (Commondec)
      (library,Q69)
      (external,orb,Extlist,crb)
      (absolute,orb,Abslist,crb)
    (Leveldec)
    (Yesno,Tracetype,trace)
    (test,SETTEST,Q68)

Q68=(Program)
  (Commondec)
  (library,CLEARTYPE,Q66)

Program=(id,BEGINPROG,begin,Body,ENDPROG,end)

Body=(D1,ENDDECS,S1)

D1=(STARTDEC,D,semi,Q50)

Q50=()
  (D1)
Syntax 2 Declarations

D=(Type,Q47)
   (Tabledec,Presetlist)
   (switch,ADDRSW,TYPESWITCH,Newid,becomes,Swlist)
   (Specialdec)
   (Overlaydec)
   (Val,Newid,BEGINPROC,colon,Newid,NEXTPSET,Procset)
   (SETTB,TYPEPROC,Q48)
   (no,SETNO,TraceType,trace)
   (Leveldec)

Type=(integer,Q59)
   (fixed,orbosb,int,NOBITS,comma,Sgint,NOAFTER,crbosb)
   (floating,TYPEFLOAT)

Q59=(osb,int,NOBITS,osb,PARTINT)
   (TYPEINT)

Sgint=(int)
   (plus,int)
   (minus,int,NEGNUM)

Q47=(Q21,Presetlist)
   (TYPEPROC,procedure,Newid,BEGINPROC,Procset)

Q21=(Idlist,DECSIZE)
   (array,TYPEARRAY,Arraylist)

Idlist=(Newid,Q52)

Q52=()
   (comma,Idlist)

Newid=(id,NEWNAME)

Presetlist=()
   (becomes,SKIPDATA,Presets)

Presets=(Presetnum,Q40)

Q40=()
   (comma,Presets)

Presetnum=(Pnumber,OUTPRESET)
   (orb,Presets,orb)
   (string,PRESETSTRING)

Pnumber=(pconst)
   (plus,pconst)
   (minus,pconst,NEGNUM)
   (ZERONUM)

Arraylist=(ZEROARRAYS,Idlist,osb,Boundpair,Morebounds,osb,ENDARRAY,Q42)

Q42=()
   (comma,Arraylist)

Boundpair=(Sgint,SETLB,colon,Sgint,SETUB)

Morebounds=(ONEDIM)
   (FIRSTDIM,comma,Boundpair,Arraytail)

Arraytail=(LASTDIM)
   (MIDDIM,comma,Boundpair,Arraytail)
Syntax 3 Expressions

\[ \text{Expr} = (\text{Cexpr})(\text{Axpr}) \]
\[ \text{zero, CONSTANT, SCALTERM} \]
\[ \text{string, STRINGEX, SCALTERM} \]
\[ \text{Axpr} = (\text{Term, Q6}) \]
\[ \text{Axpra} \]
\[ \text{Axpra} = (\text{plus, Term, Q6}) \]
\[ \text{minus, Term, UNARYMINUS, Q6} \]
\[ \text{Q6} = () \]
\[ \text{Q6a} \]
\[ \text{Q6a} = (\text{plus, Term, Add, Q6}) \]
\[ \text{minus, Term, SUB, Q6} \]
\[ \text{Term} = (\text{Factor, Q5, SCALTERM}) \]
\[ \text{Q5} = () \]
\[ \text{mult, Factor, MPY, Q5} \]
\[ \text{div, Factor, DIVIDE, Q5} \]
\[ \text{Factor} = (\text{Difference, Q24}) \]
\[ \text{Q24} = () \]
\[ \text{power, Difference, RAISE, Q24} \]
\[ \text{Differences} = (\text{United, Q4}) \]
\[ \text{Q4} = () \]
\[ \text{differ, United, NBQ, Q4} \]
\[ \text{United} = (\text{Collation, Q3}) \]
\[ \text{Q3} = () \]
\[ \text{union, Collation, ORF, Q3} \]
\[ \text{Collations} = (\text{Term, Q2}) \]
\[ \text{Q2} = () \]
\[ \text{mask, Term, MSK, Q2} \]
\[ \text{Term} = (\text{Sec, Q1}) \]
\[ \text{Pri, Q1} \]
\[ \text{Q1} = () \]
\[ \text{shift, SETSHIFT, Shifts, DOSHIFT, Q1} \]
\[ \text{Shifts} = (\text{Sec, PLUSSUB}) \]
\[ \text{Tpri, SCALTERM, PLUSSUB} \]
\[ \text{orb, Subex, orb} \]
\[ \text{Sec} = \text{Bitset, BITSIN, Pri, RHSBITS} \]
\[ \text{Bitset} = \text{bits, CLEARTYPE, osb, int, Oneormore, FIELDPOSN, UNSFIELD, osb} \]
\[ \text{Oneormore} = \text{NORBITS, NOSIG, PARTINT, comma, int} \]
\[ \text{(ONEBIT)} \]
Syntax 4  Primaries

Pri=(Tpri)
  (realcon,CONSTANT)
  (orb,STACKEXPR,SETPREF,Expr,crb,OFFSTEXPR)

Tpri=(intcon,CONSTANT)
  (id,LOOKUP,Q28)
  (CLEARTYPE,Type,TYPEEXPR,crb,Expr,crb,EXPRTYPE)
  (NONAME,Singlesubs)
  (location,orb,Locvar,LOCACT,crb)
  (label,orb,id,Q32,crb)

Q28=(Subscript,WARCHECK)
  (RHSPTEST,Rhapsel)

Rhapsel=(WARCHECK)
  (Procall)

Locvar=(id,LOOKUP,Q29)
  (NONAME,Singlesubs)

Q29=(Subscript)
  (AYMCHECK)

Q32=(LOOKUP,Singlesubs,LABSK)
  (LABL)

Singlesubs=(substart,csb,KILLEXPR)

Subscript=(Substart,Moresubs,csb,KILLEXPR)

Substart=(osb,SETUPSUB,Subex)

Moresubs=()
  (comma,SUBCOMMA,Subex,Moresubs)

Subex=(Subterms)
  (Cexpr,PLUSSUB)
  (Term,PLUSSUB,Subexcont)
  (zero)

Subexcont=()
  (Subterms)

Subterms=(plus,Term,PLUSSUB,Subexcont)
  (minus,Term,MINUSSUB,Subexcont)
Syntax 5 Conditions

Cexpr = (if IFEX Conditionlist then THENEX Expr else ELSEFI, ELSEFI Expr, ELSEFI)

Conditionlist = (Condition Q31)

Q31 = ()
   (and, OUTCJ, Conditionlist)
   (or, ORACT, Conditionlist)

Condition = (Yesno, overflow, OVRTEST)
   (STACKEXPR, ANYPREP, Lcond, Rcond, RELATION)

Yesno = (no, SETNO)
   (setYES, TYPEPROC)

Lcond = (Axpr)
   (zero, CONSTANT)

Rcond = (ro, Q30)
   (SETNE2)

Q30 = (Condterm Condterms)
   (zero, ZEROCOMP)

Condterm = (Term, CONDPLUS)
   (plus, Term, CONDPLUS)
   (minus, Term, CONDMINUS)

Condterms = ()
   (plus, Term, CONDADD, Condterms)
   (minus, Term,CONDSUB, Condterms)
Syntax 6  Assignment

S1=(St,Q39)
Q39=()
  (semi,Z1)
St=(S,ENDST)
S=()
  (id,Q38)
  (STATUSCHECK,Q26,A1)
  (Ifstat)
  (Forstat)
  (Gotostat)
  (Compound)
  (code,STATUSCODE,begi7,q35)
  (Answerstat)
Q38=(SETLAB, colon,S)
  (LOOKUP,STATUSCHECK,q33)
Q33=(Q25,A1)
  (LHSPROC,Proccall,BEHEAD)
Q25=()
  (Subscript)
Q26=(Bitset,Lhsb)
  (NONAME, Singlesubs)
Lhsb=(id,LOOKUP,LHSBITS,Q25)
  (NONAME,LHSBITS,Singlesubs)
A1=(becomes,VARCHECK,SETASS,A2,STORE,BEHEAD)
A2=(A3)
  (Expr)
A3=(Cexpr)
  (string,STRINGEX)
  (zero,STOREZERO)
  (Term,ADDA,A4)
  (plus,Term,ADDA,A4)
  (minus,Term,SUBA,A4)
A4=(SIMPLEASS)
  (SELOPTA,A5)
A5=(Q6a)
  (Axpra)
Ifstat=(if,IFS,Conditionlist,then,ENDST,St,else,ELSE5,S,FI)
Gotostat=(goto,id,Q27)
Q27=(LOOKUP,STATUSCHECK,Singlesubs,LABSK,GOTOSK)
  (GOTOL)
Syntax 7  Statements

Forstat=(for,STARTFOR,Locvar,CHECKCV,becomes,F1)
F1=(A3,ASSCV,F2)
F2=(F3)
  (while,Conditionlist,WHILEL,F3)
  (step,Expr,STEPUNT,until,Expr,STEPUNT,F3)
F3=(do,DOCS,SEE,ENDFOR)
  (comma,MOREFOR,F1)
Compound=(STATUSCHECK,begin,Q37)
Q37=(S1,end)
  (BEGINBLOCK,Body,end,ENDBLOCK)
Q35=(Cs1,end)
  (BEGINBLOCK,D1,ENDDECS,Cs1,end,ENDBLOCK)
Cs1=(Cs,Q34)
Q34=()
  (semi,Cs1)
Cs=()
  (Compound,STATUSCODE)
  (inst,INSTTYPE,Npart,OUTCODE)
  (id,CODELAB,colon,Cs)
Npart=(Nread)
  (Nwrite)
  (Nshift)
  (Njump)
  (Nloc)
Nread=(Nwrite)
  (Sgint,CONSTANT)
  (CLEARTYPE,Type,orb,Pnumber,SCALECON,crb,CONSTANT)
Nwrite=(anyacc,NISACC)
  (id,LOOKUPD,Nw1)
  (NONAME,Nw2)
Nw1=()
  (Nw2)
Nw2=(osb,Nw3,INCTOS,csb)
Nw3=(Sgint)
  (modacc,NISMOD,Zint)
Zint=(Sgint)
  (ZERONUM)
Nshift=(int,CODESHIFT)
  (modacc,Zint,CODESHIFT,NISMOD)
Njump=(id,LABL)
  (self,Zint,RELADD)
Nloc=(Nwrite)
  (string,STRINGEX)
Syntax 8  Table Declarations

Tabledec=(table,TYPEINT,TYPEARRAY,Newid,osb,int,TABLESIZE,csb,
Tabledetail,CLEARTYPE,TYPEINT)

Tabledetail=(osb,TABADD,Fieldlist)

Fieldlist=(CLEARTYPE,id,Fieldtype,TYPEARRAY,NEWNAME,Q60)

Q60=(csb)
  (semi,Q60a)
Q60a=(Q60)
  (Fieldlist)

Fieldtype=(Type,Sgint,FIELDDISP)
  (Partfield)
    (unsigned,Partfield,UNSFIELD)

Partfield=(orboab,int,NOBITS,NOSIG,Intfixfield,crbcosb,
Sgint,FIELDDISP,commab,int,FIELDPOSN)

Intfixfield=(comma,Sgint,NOAPTER)
  (PARTINT)

Swlist=(id,SPECLAB,Q41)

Q41=()
  (comma,Swlist)

Overlaydec=(overlay,Base,with,OVERON,Datadec,OVEROFF)

Base=(id,LOOKUPD,Q43)
  (NONNAME,osb,int,INCTOS,csb)
Q43=()
  (osb,Sgint,INCTOS,csb)

Datadec=(Type,Q21)
  (Tabledec)

Specialdec=(special,array,TYPEINT,TYPEARRAY,ADDRSPEC,Newid,becomess,Specialist)

Specialist=(Specitem,Q46)

Q46=()
  (comma,Specialist)

Specitem=(Zint,Q45)
  (string,SPECSTRING)
    (CLEARTYPE,Type,orb,Pnumber,SPECNUM,Morespec)

Morespec=(orb)
  (comma,Pnumber,SPECNUM,Morespec)
Q45=(SETEN,Specadd)
  (CLEARTYPE,TYPEINT,SPECNUM)

Specadd=(colon,id,SPECLAB)
  (self,Sgint,SPECREL)
    (div,Q44)
Q44=(Sgint,NONAME,INCTOS,SPECCON)
  (Base,SPECCON)
Syntax 9  Procedure Declarations

Leveldec=(level,int,SETELEVEL)

Tracetype=(assignment,ASSTRACE)
  (loop,PORTRACE)
  (label,LABTRACE)
  (procedure,PROCTRACE)

Q48=(assignment,ASSTRACE,trace)
  (loop,PORTRACE,trace)
  (label,LABTRACE,trace)
  (procedure,Q49)

Q49=(Newid,BEGINPROC,Procrest)
  (PROCTRACE,trace)

Valproc=(value,TYPEVPROC,procedure)

Procrest=(Parameterpart,",semi,Procbody,ENDPROC)

Parameterpart=()
  (orb,Parameters,crb)

Parameters=(Parameterset,Q65)

Q65=()
  (semi,NEXTPSET,Parameters)

Parameterset=(Type,Q63)
  (location,TYPELOC,Q62)
  (label,TYPELAB,Idlist)
  (switch,TYPEISWITCH,Idlist)
  (table,TYPEINT,TYPEIARRAY,Newid,osb,int,csb,PARAMTAB,Tabledetail,PARAMTAB)
  (value,Q64)
  (STYPESPEC,Pairlist)
  (TYPEVPROC,procedure,Procparamlist)

Q64=(Type,Idlist)
  (TYPESPEC,Pairlist)
  (TYPEVPROC,procedure,Procparamlist)

Pairlist=(Newid,colon,Newid,Q61)

Q61=()
  (comma,Pairlist)

Q62=(Type,Idlist)
  (TYPESPEC,Pairlist)

Q63=(array,TYPEIARRAY,Idlist)
  (TYPESPEC,procedure,Procparamlist)

Procparamlist=(Newid,Paramlev2,Q58)

Q58=()
  (comma,Procparamlist)

Paramlev2=(BEGINPSPEC,Q57)

Q57=(orb,Paramtypelist,crb,ENDPSPEC)
  (ENDPSPEC)

Paramtypelist=(Paramtype,NEXTPARAM,Q56)

Q56=()
Syntax 10 Procedure Calls

Paramtype=(Type,Q54)
  (location,TYPETLOC,Q53)
  (label,TYPETLAB)
  (switch,TYPETSWITCH)
  (table,TYPETENTRY)
  (value,Q55)
  (SET,TYPEPROC,procedure)

Q55=(Type)
  (TYPESPEC,NEXTPARAM)
  (TYPEPROC,procedure)

Q54=(array,TYPETARRAY)
  (TYPEPROC,procedure)

Q53=(Type)
  (TYPESPEC,NEXTPARAM)

Procbody=(code,begin,KILLPARAMS,0d1,ENTERPROC,S1,end)
  (SETPARAMS,Q51)

Q51=(begin,0d1,PROCENTRY,S1,end,EXITCHECK)
  (ENDDECS,PROCENTRY,Ss,EXITCHECK)

0d1=(D1,ENDDECS)
  (ENDDECS)

Ss=()
  (Ifstat)
  (Forstat)
  (Answerstat)
  (Gotostat)
  (id,LOOKUP,STATUSCHECK,Q33)
  (STATUSCHECK,Q26,A1)

Answerstat=(STATUSCHECK,answer,SETANS,Expr,ANSCHECK)

Proccall=(SETUPPROC,Selparams,CALLPROC,FINISHPROC)

Selparams=()
  (orb,Paramloop)

Paramloop=(NEXTTYPE,Selptype,ANYMORE,Selmore)

Selptype=(Expr)
  (Locvar)
  (id,LOOKUP)
  (id,Q32)

Selmore=(Multitest)
  (comma,Paramloop)

Multitest=(orb)
  (comma,CALLPROC,MULTICALL,Paramloop)
Syntax 11 Common Communicators

Commondec=(common,orb,COMON,Commonlist,COMOFF,crb)

Commonlist=(STARTDEC,Commonitem,Q23)

Q23=()
  (semi,Commonlist)

Commonitem=()
  (Type,Q22)
  (Valproc,Comprocs)
  (Tabledec,Presetlist)
  (Overlaydec)
  (Specialdec)
  (label,TYPELAB,Comspecidlist)
  (switch,TYPESWITCH,Comspecidlist)
  (SETYES,TYPEPROC,procedure,Comprocs)

Q22=(Q21,Presetlist)
  (TYPEPROC,procedure,Comprocs)

Comspecidlist=(Comspec,Comspecidcont)

Comspecidcont=()
  (comma,Comspecidlist)

Comspecid=(ADDRSPEC,Newid,SPECONE)

Comprocs=(Comspecid,Paramlev2,Q20)

Q20=()
  (comma,Comprocs)

Abslist=(CLEARTYPE,Absitem,Q11)

Q11=()
  (semi,Abslist)

Absitem=()
  (Type,Q10)
  (Valproc,Absprocs)
  (label,TYPELAB,Absidlist)
  (switch,TYPESWITCH,Absidlist)
  (table,TYPEINT,TYPEARRAY,Absid,osb,int,csb,Tabledetail)
  (SETYES,TYPEPROC,procedure,Absprocs)

Q10=(Q9,Absidlist)
  (TYPEPROC,procedure,Absprocs)

Q9=()
  (array,TYPEARRAY)

Absidlist=(Absid,Q7)

Q7=()
  (comma,Absidlist)

Absid=(id,div,int,ABSADD,NEWNAME)

Absprocs=(Absid,Paramlev2,Q8)

Q8=()
  (comma,Absprocs)
Syntax 12 Library

Q69=(orb,Liblist,orb)
   (CLEARTYPE,Q66)

Q66=(Type,Q67)
   (Valproc,Libid,SETLIBSEG,colon,Newid,NEXTSET,Procrest,FINISHSEG)
   (SETYES,TYPETPROC,procedure,Libid,SETLIBSEG,Procrest,FINISHSEG)

Q67=(Libid,becomes,Pnumber,SCALECON,SETLIBVAR)
   (TYPETPROC,procedure,Libid,SETLIBSEG,Procrest,FINISHSEG)

Liblist=(CLEARTYPE,Libitem,Q19)

Q19=()
   (semi,Liblist)

Libitem=()
   (Type,Q18)
   (Valproc,Libprocs)
   (SETYES,TYPETPROC,procedure,Libprocs)

Q18=(Libidlist)
   (TYPETPROC,procedure,Libprocs)

Libidlist=(Libid,Q16)

Q16=()
   (comma,Libidlist)

Libid=(id,div,int,LIBADD,NEWNAME)

Libprocs=(Libid,Paramlev2,Q17)

Q17=()
   (comma,Libprocs)

Extlist=(CLEARTYPE,Extitem,Q15)

Q15=()
   (semi,Extlist)

Extitem=()
   (Type,Q14)
   (Valproc,Extprocs)
   (label,TYPESLAB,Extidlist)
   (switch,TYPESWITCH,Extidlist)
   (table,TYPESINT,TYPESARRAY,Extid,osb,int,osb,Tabledetail)
   (SETYES,TYPETPROC,procedure,Extprocs)

Q14=(Q9,Extidlist)
   (TYPETPROC,procedure,Extprocs)

Extidlist=(Extid,Q12)

Q12=()
   (comma,Extidlist)

Extid=(id,div,SETTN,int,Zint,EXTADD,NEWNAME)

Extprocs=(Extid,Paramlev2,Q13)

Q13=()
   (comma,Extprocs)
'INTEGER' 'PROCEDURE' ACCOF

('VALUE' 'INTEGER' REF);

This procedure is used to find out whether the Arithmetic Operand referred to by REF, is held in an accumulator. If it is, then the result is the number of the accumulator, otherwise zero.

'PROCEDURE' ACCPICK

('VALUE' 'INTEGER' REF,ACC);

This procedure copies the Arithmetic Operand referred to by REF into the specified accumulator ACC. If it already exists in that accumulator, then no action is taken; but if the accumulator is occupied by another operand, then this is first stored. This procedure is used exclusively with floating point operations.

'PROCEDURE' ACCUPDATE

('VALUE' 'INTEGER' REF);

If the Arithmetic Operand on the stack, referred to by REF is not a constant, then a check is made to see if either the direct or indirect address is an accumulator.

If one is, then the appropriate entry is made in ACCS to keep it up to date.

'INTEGER' 'PROCEDURE' ADDADD

('VALUE' 'INTEGER' ADD,INC);

This procedure adds an increment INC, to the index field of an address ADD, ensuring that any carry from the index field (14 bit) does not overflow into the other fields.

'PROCEDURE' ADDRARRAY

('VALUE' 'INTEGER' START);

This procedure is called to insert the addresses into the identifier records of one or more arrays declared with the same bounds. It is called from ONEDIM in the case of single dimensional arrays, and from FIRSTDIM in the case of multi-dimensional arrays. The procedure operates by calling DODIM with the local procedure MINE as the procedure parameter. This procedure parameter is applied to each identifier record, in the reverse order in which the identifiers were declared. Thus the parameters supplied to DODIM have to ensure that the highest address is supplied to MINE first, and the lowest last. The number of identifier records to be addressed is given in ARRAYS.
'PROCEDURE' ADDSUB

('VALUE' 'INTEGER' FUN);

This procedure is called from ADD with a value of FUN of 2, and from SUB with a value of FUN of 3. This procedure generates the instructions required to perform addition and subtraction operations. The scale to which this will be carried out is given by EXPSCALE. The left hand operand is copied into an accumulator, unpacked, and rescaled as required by means of a call of PICK with GOODACC providing the accumulator number. If the right hand operand is a constant then it is rescaled if required. Otherwise it is copied into a second accumulator if it requires unpacking or rescaling. The instruction with the function code FUN is then output. The top operand record is removed from the stack and the type of the other operand record changed to that given by EXPSCALE.

'PROCEDURE' ANSLINK;

This procedure is called in the case where an 'ANSWER' statement is not the last statement of a procedure. If the use of the procedure trace facility is not required, an exit instruction is output by means of OUT27, and the fact recorded by setting EXITCH to one. If however trace is required, an unconditional jump to the end of the procedure, where the call of L6 is made, is output. This jump is chained up using EXITCH as the address, which is then updated.

'INTEGER' 'PROCEDURE' ANYMORE;

This procedure is a selector action called directly from the syntax, after each parameter of a procedure call. If the procedure requires further parameters, a result of 1 is returned, otherwise 0. Before returning to the syntax analyser, however, a check is made to ensure that the parameter just processed is of the correct type.

If the parameter required is an untyped 'VALUE' or 'LOCATION', the type of the actual parameter is copied into EXPSCALE. A switch is then made, using PARAMCLASS, to check the type of the parameter. In the case of 'VALUE' parameters, if the type of the actual parameter is not that of the formal parameter, the type of the actual parameter is changed to that required. If the parameter is a constant, then this is simply rescaled. Otherwise the parameter is evaluated and rescaled, the result being left in an accumulator, which will be the one required for passing the parameter, unless it is already in use.

In the case of 'LOCATION' parameters a check is made to ensure that the actual parameter is not a partword or array, and is of exactly the required type and scale.

In the case of 'ARRAY' and 'SWITCH' parameters a check is made that they are of the exact type required.

If the actual parameter is any type of procedure then a check is made that it is not being used in a potentially recursive situation. This is detected by examining the sign bit of its PARAMSPEC, which will be set if the procedure
is not yet complete, i.e. the use of the procedure identifier occurs within the procedure body.

A check is then made to ensure that the procedure supplied is allowable. If the formal parameter is an untyped procedure then the actual parameter may be any type of procedure except a 'VALUE' 'PROCEDURE', otherwise the actual parameter must be exactly of the right type. No check is made that its parameter requirements match, as insufficient information is kept to enable this to be done.

If the formal parameter is a 'LABEL' then no checking is required, as it is done elsewhere.

After the type checking has been completed, the parameter Arithmetic Operand is processed by LOCACT if a location is required to be passed. If the parameter has a non zero indirect address then the parameter is copied into an accumulator. The parameter specification pointer, PSP, is incremented. If the next parameter is the second, type, parameter of a non-standard parameter (pair) then a call is made on MAKEPARAM to set up an Arithmetic Operand record for this second parameter. A flowchart is given for this procedure.

The result of this procedure is the address of first location of the (last dimension of the) array currently being declared. In the case of a non-overlay declaration this is given by the value of DATAMAX, or in the case of an overlay declaration, OVERBASE. At this point the appropriate variable has not yet been incremented by the total data requirement of the array.

This procedure is called at the start of a block and performs the housekeeping associated with the block structure. It is called directly from the syntax for blocks within a segment. It is also called from PROCSTACK to make up the body of a procedure a block, even if it has no declarations at its head, and also from BEGINFROG.

This procedure starts with a call of BEGLABELBLOCK to set up a block level for labels. This is followed by a call of ONSTACK to place on the stack the values of the following variables whose values will be reset at the end of the block:

BLOCKCHAIN, LEVEL, TRACE, LOCALLIMIT, DATASTART, DATAPTR, DATAMAX, and PRESETOK.

The value of LOCALLIMIT is set to the value ofDECLIST at the start of the block. The variable DATAMAX will be used to allocate workspace to the declarations following, and is set to the value of DATASTART of the outer block at this point. (See also ENDBLOCK)

This procedure begins a block as far as the administration of labels is
concerned. It is called from BEGINBLOCK at the start of a normal block, and also from STARTFOR to prevent the possibility of a jump being made to a label within a 'FOR' statement from a point outside. The current values of LABDECLIST and LABSTACKPTR are stacked using a LABCHAIN. LABDECLIST is then set to zero, thus labels set or referred to in the enclosing block are rendered inaccessible.

'PROCEDURE' BEHEAD;

This procedure removes the top Arithmetic Operand from the stack and deletes any reference to it in ACCS. LH and RH are reset. This procedure is called from the syntax, and from other compiling actions.

'PROCEDURE' CALLLIB

('VALUE' 'INTEGER' LNO,SPIEL);

This procedure is used to set up a call of the Library Procedure whose reference number is given by LNO. This procedure is used by the compiler to set up calls on tracing and floating point arithmetic procedures, it is not used to set up calls of Library Procedures explicitly referred to. The procedure forms the required address by adding an L tag to the LNO and calling OUT27.

'INTEGER' 'PROCEDURE' COPYINACC

('VALUE' 'INTEGER' REF);

This procedure is used to find out whether the Arithmetic Operand referred to by REF is in an accumulator, or if there is a suitable copy of the operand in an accumulator. If the Operand is a constant, or is a location, or is indirectly addressed then the result is zero, otherwise a call is made on FINDOUT and its result taken.

'PROCEDURE' DIAG

('VALUE' 'INTEGER' TA, TB);

This procedure is used to output, on the printer, a commentary of important points reached during compilation. The current program address PTA is printed out, followed by the two strings TA and TB separated by " : ". This is used to indicate the positions of labels and the starting point of procedures and loops. If the value of the variable LEVEL indicates that this information is required at load time, then the string TB is output as part of the relocatable binary output, together with the appropriate directive tag (35).

'INTEGER' 'PROCEDURE' DIRINDADD

('VALUE' 'INTEGER' REF);

The result of this procedure is the direct address of the Arithmetic Operand referred to by REF. If however this is null then the result is the indirect address. This procedure is used where it is known that one address has a value, and the other is null. A similar action takes place in LOOKUPFD.
"PROCEDURE" DODIM

('VALUE' 'INTEGER' START,INC;
 'PROCEDURE' ACT('VALUE' 'INTEGER'));

This procedure is used for processing array declarations. When called at
the first (or only) dimension, the number of array identifiers declared with
the same bounds is given by ARRAYS. If there is only one dimension then
DODIM is called by ADDARRAY which is called by OWEDIM. The effect is to
proceed down the list of identifier records inserting the direct address.
The total store requirement for the arrays is calculated by ARRAYS * NUMBER.

If however there is more than one dimension, each dimension except the last,
requires the setting up of Iliffe vectors in Special Data. In this case the
first dimension causes FIRSTDIM to call ADDARRAY and hence DODIM to insert
the addresses of the first level of the Iliffe vectors into the identifier
records. If there are three or more dimensions then further levels of Iliffe
vectors are set up by MIDDIM calling DODIM with OUT1014CS as the parameter ACT.
This outputs all but the last level of vector. In this case the vectors point
at the next level vectors. For the last ( or second ) dimension LASTDIM calls
DODIM, again with OUT1014CS as a parameter, to output the highest level of
vector. This level point at the actual data area occupied by the array data.
Note that even if the arrays are declared as part of an overlay declaration
separate vectors are set up as required.

"PROCEDURE" DOSTRING

('VALUE' 'INTEGER' STRING;
 'PROCEDURE' PROC('VALUE' 'INTEGER'));

This procedure applies the procedure PROC to each word in turn of the string
STRING. The parameter PROC will be either OUT24CS or OUTCONT.

"PROCEDURE" DUEERR;

This procedure is called in the case where a statement is not labelled, and
not directly entered from the previous statement. If in this case DUECHAIN
is null, the statement cannot be ( legally ) entered, and a message is output
to this effect.

"PROCEDURE" DUMMY;

This procedure does nothing. It is used as a dummy parameter in a call of
STATUSTEST by STATUSCODE.

"PROCEDURE" DUMPACC

('VALUE' 'INTEGER' A);

This procedure outputs an instruction to dump the contents of the specified
accumulator A in an anonymous temporary workspace, allocating this as
required. The direct address of the relevant Arithmetic Operand is set to
the address of the workspace and flagged as being temporary. The entry
in ACCS for the accumulator A is cleared. This procedure is called when it
is required to preserve an Arithmetic Operand which exists only in an accumulator.
PROCEDURE DUMPACCS;

This procedure outputs instructions to store in anonymous temporary workspace all Arithmetic Operands which exist only in accumulators.

PROCEDURE ENDBLOCK;

This procedure is called at the end of a block and performs the housekeeping associated with the block structure. It is called directly from the syntax for blocks within a segment. It is also called from ENDPROC at the end of a procedure, and also from ENDPROC.

The value of DATAMAX at the end of the block is temporarily stored in MAX, and DECLIST reset to its value at the start of the block, this value being held in LOCALLIMIT. OFFSTACK is then called to reset the values of BLOCKCHAIN, LEVEL, TRACE, LOCALLIMIT, DATASTART, DATAPTR, DATAMAX, and PRESETOK. This also has the effect of removing from the stack the identifier records set up inside the block, as these identifiers are now out of scope. If the value of MAX, i.e. the highest numbered data space allocated within the block, exceeds the restored value of DATAMAX, i.e. the highest numbered data space allocated so far, then DATAMAX is set to this larger value. Finally a call is made on ENDLABELBLOCK to compact the label stack and remove any labels now out of scope.

PROCEDURE ENDLABELBLOCK;

This procedure ends the compilation of a block as far as the administration of label records is concerned. It is called from ENDBLOCK and ENDPOR. The previous values of LABDECLIST and LABSTACKPTR are unstacked from the main stack, and the list of label records previously referred to by LABDECLIST processed. These may be divided into three categories:

a) Labels set in the inner block.

b) Labels unset in the inner block and referred to in the outer block.

c) Labels unset in the inner block and not referred to in the outer block.

In the case of (a) the record is simply discarded, and in the case of (b) the chains are joined with JOINCHAINS before the record is discarded. In the case of (c) the record is moved up, and joined on to the label list of the outer block. A flowchart is given for this procedure.

PROCEDURE ENDSIG;

This procedure is called at the end of a segment. It outputs an end of segment tag(39) with the checksum for the characters forming the segment. This is then followed by a tail of blank, a stop directive, erases and blank by means of a call of TAIL. The variable TEST is set to zero in case it had been set by a 'TEST' directive, or a call of STOPOP.

PROCEDURE ENTERPROC;

This procedure is called when the entry point of a procedure is reached. A call is made on SETDUE to set the entry address of the procedure. In the case of a Library Procedure this address is set into the L area by the loader.
END LABEL BLOCK

COPY POINTER TO LABEL LIST

UNSTACK VALUES OF LABDECLIST AND LABSTACKFOR

ANY RECORDS LEFT ON OLD LIST

RETURN

Yes

IS LABEL SET

No

Does label match record on new list

No

MOVE RECORD TO NEW LIST

Yes

JOIN LABEL CHAINS

REMOVE RECORD FROM LIST

Yes

RETURN
clearing the unsatisfied reference, and in the case of a procedure within a
segment this address is set into the location allocated within the Special
Data area. STATUS is set to 0 to indicate that control is passed to the
first statement of the procedure.

This procedure exchanges the values of IDTYPE and IDTYPE2, and also PARAMPTR
and PARAMPTR2. This action is required in the case of the parameter
specifications of a 'PROCEDURE' parameter in the heading of a procedure
declaration. As the action of the procedure is symmetrical, separate procedures
are not required to store and restore the values of IDTYPE and PARAMPTR.

This is the main accumulator allocation algorithm for finding free accumulators.
Starting with the accumulator specified by the parameter A, a scan is made,
with a decreasing index of the array ACCS to find a suitable accumulator.
A mapping function is applied to each entry of ACCS to produce a 24 bit "value"
for that accumulator. The accumulator with the least value, or the higher
numbered one of equal values, is chosen. The mapping function is simple
and arranges that accumulators are chosen in the following order:

- Free accumulators
- Accumulators containing a copy of a stored word
- Accumulators containing (the only copy) an Arithmetic Operand.

If an accumulator of category (c) is chosen, its contents are dumped using
DUMPACC. Within this category the accumulator holding the Arithmetic Operand
furthest down the stack will be chosen. Within category (b) data local to
the segment is regarded as having a greater "value" than data external to it;
and within a segment, data having a higher address (i.e. declared nearest )
is regarded as having a higher "value".

The result of the procedure is the number of the chosen accumulator.

This procedure is used to find whether the address given by the parameter ADD
is an accumulator address or if there is a copy of the word already in an
accumulator. If the address is an accumulator address, the result is the
number of the accumulator. Otherwise a search is made through ACCS, starting
with the lowest numbered accumulator. If the entry in ACCS matches ADD then
the result is the accumulator number. If no match can be found the result is
zero. As a safeguard a zero result is returned for a zero address.
'INTEGER' 'PROCEDURE' FINDPREF;

This procedure is used to find a free accumulator. If the preferred accumulator for the expression, whose number is given by PREFACC, is free then this one is taken. Otherwise a call is made on FINDACC.

'PROCEDURE' FINISHPSPEC;

This procedure is called to complete the parameter specification record of a procedure identifier record on the stack. Any unused space in the parameter specification record is handed back to the stack, and the anti-recursion marker in the PARAMSPEC field deleted.

'PROCEDURE' FINISHSEG;

This procedure is called from ENDPREG at the end of a program segment, and from COMOFF at the end of a 'COMMON' segment. It is also called from the syntax at the end of a 'LIBRARY' 'PROCEDURE' segment. If any labels remain unset in program or library segments, (LABDECLIST not used in 'COMMON') then a warning is printed and the labels set to calls of L2 with the label identifier string as a parameter. ENDPREG is called to output the end of segment directive and tail on the tape. This is then followed by the "size block" output on tape, giving the size of the Program, Data, and Special areas; this information also being printed. A tail is punched on the tape, and "END OF SEGMENT" printed.

'PROCEDURE' FIXCON;

If the operand on the top of the stack is an integer constant, then this is converted to a fixed constant with the minimum number of integer bits required to hold the number without loss of significant bits. This number is given by SGB-1.

'PROCEDURE' FIXLABEL

('VALUE' 'INTEGER' LAB);

This procedure is used when the label specified by the pointer LAB is to be set to the current program address. If the label record indicates that it has been previously set then a call is made on STOPOP to print a message. Otherwise calls are made on SETCHAINOPTA to set the program area and (special) data area chains, if any. The two addresses in the label record are set to the current program address and marked (sign bit set) to indicate that the label is now set. A call is made on ZEROACCS to clear the accumulator record ACCS, even though this might also be called by SETCHAINOPTA. Finally a call is made to DIAG to print out the label identifier and the current program address. FIXLABEL is called from SETLABEL to set labels normally, from FINISHSEG to set labels unset in a segment, and from CODELAB to set labels in code.
PROCEDURE FLOTTT
('VALUE' 'INTEGER' REF)

This procedure ensures that the Arithmetic Operand referred to by REF is of type floating. If the operand is a constant then it is rescaled using SCALENUM. Otherwise if a conversion is required, a call is set up to L10 to float the value.

PROCEDURE FLOATOP
('VALUE' 'INTEGER' OP);

This procedure outputs the instructions to load the operands into @7 and @6, and to call the appropriate Library procedure. If the right hand operand is already in @7, then a procedure which expects the operands to be reversed is called. A table is given below showing the relation between the operators, the value of OP, and the Library procedure called.

<table>
<thead>
<tr>
<th>Operator</th>
<th>OP</th>
<th>Normal</th>
<th>Reversed</th>
</tr>
</thead>
<tbody>
<tr>
<td>*</td>
<td>1</td>
<td>L11</td>
<td>L11</td>
</tr>
<tr>
<td>+</td>
<td>2</td>
<td>L12</td>
<td>L12</td>
</tr>
<tr>
<td>-</td>
<td>3</td>
<td>L13</td>
<td>L17</td>
</tr>
<tr>
<td>/</td>
<td>4</td>
<td>L14</td>
<td>L18</td>
</tr>
<tr>
<td>†</td>
<td>5</td>
<td>L15</td>
<td>L19</td>
</tr>
<tr>
<td>† (integer)</td>
<td>6</td>
<td>L16</td>
<td>L20</td>
</tr>
</tbody>
</table>

After the instructions have been output, the top operand record is removed from the stack, the type of the other operand record changed to floating, and the accumulator record updated.

PROCEDURE FORCON;

This procedure is called at the end of each element of a forlist containing more than one element. After masking the bottom two bits, FORSTATE will be zero for a simple expression, and two for a step-until; otherwise the element is a while.

If the element is a step-until a call is made on FORTEST to output the instructions to test the control variable against the limit. A call of the controlled statement, as an anonymous procedure, is then output. If the element is a step-until, a call is then made on FORINC to output instructions to increment the control variable. Unless the element is of the form Expression, a call is then made on UJBACK to output an unconditional jump back to the step-until or while test. FORSTATE is then set to four prior to processing the next element.

PROCEDURE FORINC;

This procedure outputs the code to increment the control variable in a step-until element of a For statement. ASSFUN is set to OCTAL(12) and a call made on STOREAWAY. This causes instructions to be output which copy the step value into an accumulator and add into store.

125
'INTEGER' 'PROCEDURE' FORMMASK

('VALUE' 'INTEGER' REF);

This procedure is used to form the masks (bit patterns) required for the packing and unpacking in part word operations. The parameter REF is a reference to an Arithmetic Operand on the stack, whose PARTWORD field specifies the location of the required part word within the whole word. This information is obtained from the two fields LSS and MSS of PARTWORD. LSS specifies the number of shifts required to align the field with the least significant end of the word, and MSS the number of shifts required to align the field with the most significant end. Thus for whole word fields these are both zero.

The result of the procedure FORMMASK is a bit pattern consisting of all ones in bit positions corresponding to the specified field, and all zeros elsewhere.

'PROCEDURE' FORTEST;

This procedure outputs the instructions for the testing of the control variable against the limit value of a step-until element of a for statement. This test is always performed using accumulator 7, and the test instruction is always a jump-if-negative; this jump taking place when the element is exhausted. This procedure is used in all step-until cases except where the for statement has only one element, which is a step-until with all three expressions being constants.

If the step is a negative constant, an instruction is output to copy the control variable into accumulator 7; otherwise an instruction is output to copy it negatively. Provided that the limit is not a zero constant, a further instruction is output; if the step is a negative constant, this subtracts the limit value, otherwise it adds it. If the step is not a constant, an instruction is output to multiply the result by the step value. The test instruction is then output, its address being recorded in SKIPCHAIN. Finally the Arithmetic Operand for the limit is deleted from the stack. The code generated is summarised below:

If limit is not zero constant

If step not a constant (Limit-Controlvar)*Step
If step a +ve constant +Limit-Controlvar
If step a -ve constant +Controlvar-Limit

If limit a zero constant

If step not a constant -Controlvar*Step
If step a +ve constant -Controlvar
If step a -ve constant +Controlvar
'PROCEDURE' GIVEUP

    ('VALUE' 'INTEGER' S);

This procedure is called in case of irrecoverable errors. The contents of
the input buffer are printed using PRINTBUFF. This is then followed by
the characters "FAILED : " and the string S. The procedure does not return
control to the point of call but jumps to the label STARTUP to re-initialise
the compiler.

'INTEGER' 'PROCEDURE' GOODACC

    ('VALUE' 'INTEGER' REF);

This procedure choses a suitable accumulator for use with the Arithmetic
Operand referred to by REF. If the operand exists in an accumulator, then
this is chosen. Otherwise a call is made to FINDPREF to find a free
accumulator.

'INTEGER' 'PROCEDURE' GOODONE

    ('VALUE' 'INTEGER' REF,LIM);

This procedure choses a suitable accumulator number for the evaluation of a
subscript expression incorporating the Arithmetic Operand referred to by REF,
which is less than or equal to the value LIM. A detailed flowchart is given.

'PROCEDURE' GOODPICK

    ('VALUE' 'INTEGER' REF,NEG);

This procedure loads the Arithmetic Operand referred to by REF, into a
suitable accumulator. The operand is not rescaled but unpacked if a partword.
If NEG is non zero then the operand is negated.

'INTEGER' 'PROCEDURE' GRABSTACK

    ('VALUE' 'INTEGER' N);

This procedure allocates the number of words given by N, of workspace on the
compiler's main stack. The result is the starting point of this space. A
call is made on STACKCHECK to test for stack collision.

'PROCEDURE' GRABVAR;

This procedure obtains five words on the top of the stack to form a new
Arithmetic Operand. This area is initialised by setting the PARTWORD,DIRADD,
and INDADD words to zero.

'PROCEDURE' HALT

    ('VALUE' 'INTEGER' STRING);

This procedure outputs the characters "HALTED - ", the string STRING, and a
newline. It then waits for Handswitch 0 to be moved from 0 to 1.
'PROCEDURE' INACC

('VALUE' 'INTEGER' ACC,REF);

This procedure updates the Arithmetic Operand on the stack whose reference is REF, to show that it is now held in the accumulator whose number is given by ACC. A call of ACCUPDATE is made to update ACCS.

'PROCEDURE' INST

('VALUE' 'INTEGER' REF,ACC,FUN);

This is the general procedure for the output of instructions whose address and modifier parts are derived from an Arithmetic Operand record on the stack. The required accumulator and function number are passed as ACC and FUN, and the parameter REF refers to an Arithmetic Operand. The required modifier is obtained by the use of ISMOD. The address is obtained from the direct address, but in certain cases an intermediate transformation is required. Where the Arithmetic Operand is a constant, this is output by OUTWCONST and the address supplied by it (WO) used. Where the address is a label address, LBM is set, and USELAB is called to obtain the actual or the chain address. Where a location is required, LCM will be set, and four is added to the function code. In this case the function code will always be less than four.

'INTEGER' 'PROCEDURE' INSTTYPE;

This procedure is a selector action called directly from the syntax analyser, after the three letter function in code. The procedure returns a number in the range 0 to 4 as follows:

<table>
<thead>
<tr>
<th>Function</th>
<th>Type</th>
</tr>
</thead>
<tbody>
<tr>
<td>LDX</td>
<td>0</td>
</tr>
<tr>
<td>NIX</td>
<td>0</td>
</tr>
<tr>
<td>ADD</td>
<td>0</td>
</tr>
<tr>
<td>SUB</td>
<td>0</td>
</tr>
<tr>
<td>LOC</td>
<td>4</td>
</tr>
<tr>
<td>LAB*</td>
<td>3</td>
</tr>
<tr>
<td>STO</td>
<td>1</td>
</tr>
<tr>
<td>STN</td>
<td>1</td>
</tr>
<tr>
<td>ADS</td>
<td>1</td>
</tr>
<tr>
<td>SSB</td>
<td>1</td>
</tr>
<tr>
<td>EXC</td>
<td>1</td>
</tr>
<tr>
<td>AND</td>
<td>1</td>
</tr>
<tr>
<td>NEQ</td>
<td>0</td>
</tr>
<tr>
<td>ORP</td>
<td>0</td>
</tr>
<tr>
<td>JZE</td>
<td>3</td>
</tr>
<tr>
<td>JNZ</td>
<td>3</td>
</tr>
<tr>
<td>JGE</td>
<td>3</td>
</tr>
<tr>
<td>JLE</td>
<td>3</td>
</tr>
<tr>
<td>CVR</td>
<td>3</td>
</tr>
<tr>
<td>JRS</td>
<td>3</td>
</tr>
<tr>
<td>OUT</td>
<td>1</td>
</tr>
<tr>
<td>JCS</td>
<td>1</td>
</tr>
</tbody>
</table>
'INTEGER' 'PROCEDURE' INSTTYPE

<table>
<thead>
<tr>
<th>Function</th>
<th>Type</th>
</tr>
</thead>
<tbody>
<tr>
<td>SRA</td>
<td>2</td>
</tr>
<tr>
<td>SLA</td>
<td>2</td>
</tr>
<tr>
<td>SRL</td>
<td>2</td>
</tr>
<tr>
<td>SLC</td>
<td>2</td>
</tr>
<tr>
<td>SLL</td>
<td>2</td>
</tr>
<tr>
<td>SLV</td>
<td>1</td>
</tr>
<tr>
<td>MPT</td>
<td>0</td>
</tr>
<tr>
<td>DIV</td>
<td>0</td>
</tr>
</tbody>
</table>

Note: LAB is initially given the function code 05, to differentiate it from LOC. This procedure changes the code to 04.

'INTEGER' 'PROCEDURE' ISINACC

('VALUE' 'INTEGER' REF);

This procedure is used to find out whether the Arithmetic Operand referred to by REF is in an accumulator or if there is a suitable copy of the operand in an accumulator. If there is the both the Arithmetic Operand and the accumulator record, ACCS, are updated. A call is made on COPYINACC whose result is made the result of ISINACC. If this is non zero, a call is made on INACC to update the record.

'PROCEDURE' ISMOD

('VALUE' 'INTEGER' REF);

The result of this procedure is the number of the modifier holding the indirect address of the Arithmetic Operand referred to by REF. If the indirect address is null, the result is zero. If the indirect address is a modifier number, or if there is a copy of the required location already in a modifier, the result is this modifier number. Otherwise a call is made on FINDACC to obtain a modifier, and an instruction output to load it. Where the modifier contains a copy of a store location then the appropriate entry in ACCS is updated, otherwise it is cleared. ISMOD is called from INST, (the most common case) from SUBTERM in evaluating subscript expressions, and from STOREAWAY in the case of assignment and 'FOR' statements. A flowchart is given for this procedure.

'PROCEDURE' JOINCHAINS

('VALUE' 'INTEGER' CHAIN;
 'LOCATION' 'INTEGER' MASTER);

This procedure joins the two jump chains, the chain CHAIN being joined on to the end of the chain MASTER. If the chain CHAIN is null, the procedure has no effect, and if the chain MASTER is null it is assigned the chain CHAIN. Otherwise a two-part directive is output to the loader, instructing it to join the chains. If however the chain MASTER is marked (-ve) this represents a label which has been set, and the chain CHAIN is set to this address.
'PROCEDURE' KILLEXPR;

This procedure is called at the end of nested expressions in cases where the result is no longer required. It is called directly from the syntax at the end of subscript expressions, from RELATION at the end of expression in conditions, and from DOSHIFT in the special case of expressions following shift operators. The stacked values of EXPRCHAIN, PREFFACC, EXPSCALE, SCALEFIRM, and BITSHIFT are restored by means of a call of OFFSTACK, and the pointers LH AND RH reset by means of a call of SETLHRH.

'INTEGER' 'PROCEDURE' LHACC;

This procedure is used to find whether the Arithmetic Operand next to the top of the stack (ie the LH operand) is held in an accumulator. If it is, then the result is the number of the accumulator, otherwise zero.

'INTEGER' 'PROCEDURE' LHEQQRH;

This procedure is called after the first Term of the right hand side of an assignment statement. If the Term is identical to the variable on the left hand side, and it is neither floating point nor part-word, then the result is 1, otherwise the result is zero. This is used to detect cases such as: a=a+b.

'PROCEDURE' LIBTRACE

(('VALUE' 'INTEGER' LNO, ADD, SPIEL);

This procedure is used to set up calls on trace procedures having one parameter, usually a string. The address of the string (ie run time address) is passed as the parameter ADD, LNO being the procedure reference number, and SPIEL the commentary. An instruction to load the address into accumulator 7 is output, followed by a call of CALLLIB to output the call of the procedure.

'PROCEDURE' LOCACT;

This procedure acts upon the Arithmetic Operand on the top of the stack. It alters the record, so that when used, the location of the operand rather than its value is obtained. LOCACT is called from the syntax in the case of the use of 'LOCATION'( ), and also from ANYMORE in the case of 'LOCATION' parameters.

If the direct address is non-zero then the type is changed to integer, and the marker bit LCM set. Otherwise the indirect address is copied into the direct address, the indirect address set to zero, and the type changed to integer, but without the location marker set. This second case optimises occurrences of anonymous references in location parameter positions.
'PROCEDURE' LOOKUP;

This is the procedure for looking up the records of non-label identifiers in a non-declaration context. This procedure forms an Arithmetic Operand on top of the stack and copies into it the four full word fields: TYPEBITS, PARTWORD/PARAMSPEC, DIRADD and INDADD. The commentary pointer SPIEL is set to the address of the identifier string in the stored record. If the identifier is used, but undeclared, a warning message is output, and the Arithmetic Operand set to type Integer, with a SPIEL of "(UNDECLARED)".

'INTEGER' 'PROCEDURE' LOOKUPLAB;

This procedure is used to scan the label list of the current block to find whether there is a label record with the same identifier string as that held in NAME. If there is not, then a new record is formed and added to the list by means of ONLAB. Thus the result of LOOKUPLAB is always a pointer to a label record.

'INTEGER' 'PROCEDURE' LOOKUPNAME

('VALUE' 'INTEGER' LIMIT,REF);

This procedure is used to search down the main declaration list, starting with the value held in DECLIST, and stopping when the value of the chain equals that of LIMIT. The identifier string of each record is compared with that of the record referred to by REF. If a match is found, then the result is a pointer to that record, otherwise zero.

'PROCEDURE' MAKEPARAM

('VALUE' 'INTEGER' TYPE);

This procedure is used to create the "hidden" type/scale parameters for untyped parameters and 'VALUE' procedures, and also for the result of intermediate calls of a multiple call of a typed procedure.

An Arithmetic Operand is created on the stack by means of a call of GRABVAR. Its TYPEBITS are set to 'INTEGER' and the accumulator number field inserted. This is obtained from the parameter specification record, the relevant entry of which is pointed at by PSP, which is incremented. The commentary is set to "(TYPE)".

If the parameter TYPE is zero, the record being set up is that of the result of an intermediate call of a multiple call. In this case the direct address of the Arithmetic Operand, and the appropriate entry in ACCS are updated by means of a call of TOPACC. Otherwise the typebits are marked to denote that the Arithmetic Operand is to be treated as a constant, and the direct address field set to the type/scale. If the parameter TYPE is two, this occurring for untyped parameters, the type is obtained from EXPSCALE. Otherwise (TYPE=1) the type/scale is obtained from the type of the procedure being called, which will have been changed to that of the context.
'PROCEDURE' MAKESP\text{}EC;

This procedure prepares a parameter specification record for a procedure identifier record. It is called from PROCSTACK in the case of a procedure declaration, and also from BEGINSPEC in the case of a procedure specification.

The address of a seven word area of stack is set into PARAMPTR2, and the address plus the marker MARK set into the PARAMSPEC word of the procedure identifier record on the head of DECLIST. The marker indicates that the procedure has not yet been completely processed, and is used in the detection of recursion. The PAC field of IDTYPE2 is set to 7, and if the procedure is a 'VALUE' 'PROCEDURE' the fields SPB and PVL are also set for the context parameter. The values set up in PARAMPTR2 and IDTYPE2 will be placed in PARAMPTR and IDTYPE by a subsequent call of EXCP\text{}EC. This precaution is required in the case of a 'PROCEDURE' parameter of a procedure declaration.

'PROCEDURE' MASKINST

\text{}( 'VALUE' 'INTEGER' REF,ACC,SENSE);

This procedure is used to output mask instructions (function 15) for the packing and unpacking of data. The parameter REF is a reference to an Arithmetic Operand on the stack which is required to be packed or unpacked using the accumulator specified by ACC. The parameter SENSE is either all zeros, in which case the required field is unaltered, and the rest of the word cleared; or the parameter is all ones (-1), in which case the required field is cleared and the rest of the word unaltered.

'PROCEDURE' MOVE

\text{}( 'VALUE' 'INTEGER' N;
 'LOCATION' 'INTEGER' FROM,TO);

This procedure moves the number of words given by the parameter N, from successive locations starting from the location given by the parameter FROM, to successive locations starting from the location given by the parameter TO. The higher numbered locations being moved first. This is essential in the case where MOVE is called from ONL\text{}AB which is called by ENDL\text{}ABBLOCK. In this case the areas given by TO and FROM may overlap, the data being moved upwards.

'PROCEDURE' NEWLINE;

This procedure outputs a newline (CrLf) using OUT\text{}AR

'INTEGER' 'PROCEDURE' NEXTCHAR

\text{}( 'LOCATION' 'INTEGER' Z);

This procedure unpacks one six-bit character, packed four characters to a word. The parameter Z specifies the location of a pointer to the packed characters. This pointer is incremented by one character position each time a character is unpacked. The two most significant bits of the pointer being used as an index to the character within the word. The bit pattern 00 specifies the character in the most significant six bits of the word.
This procedure is used to build up the parameter specification record which is pointed at by the PARAMSPEC in a procedure identifier record. This record is a summary of the parameter requirements of a procedure, and is required to be available when a call of the procedure is to be compiled. This should be distinguished from the identifier records of the parameters themselves, which are only required whilst compiling the procedure; and are afterwards deleted.

NEXTPARAM is called directly from the syntax when dealing with procedure specifications, from NEWNAME when dealing with procedure declarations, and from BEGINPSPEC to set the (hidden) context parameter.

If, on entry, IDTYPE is marked negative (seventh parameter), then a call is made on XPARAMS with a zero parameter. This causes a diagnostic message to be printed out using the current (procedure parameter) identifier. If IDTYPE is zero, there have been more than seven parameters, and no action is taken. Otherwise the current value of IDTYPE is added to the parameter specification record, the pointer PARAMPTR incremented, and the following word in the record set to MARK (end of record marker). The parameter accumulator field PAC of IDTYPE is decremented in the following sequence 7,6,5,3,2,1,0. If the value 0 is reached, IDTYPE is set to MARK to indicate that no further parameters are allowed. If the parameter is an untyped parameter (pair) then the field SPB of IDTYPE will be non zero. If it has the value 1, it is set to 2 and the bit LCM cleared. If it has the value 2, it is set to 1 and the bit LCM set to the value of the bit LM2. This bit will only be set in the case of 'LOCATION' parameters.

This procedure is a selector action, called directly from the syntax. It returns an integer in the range 0 to 3 depending on the type requirements of the next parameter. This is obtained by means of a call of PARAMCLASS, but after the values of SCALEFIRM and EXPSCALE have been set up.

If the parameter is untyped, these are cleared, otherwise EXPSCALE is set to the type of the parameter, and SCALEFIRM set to 1.

This procedure forms an Arithmetic Operand on the top of the stack primarily for use with Anonymous References.

GRABVAR is used to set up the Arithmetic Operand, the commentary field SPIEL set to "(ANON)", and the TYPEBITS field set to Integer(Array). This procedure is called from the syntax and from other compiling actions to form an anonymous Arithmetic Operand.
'PROCEDURE' OCTLOOP

('VALUE' 'INTEGER' WORD,DIGITS);

This procedure is the basic octal print routine, it outputs the specified number of octal digits from the least significant end of the value supplied.

The parameter WORD is first shifted cyclically left so that the first octal digit to be printed is moved to occupy the most significant three bits. This is then repeatedly shifted three places left cyclically, and octal digits extracted from the least significant end. These are output using OUTCHAR. The number of digits to be output is specified by the parameter DIGITS.

PROCEDURE OFFSTACK

('VALUE' 'INTEGER' N;
'LOCATION' 'INTEGER' START);

This procedure moves the number of words of data specified by N from the main stack to the area of core starting from START. Before the data is moved the variable whose location is passed as START points at the block of data on the stack. This procedure is essentially the inverse of ONSTACK and is used to "unwind" a level at the end of a syntactically recursive situation. After the data has been moved, the value of the variable STACKPOINTER is updated, as the area on the stack occupied by this data (and any data held above it) has now been relinquished.

PROCEDURE ONLAB

('VALUE' 'INTEGER' FROM);

This procedure moves a label record from the part of the stack whose location is given by the value of FROM, on to the end of the label stack, chaining it up to the chain LABDECLIST. This is used to move the record of a label on to the label stack, on its first occurrence within a block, and also during label stack compaction at the end of a block. It should be noted that the label stack is inverted, i.e., the "top" record on the stack occupies the lowest location.

The size of the record is first calculated, and then the new value to be assigned to the label stack pointer. The record is then moved. If LABDECLIST is zero, it is set to point to the record, otherwise the record below is set to point at the top record. Finally the label stack pointer LABSTACKPTR is set to its new value.
'PROCEDURE' ONSTACK

('VALUE' 'INTEGER' N;
 'LOCATION' 'INTEGER' START);

This procedure moves the number of words of data specified by N, from the area of core starting from START, on to the compiler's main stack. The variable location is passed as START is updated to point to the word on stack containing its previous contents. This procedure is used to stack the compiler's workspace in syntactically recursive situations, eg nested 'FOR' statements. It is also used to move declarations of new identifiers on to the stack and chain them up.

'PROCEDURE' OPERATE

('VALUE' 'INTEGER' FUN);

This procedure is used to apply logical and shift operators between the two Arithmetic Operands on top of the stack. No rescaling takes place prior to the specified operation, and the scale of the result is set as 'INTEGER'.

After ensuring that the left hand operand is in an accumulator, and that the right hand operand is unpacked if required, the instruction with the function code FUN is output using INST. The top operand is removed from the stack, and the type of the remaining operand changed to integer.

'PROCEDURE' OUT1014

('LOCATION' 'INTEGER' TA;
 'VALUE' 'INTEGER' TEN,ADD);

This procedure forms a 24 bit word from the ten bit group TEN and the least significant fourteen bits of the address ADD. This is then output using OUT24 with the transfer address given by TA and the address tag of ADD. This procedure is used to output instructions (Program), and constants containing address information (Special).

'PROCEDURE' OUT1014CS

('VALUE' 'INTEGER' ADD);

This procedure is used to output constants containing address information destined for the special constants (Special) area. The ten bit group to be packed in the most significant end is obtained from the variable TOPTEN.

'PROCEDURE' OUT24

('LOCATION' 'INTEGER' TA;
 'VALUE' 'INTEGER' WORD, TAG);

This procedure is the basic output procedure for 24 bit words which are to be relocated and loaded to core by the loader. The parameter TA specifies which transfer address is to be used to load the word (Program,Data or Special), and the address tag TAG specifies how the word is to be modified. The relocatable binary tag is formed from the address tag of the transfer address and the address tag from the parameter TAG. The appropriate transfer address is incremented by one.
This procedure outputs whole word constants destined for the special constants (Special) area. These constants do not require relocation.

PROCEDURE OUT27

(VALUE INTEGER ADD, SPIEL);

This procedure is used to output instructions with a function part of OCTAL(27), an accumulator part of 4 and the address and commentary specified by ADD and SPIEL. This procedure is designed to be used to set up procedure calls, the link being set in accumulator 4. This procedure is also used for procedure exit, jumps to label parameters, and jumps to 'EXTERNAL', 'ABSOLUTE' and 'COMMON' labels. In these cases there is no need to set a link, but it is simpler to do so.

PROCEDURE OUT5

(VALUE INTEGER TAG, W IND);

This is the basic output procedure for relocatable binary. It outputs a five character group. The first six bit character is given as the parameter TAG, and the following four characters are unpacked from the parameter W IND. Before the first character is output OUTDEV is changed to OCTAL(107) to select the punch, and after the last character is output OUTDEV is reset to OCTAL(106). A checksum TSU is formed of all characters output. If the variable TEST is non-zero then the procedure does nothing.

PROCEDURE OUTCHAR

(VALUE INTEGER CHAR);

This procedure converts the character whose internal representation is given by the parameter CHAR to the appropriate IS07 character. This is then output to the peripheral whose address is given by the variable OUTDEV. The internal representation for newline is 64, this being output as Carriage Return followed by Line Feed.

PROCEDURE OUTCJ;

This procedure is used to output all conditional jump instructions associated with conditions, except those of the form 'THEN' 'GOTO' label, these being output by GOTOL. OUTCJ is called directly by the syntax to output a jump (if false) following the symbol 'AND'. It is called from ORACT, after the jump has been reversed, following the symbol 'OR' (jump if true). It is also called, following the symbol 'THEN' from STATUSTEST and from THEMEX.

This procedure uses the previously calculated values of ACCUMULATOR and FUNCTION, the address part being obtained from SKIPCHAIN, which is updated.
'PROCEDURE' OUTCONT

('VALUE' 'INTEGER' WORD);

This procedure outputs the value passed as the parameter WORD as a five character relocatable binary group. The tag in this case is zero, signifying to the loader that this group forms part of a multi-group element. OUTCONT is often used as a parameter for DOSTRING.

'PROCEDURE' OUTI

('VALUE' 'INTEGER' XFM,N,SPIEL);

This is the basic procedure for the output of instructions (Program). The accumulator, function, and modifier field are passed as a packed ten bit field XFM. The address is passed as the parameter N, and a string for commentary as SPIEL.

Where the address is the address of an accumulator, it is converted to absolute form (X+4096) and a string formed of the form "@X" which replaces the string passed as SPIEL. Otherwise if the address is flagged as being the address of temporary workspace the string passed as SPIEL is replaced by "(TEMP)".

The instruction is output using OUT1014, and if the value of the variable LEVEL indicates that a detailed commentary is required, then this is output using DOSTRING to output SPIEL.

'PROCEDURE' OUTPRESSTD

('VALUE' 'INTEGER' N,T);

This procedure is used to output preset data in data declarations. It is called from OUTPRESSET to output numeric data, and also from PRESETSTRING to output the addresses of strings. (eg In the case of: 'INTEGER' S."STRING";)

A check is made that presetting is allowed at this point, and also that there have not been more presets than data space allocated by the declaration. The preset value N is output by means of a call of OUT24, and relocated according to the address T. (Relocation is only required in the case of strings, which are stored in Special Data)

'PROCEDURE' OUTUJ;

This procedure is used to output unconditional jumps following 'ELSE', and also over procedures where required. The address part for the jump is obtained from SKIPCHAIN, which is updated.

'INTEGER' 'PROCEDURE' OUTWCONST

('VALUE' 'INTEGER' CONST);

This procedure outputs the constant CONST and returns the run-time address of the constant. As the loader optimises and allocates storage for the constants the address returned is always the same (W0). The procedure outputs the constant as a prefix-continuation to the following instruction by the use of OUTCONT. The use of OUTWCONST as a separate procedure allows other methods of dealing with constants to be used if desired.
PROCEDURE OUTXFMN

('VALUE' 'INTEGER' X,F,M,N,SPIEL);

This procedure is used to output instructions (Program) where the fields of
the instruction have been obtained separately. It packs the accumulator (X),
function (F), and modifier (M) fields as a ten bit group. This together with
N and SPIEL are passed to OUTI for subsequent output.

PROCEDURE PARAMCLASS;

This procedure returns an integer, in the range 0 to 3, depending upon the
class of the parameter currently being pointed at, in a parameter specification
record, by PSP. The classification is as follows:

<table>
<thead>
<tr>
<th>Result</th>
<th>Parameter Type</th>
</tr>
</thead>
<tbody>
<tr>
<td>0</td>
<td>'VALUE' (type)</td>
</tr>
<tr>
<td>1</td>
<td>'LOCATION' (type)</td>
</tr>
<tr>
<td>2</td>
<td>'ARRAY','SWITCH', and 'PROCEDURE'</td>
</tr>
<tr>
<td>3</td>
<td>'LABEL'</td>
</tr>
</tbody>
</table>

PROCEDURE PERM;

This procedure exchanges the top two Arithmetic Operands on the stack
(i.e. the LH and RH operands), and updates the accumulator record, ACCS, by
means of calls on ACCUPDATE.

PROCEDURE PICK

('VALUE' 'INTEGER' REF,ACC,SCALE,NEG);

This is the procedure responsible for the output of instructions for copying
the Arithmetic Operand referred to by REF, into the accumulator ACC. It is
unpacked, if required, rescaled to the type/scale SCALE, and negated if NEG
is non zero. Conversions from 'FIXED' to 'INTEGER' are rounded.

This procedure is not used with 'FLOATING' operands.

Detailed flowcharts describing the operation of the procedure are given.

PROCEDURE POWERTWO;

If the top operand on the stack is a positive integer constant which is a
positive power of two, then the result of this procedure is the power of two,
otherwise the result is zero.

This procedure takes advantage of the fact that such a constant, \(2^n\), requires
\(n+2\) significant bits to be represented as a signed integer. This value is
obtained from SGB. The cyclic shift used is equivalent to a division by \(2^n\),
but with the remainder placed in the top of the word. Only if the constant
is positive, and a power of two, will the result be one.
PICK

IS OPERAND A CONSTANT? Yes/No

RESCALE TO REQUIRED SCALE

CALCULATE RESCALING SHIFTS

IS THERE A COPY OF THE OPERAND IN THE REQUIRED ACCUMULATOR? Yes/No

DOES OPERAND REQUIRE UNPACKING? Yes/No

IS NEGATION REQUIRED? Yes/No

DELETE NEGATION MARKER

OUTPUT INSTRUCTION TO COPY OPERAND INTO REQUIRED ACCUMULATOR

OUTPUT INSTRUCTION TO COPY OPERAND NEGATIVELY INTO REQUIRED ACCUMULATOR

Stop
'PROCEDURE' PRINTADD

('VALUE' 'INTEGER' ADD);

This procedure prints the address given as the parameter ADD in the form of a single letter representing the tag type, and five octal digits representing the index. This is followed by the characters ": ".

'PROCEDURE' PRINTBUFF;

This procedure is used, after the detection of an error, to print the contents of the cyclic buffer INBUFF. This contains the last 64 non ignored characters read from the input tape. If less than 64 characters have been read, then only those will be output. The output consists of the characters ": "., the contents of the buffer, and the characters "++". Where the buffer contains a newline, this is followed by the character ":. This is therefore the first character on any line output by this procedure.

'PROCEDURE' PROCSTACK;

This procedure is called to perform the housekeeping required at the start of a procedure declaration. It is called from BEGINPROC for procedures within a segment, and also from SELLIBSEG for Library Procedure segments. STATUS is set to 1, as the procedure is entered by a jump, and a parameter specification record prepared by means of a call of MAKESPEC. A call is made on ONSTACK to move the current values of PROCCHAIN, PROCPTR, PARAMPTR, VALUECH, and DUECHAIN. This is only necessary where a procedure is declared within a procedure, but is carried out in all cases for the sake of simplicity. BEGINBLOCK is then called to set up an outer block for the procedure which encloses the parameters (if any). PROCPTR is set to point to the procedure identifier record, and LOCALLIMIT moved one record beyond this. This, and the call of BEGINBLOCK, prevent the redeclaration of the procedure and parameter identifiers within the outer block of the procedure, thus rendering the parameters inaccessible. EXITCH and PROCPTR are set to zero to inhibit procedure trace unless specifically required.

'PROCEDURE' PUNCHLOOP

('VALUE' 'INTEGER' CHAR,TIMES);

This procedure outputs the binary character CHAR directly to the punch, the number of repetitions being given by the parameter TIMES. This is used only to obtain blank tape and groups of erases. No characters are output if the variable TEST is non zero.

'INTEGER' 'PROCEDURE' READTAPE;

This procedure is the basic input procedure of the compiler. It reads one character at a time from paper tape and converts it to internal representation. Line feed is treated as newline (64), all other control characters and erase being ignored. Lower case letters are mapped onto the corresponding upper case letters. After conversion the character is also deposited into a 64 character cyclic buffer INBUFF, using an integer INCTR as the pointer to the current character position. This pointer always has a value in the range 0 to 63 inclusive. A flowchart is given for this procedure.
'INTEGER' 'PROCEDURE' RESCALE

('VALUE' 'INTEGER' NUMBER, OLDSCALE, NEWSCALE);

This procedure rescales the constant NUMBER whose type/scale is given by OLDSCALE to the type/scale required by NEWSCALE. Type changing from 'INTEGER' to 'FIXED' or 'FLOATING', or from 'FIXED' to 'FLOATING' is allowed, other type changes are not. This means that constants with a decimal point cannot be used in a strictly 'INTEGER' context, e.g. as a subscript. The number type is given by 'BITS' [3,10] and the scale by 'BITS' [6,18] of OLDSCALE and NEWSCALE.

Provided that rescaling is required, the number is first normalised and its scale adjusted as required. If a 'FLOATING' result is required then the number is truncated, rounded, and packed. Otherwise the number is fixed to the required scale, provided this can be done without total loss of significance; and rounded. A flowchart is given for this procedure.

'PROCEDURE' REVCHAIN;

This procedure exchanges the chains DUECH, TH and SKIPCHAIN.

'PROCEDURE' REVCGJ;

This procedure is used to reverse the type of conditional jump instruction calculated but not yet output. Function 20 is changed to 21, and function 22 to 23 and vice versa. In the case of function 24 the accumulator number is changed from 0 to 1, or from 1 to 0. This reversal of the test is required following the symbols 'THEN', 'ELSE', 'THEN' 'GOTO', and 'OR'.

'INTEGER' 'PROCEDURE' RHACC;

This procedure is used to find whether the Arithmetic Operand on the top of the stack (i.e. the RH operand) is held in an accumulator. If it is, then the result is the number of the accumulator, otherwise zero.

'INTEGER' 'PROCEDURE' RHSPTEST;

This procedure is a selector action called directly by the syntax, after the occurrence of an unsubscripted identifier in an expression. If the identifier is the name of a procedure, a result of 1 is returned to the syntax analyser, otherwise a result of 0 is returned. This allows the case of procedures without parameters to be detected, and differentiated from simple variables.

'INTEGER' 'PROCEDURE' SCALECON;

This procedure rescales the constant held in NUMBER, and whose scale is held in NUMBERSCALE, to the scale required by the current context which is held in IDTYPE. This procedure is called by the syntax to rescale Library Preset variables and scaled constants occurring in the operand field of code instructions. It is also called from other compiling actions which output preset values.
This procedure rescales the constant held in the Arithmetic Operand record referred to by REF, to the type/scale SCALE; changing the TYPEBITS of the record to the required scale. This is performed by means of a call of RESCALE.

This procedure is supplied with two scale/type words A and B. It checks these for compatibility, and jumps to the appropriate entry of the switch S according to the mode of the result of an operation with two operands of the specified types. If the four types considered are represented as follows:

- Io = 'INTEGER' of unspecified significance
- Is = 'INTEGER' of specified significance
- Fx = 'FIXED'
- Fl = 'FLOATING'

then the operation of the procedure may be described in the table below:

<table>
<thead>
<tr>
<th>Case</th>
<th>Result</th>
<th>Entry of S</th>
</tr>
</thead>
<tbody>
<tr>
<td>Io,Io</td>
<td>Io</td>
<td>1</td>
</tr>
<tr>
<td>Io,Is &amp; Is,Is</td>
<td>Io,Is</td>
<td>1</td>
</tr>
<tr>
<td>Is,Is</td>
<td>Is</td>
<td>2</td>
</tr>
<tr>
<td>Is,Fx</td>
<td>Fx</td>
<td>3</td>
</tr>
<tr>
<td>Fx,Is</td>
<td>Fx</td>
<td>4</td>
</tr>
<tr>
<td>Fx,Fx</td>
<td>Fx</td>
<td>5</td>
</tr>
<tr>
<td>Fl,Any</td>
<td>Fl</td>
<td>6</td>
</tr>
<tr>
<td>Any,Fl</td>
<td>Fl</td>
<td>6</td>
</tr>
</tbody>
</table>

Where the two scales are Io and Fx, in either order, a warning message is output as the result of the operation may be ill defined because the number of significant bits in the Integer Operand is not known, and hence its tolerance to rescaling cannot be predicted.

This procedure scans the label list of the current block to discover whether there is a label record with the same identifier string as that of the record on the main ( declaration ) stack pointed at by PTR. If there is, then the result is the pointer to the label record, otherwise zero. This procedure is called by LOOKUPLAB to deal with the normal use of labels within a segment. It is also called by ENDPREC to check whether labels unset within a segment are 'COMMON' labels.
'INTEGER' 'PROCEDURE' SELOPTA;

This procedure is a selector action called by the syntax, after the first, but before the second, term on the right hand side of an assignment statement.

If the assignment function, ASSFUN, is marked negative, this marker is deleted, and if the resulting value is OCTAL(10) a call is made on UNARYMINUS. This case occurs where the right hand side commences with an unary minus sign, and the first term is not identical to the left hand variable.

If, now, the value of ASSFUN is OCTAL(10), then assignment optimisation is not required, and a result of zero returned to the syntax analyser. This causes it to chose the rule for the continuation of the expression in the normal manner. If however assignment optimisation is invoked, the top Arithmetic Operand record is deleted from the stack, and the result 1 returned to the syntax analyser. This causes it to chose the rule which effectively restarts the expression at the second Term, to evaluate from this Term onwards, ignoring the first. The result of this expression will be added to, or reverse subtracted from, store.

'INTEGER' 'PROCEDURE' SZTASS;

This procedure is a selector action called by the syntax after the Becomes symbol in an assignment statement.

The value of ASSFUN is set to its initial value of OCTAL(10), the required scale for the right hand side, EXPSCALE, obtained from the scale of the left hand side, and SCALEFIRM set to 1, to indicate that this scale is mandatory. If assignment trace is required then PREFACC is set to 7, as the trace procedure requires it in this accumulator. The result 1 is returned to the syntax analyser. This causes the rule for expression which does not have assignment optimisation actions embedded to be selected.

If no trace is required then PREFACC is set to a value supplied by FINDACC, and the result zero returned to the syntax analyser. This causes the rule for expression which includes assignment optimisation actions to be selected.

'PROCEDURE' SETCHAINTOPTA

('VALUE' 'INTEGER' CHAIN);

This procedure sets the jump chain CHAIN to the current program address. As this represents a point which will be jumped to in the object program, the array ACCS is cleared. The chain is set by outputting the appropriate directive tag (17) to the loader. If the chain is null the procedure has no effect.

'PROCEDURE' SETDUE;

This procedure is used to set the chain DUECHAIN to the current program address. After the chain has been set DUECHAIN is set to zero.
'PROCEDURE' SETLABEL

('VALUE' 'INTEGER' LAB);

This procedure is used when the label specified by the pointer LAB is to be set to the current program address. This is done by means of a call of FIXLABEL. If label tracing is required a call is set up on L3 with the label identifier string as a parameter, by means of a call of TRACESTRING.

SETLABEL is called from SETLAB to set labels explicitly set in a segment, from ENDPTRG to set labels which are supplied by a communicator, and from PROCENTRY to set labels within a procedure which are supplied as 'LABEL' parameters.

'PROCEDURE' SETLHHR;

This procedure sets the pointers LH (left hand) and RH (right hand) to the two top Arithmetic Operands on the stack. This procedure is used frequently, to ensure that these pointers are always up to date whilst compiling expressions.

'PROCEDURE' SESTRT;

This procedure is used in the compilation of For statements. It is called when it is discovered that the For statement is of such complexity that the controlled statement is required to be treated as an anonymous procedure. This condition occurs where the for-list is more complex than the form.

expression comma expression while.

If the first element of the for-list is a simple expression, then the assignment of this value is followed by an unconditional jump whose address is recorded in RTA. If on entry to the procedure, RTA is non zero, a special directive (35) is output to the loader requesting it to change the unconditional jump into a jump-setting-link instruction. A location is allocated in the Special Data area, its address being recorded in RTA, and the location initially set to zero. This will later be set to the entry address for the controlled statement.

'INTEGER' 'PROCEDURE' SETUPPROC;

This procedure is a selector action called directly from the syntax, after the occurrence of an identifier which has been found to be the name of a procedure. If the procedure requires parameters, a result of 1 is returned, otherwise 0.

Before returning to the syntax analyser, the opportunity is taken to perform the housekeeping required to prepare for the generation of a procedure call.

If the call is recursive an error message is output. The values of EXPCHAIN, PREPACC, EXPSCALE, SCALEFIRM, BITSHIFT, PNP, FSP, and FPP are placed on the stack by means of a call of ONSTACK. This protects the workspace of any enclosing expression or procedure call. PNP (procedure name pointer) is set to point at the Arithmetic Operand record of the procedure to be called, FPP (first parameter pointer) set to point to the place on the stack where
'INTEGER' 'PROCEDURE' SETUPPROC/

the record for the first parameter will be created, and PSP (parameter specification pointer) set to point (one back) at the first entry in the parameter specification record of the procedure. The type of the procedure held in the Arithmetic Operand record is then changed to the type of the result of the procedure. If the procedure is a 'VALUE' 'PROCEDURE' this type will depend on the context in which the call is made. This is determined from the variables of the enclosing expression context, which at this point have not yet been altered. If the scale is firm, then this scale is used, otherwise the type 'FLOATING' is used. A call is made on MAKEPARAM to create an Arithmetic Operand record for the context parameter.

'INTEGER' 'PROCEDURE' SPECSTRING;

This procedure outputs the string held in the base of the stack, starting in the location overlayed with the identifier NAME. This string is output destined for the special constants (Special) area. The result of the procedure is the value of the Special Transfer Address (STA) before the string had been output. This procedure is called from the syntax analyser in cases where a string occurs in a 'SPECIAL' 'ARRAY'. In this case the result is irrelevant.

'PROCEDURE' STACKCHECK;

This procedure checks whether the main stack, which extends upwards, has collided with the label stack, which extends downwards. If a collision occurs the compilation is halted.

'PROCEDURE' STACKEXPR;

This procedure is called where it is required to nest expressions. It is called from the syntax in the cases of untyped bracketed expressions and expressions in conditions, from TYPEEXPR in the case of typed bracketed expressions, and from SETUPSUB in the case of subscript expressions. It should be noted that a bracketed expression following a shift operator is treated as a subscript expression, as the result is required to be of integer scale, and be in a modifier if evaluated.

The values of EXPRCHAIN,PREFACC,EXPSCALE,SCALEFIRM and BITSHIFT are stacked by means of a call of ONSTACK. EXPSCALE and SCALEFIRM are then cleared in case the expression to be evaluated is untyped.

'PROCEDURE' STARTSEG

('VALUE' 'INTEGER' TYPE);

This procedure is called at the start of Program (TYPE=1), Library (TYPE=2), and Common (TYPE=3) segments. It punches the segment header on tape, and prints a start of segment message. The transfer addresses PTA,DTA,STA, and the allocation variables DATASTART and DATAMAX set to their respective initial values. ZEROACC is called to clear ACCS.
'PROCEDURE' STATUSCHECK;

This is a compiling action which is called directly from the syntax, and also from other compiling actions.

It is called at the start of every statement except 'CODE' and 'GOTO' statements. STATUSTEST is called, with DUEERR as a parameter, so that if STATUS is 1 or 2 and DUECHAIN is null, then an error is reported. If STATUS is 2, 3, or 4 then the deferred action is carried out.

PROCEDURE' STATUSTEST

This is a compiling action which is called at the start of 'CODE' statements. It ensures that any deferred action associated with values of STATUS of 2, 3, or 4 are carried out. This is performed by means of a call of STATUSTEST with DUMMY as a parameter, as a code statement need not be entered at the beginning. A call is also made on ZEROACCS to clear the accumulator record ACCS.

STATUSCODE is also called from SETLAB as the above actions are required at this point, but for different reasons.

PROCEDURE' STATUSTEST

(PROCEDURE' PROC);

This procedure is called at the start of every statement except 'GOTO' statements. The value of the variable STATUS indicates the type of the preceding statement, as follows:

<table>
<thead>
<tr>
<th>STATUS</th>
<th>Preceded By</th>
</tr>
</thead>
<tbody>
<tr>
<td>4</td>
<td>'THEN' (or 'THEN' 'ELSE')</td>
</tr>
<tr>
<td>3</td>
<td>'ELSE'</td>
</tr>
<tr>
<td>2</td>
<td>'ANSWER' statement</td>
</tr>
<tr>
<td>1</td>
<td>'GOTO' statement (also certain cases of 'DO' statement, and also where the statement is the first statement of block containing the declaration of a procedure)</td>
</tr>
<tr>
<td>0</td>
<td>All other cases. (And also where the statement is labelled)</td>
</tr>
</tbody>
</table>

The use of this method ensures that the number of unconditional jumps generated is reduced as far as possible. In this case it is necessary to defer the actions associated with certain symbols until the following symbol(s) have been processed. A value of 2, 3, or 4 for the variable STATUS indicates that an action has been deferred. A value of 0 indicates that control is passed normally to the statement, but a value of 1 indicates that control is not passed normally to the statement, and if in this case DUECHAIN is null then the statement is inaccessible. In this case a call is made on the procedure parameter PROC. STATUSTEST is called from STATUSCHECK with DUEERR as a parameter, and from STATUSCODE with DUMMY as a parameter. In the latter case it is assumed that a code statement which is not directly accessible, is entered via a relative jump.
'PROCEDURE' STOPOP

('VALUE' 'INTEGER' TA,TB);

This procedure is called in cases where a recoverable error, but of sufficient gravity to prevent the compiled program from running safely, is discovered. This procedure outputs the two strings TA and TB separated by " : ". If the variable TEST is zero then it is set to one and the message " : COMPILATION INHIBITED" output. The effect of setting TEST to a non zero value is to inhibit further output on the punch until the end of the segment.

This is followed by a newline, and PRINTBUFF is called to print the contents of the input buffer.

PROCEDURE' STOREAWAY

('VALUE' 'INTEGER' TFLAG);

This procedure is the procedure that generates the instructions to assign an expression to a variable. It is called from STORE in the case of an assignment statement, and from FORINC and ASSCV in 'FOR' statements. The parameter TFLAG indicates whether assignment trace is required, and is zero in the second two cases. The variable ASSFUN indicates which function is to be used for the assignment. Where this is marked negative, the assignment is of zero, and is treated as a special case. Where ASSFUN is zero, no code is generated; this occurs in the case of dummy assignments, e.g. 'FOR' I=I 'WHILE' . .

Provided that ASSFUN is non zero, the procedure starts by detecting whether the assignment is of the form partword becomes constant, and that trace is not required. If this case is detected then FFLAG is set. Next an address ADD, is calculated. This will be inserted into the entry in ACCS of the accumulator used for the assignment. If the left hand variable is not partword and not indirectly addressed then this address is used, otherwise the address of the right hand side is used if possible. A suitable accumulator is then chosen, bearing in mind that the right hand side may already be in an accumulator. If trace is required then the accumulator must be 7, otherwise if the assignment is of the form wholeword becomes zero, then accumulator 0 is chosen.

If the indicator FFLAG is clear then the expression on the right hand side is copied into the chosen accumulator, unpacked, and rescaled as required. This is omitted if the accumulator is zero. Otherwise if the left hand side is indirectly addressed the entry in ACCS for the chosen accumulator is set to all ones to inhibit the possible choice of this accumulator as a modifier.

If trace is required then the instructions required are output. If the index of the left hand side is in a modifier, this is first dumped. A call on L7 is then output, with the result in accumulator 7, the type/scale in accumulator 6, and the address of the string, which is output to Special Data, in accumulator 5.

If the assignment is to a wholeword then the assignment instruction is then output; this will use one of the four functions STO,STN,ADS, or SSB. Otherwise a modifier is chosen, (this being zero if modification is not required) and the number of shifts to align the expression to the field in which it is to be stored, calculated. If FFLAG is clear then instructions are output to perform the assignment. The expression is first shifted, if required, then masked, and exchanged with the word of
'PROCEDURE' STOREAWAY/

store. This is then masked to clear the field to which the expression is being assigned, and the remaining fields replaced by adding them back to store. Otherwise if FFLAG is set, the constant is shifted and masked to fit the field. If the left hand side is not already in the accumulator then instructions are output to copy it into the chosen accumulator. Provided the assignment is not all ones to the field, an instruction is then output to mask the field. Provided the assignment is not all zeros, then an instruction is output to or the constant into the field. Finally the word is stored.

At the end of the procedure the record of any previous copies of the left hand side is deleted, the right hand Arithmetic Operand deleted from the stack, and the accumulator record for the accumulator ACC updated with the address ADD.

A simplified flowchart is given.

'PROCEDURE' SUBTERM

('VALUE' 'INTEGER' NEG);

This procedure is called after each Term of a subscript expression. From PLUSSUB with NEG equal to zero, and from MINUSSUB with NEG equal to one. This deals with each Term individually in order to reduce the amount of code generated. When the procedure is called, the pointer RH points at the Arithmetic Operand which is the Term of the subscript expression, and the local pointer LH is set to point at the Arithmetic Operand of the variable, (which may be anonymous) which is to be subscripted. If the Term is a constant then this is incorporated in the direct address of the variable. If the term is not a constant, but the indirect address of the variable is null, then this is set to the direct address of the term, which is evaluated in a modifier if required. Otherwise a suitable modifier is chosen, and the term added or subtracted from this modifier. A detailed flowchart is given.

'PROCEDURE' SUSPIEL;

This procedure is called after the expressions for the step and until elements of a For statement have been processed. This alters the commentary of the two top Arithmetic Operands on the stack to "(STEP)" and "(LIMIT)".

'PROCEDURE' SWOP

('LOCATION' 'INTEGER' X,Y);

This procedure exchanges the variables whose locations are passed as the parameters X and Y.

'PROCEDURE' SWOFOPT;

This procedure is called before the application of reversible operators, to reverse the order of the two Arithmetic Operands on top of the stack, if this would reduce the amount of code subsequently generated. The aim is to have the LH operand in the preferred (or any) accumulator, and to avoid having to unpack or rescale the RH operand.
Storeaway

is assignment from dummy? Yes

No

Calculate field optimisation flag (FFLAG)

Calculate address to be stored in accumulator record (ADD)

Select accumulator (ACC)

Output instruction to copy R.H.S. into accumulator if required

Is assignment to partword?

Yes

No

Output store or add/subtract from store instruction

Delete record of copies of LHS
Delete R.H.S. from stack
Update accumulator record

Exit

Delete R.H.S. from stack

Exit

Output trace instructions if required

Is assignment of constant?

Yes

No

Output instructions to:
shift mask exchange mask add back

Shift and mask constant. Output instructions to:
copy LHS mask insert constant store
"INTEGER" "PROCEDURE" TESTSTRING

\((\text{VALUE}\ 'INTEGER\ S1,S2)\);

This procedure compares two strings, pointers to which are passed as the parameters \(S1\) and \(S2\). If the strings are identical the result of the procedure is zero, if they are not identical the result is non zero.

"PROCEDURE" TEXT

\((\text{VALUE}\ 'INTEGER\ STRING)\);

This procedure outputs the string whose address is given by the parameter STRING. The length of the string is obtained by use of the procedure NEXTCHAR, the first character of the string being its length. The characters forming the string are obtained by the repeated use of NEXTCHAR and output by use of the procedure OUTCHAR. It should be noted that the use of NEXTCHAR alters the value of the parameter, STRING. This does not of course alter the actual parameter passed from the point of call, or the string itself, but merely the local value of the parameter. Zero length strings cause no characters to be output.

"PROCEDURE" TEXTLINE

\((\text{VALUE}\ 'INTEGER\ S)\);

This procedure outputs the string \(S\), followed by a newline.

"PROCEDURE" THEAD

This procedure outputs ten inches of blank tape on the punch, using PUNCHLOOP. It also resets the checksum for the characters output on the tape, TSUM, to zero.

"PROCEDURE" TOPACC

\((\text{VALUE}\ 'INTEGER\ ACC)\);

This procedure is used to update the Arithmetic Operand on the top of the stack (i.e. the RH operand), to show that it is now held in the accumulator whose number is given in ACC. This is performed by means of a call of INACC.

"INTEGER" "PROCEDURE" TOSADD;

The result of this procedure is the direct address of the top Arithmetic Operand on the stack, which is deleted.

"PROCEDURE" TRACESTRING

\((\text{VALUE}\ 'INTEGER\ LNO,STRING)\);

This procedure is used to set up calls on trace procedures having a string as a parameter. The library reference number is passed as the parameter LNO, and the compile time address of the string as the parameter STRING. The call of the trace procedure is output using LIBTRACE and then the string is output, destined for the special constants area (Special), using DOSTRING with OUT24 as a parameter.
PROCEDURE TTAIL;

This procedure outputs a tail on paper tape. This consists of ten inches of blank tape, a stop directive (STOP), twenty erases, and another ten inches of blank tape.

PROCEDURE UJBACK;

This procedure outputs an unconditional jump using the variable BJA (back jump address) as the address. This is output where it is required to jump back and repeat after testing the condition for continuation.

PROCEDURE UNARYMINUS;

This procedure is called where it is required to negate the Arithmetic operand on the top of the stack. If it is a constant then it is negated within the compiler. Otherwise instructions are output to negate it and leave the result in an accumulator. This procedure is called from the syntax, and also from SELOPTA,SUBA, and SIMPLEASS in the case of a unary negation operator; and also from CONDMINUS and ZEROCOMP when compiling conditions.

INTEGER PROCEDURE USELAB

(VALUE INTEGER LAB):

This procedure is called, where it is required to use the address of the label whose record is specified by LAB, as the address part of an instruction to be output. The result is either the address of the label, if it is set, or the chain address of the label if it is not. If this is the first reference to the label in the current block then the result is zero, but this requires no special action at the point of call. If the label is unset then the label record is updated, to include the instruction about to be output, on the chain for the label.

PROCEDURE WARN

(VALUE INTEGER TA,TB);

This procedure is called in cases where a recoverable error, but of insufficient gravity to prevent the compiled program from running safely, is discovered. This procedure output the two strings TA and TB, if non zero, separated by " : ". This is followed by a newline, and PRINTBUFF is called to print out the contents of the input buffer.

PROCEDURE XSPARMS

(VALUE INTEGER PTR);

This procedure is called if the procedure identifier whose declaration record is pointed at by PTR is specified to have more than the number of parameters allowed by the implementation. After the sixth parameter, IDTYPE is marked negative. If there is a seventh parameter, a call is made to XSPARMS which prints out a warning message; and also clears IDTYPE so that no further calls are made for the eighth and subsequent parameters.
'PROCEDURE' ZEROACCS;

This procedure is called to clear the array ACCS, which holds a memory of the current contents of the accumulators of the object program. This is necessary where there is a flow of control which invalidates the contents of ACCS.
Label ABSADD:

This section is called in 'ABSOLUTE' communicators, after the integer constant specifying the absolute address for the preceding identifier has been processed. This address is inserted into the direct address field of the identifier record being formed at the base of the stack. The address is masked to ensure that it does not overflow into the tag field, which in this case is zero.

Label ADD:

This section is called to output instructions for the addition of two terms. This is performed by means of a call of ADDSUB with the function code 02 as the parameter. As this function may be applied with its operands in either order, a call is first made to SWOPOPT for optimisation purposes.

Label ADDA:

This section is called after the first term of the expression on the right hand side of an assignment symbol, if the term is unsigned or preceded by a + symbol. If the term is a variable, and is the same variable as the one on the left of the assignment symbol, and is not a partword or of type floating, ASSFUN is set to function 12, so that an "add-to-store" assignment will be used. This section is not called if assignment trace is required.

Label ADDRspec:

This section is called to insert the current special data address into the direct address of the identifier record being prepared at the base of the stack. This is used in all cases of 'SPECIAL' 'ARRAY's, and also for 'COMMON' 'LABEL's, 'SWITCH'es, and 'PROCEDURE's.

Label ADDRSW:

This section is called after the occurrence of the symbol 'SWITCH' in a switch declaration. The address of the (virtual) zeroth element of the switch is calculated, and stored in the base of the stack in preparation for generating the identifier record. The variable TOPTEN is cleared to ensure that the top ten bits of each entry in the switch list are clear.

Label ANScheck:

This section is called at the end of an 'ANSWER' statement, to ensure that the expression is evaluated to the required type and scale, given by EXPSCALE, in accumulator 7. Before returning to the syntax analyser via BEHEADEXIT, to delete the Arithmetic Operand record of the expression; STATUS is set to 2 to indicate that the exit instruction has not yet been output.

Label ANYPREF:

This section is called to select an accumulator for the evaluation of Comparisons within Conditions. An accumulator number is supplied by FINDACC, and stored in PREFACC.
Label ASSCV:

This section is called after the (first) expression of a for element. CCC is set to the TYPEBITS of this expression, only the sign bit (constant marker) being relevant. The instructions to assign this expression to the control variable are output by means of a call of STOREAWAY with a parameter of zero, as trace is not required at this point. If this is the second for element, and the first was simply an expression, FORSTATE will have a value of 2. In this case, if the next symbol is not 'WHILE', a call is made of SETRT, to cause the controlled statement to be treated as an anonymous procedure; and FORSTATE is reset to 4. (If the following symbol is 'WHILE', FORSTATE will be set to 3 by WHILEL.) In order to avoid the inefficiency of treating the controlled statement as a procedure in cases where the for list is of the form:

Expression Comma Expression While

if the first element is an expression, the decision to treat the controlled statement as an anonymous procedure is delayed until after the expression of the second element.

Label ASSTRACE:

This section sets 'BIT'[3]TRACE to the value of ACCUMULATOR. It is called in the case of a 'NO' 'ASSIGNMENT' 'TRACE' directive, in which case the value of ACCUMULATOR will be zero; and also in the case of an 'ASSIGNMENT' 'TRACE' directive, in which case the value of ACCUMULATOR will be 1.

Label AYMCHECK:

This section is used to check that an identifier, which is used without a subscript, does not require one. If it does, a warning message is output.

Label BEGINPROC:

This section is called at the start of a procedure declaration. If the procedure body requires a "jump round" this is output, or if there is a deferred 'ANSWER' action, this is completed. PROCSTACK is called to perform housekeeping on the variables used for procedure declarations. The LINK address is set to the data location which had been allocated to the procedure identifier record, this direct address entry is changed to the current value of the Special Data allocation address, DUECHAIN being also set to this address. This location, which will be used to hold the entry point of the procedure, is allocated, and cleared by OUT24CS. If procedure trace is required, the procedure identifier string is output, destined for the Special Data area, by SPECSTRING, and its address stored in PROCSTRING, which acts as a trace flag.
Label BEGINPROG:

This section is called at the start of a program segment. The heading for the paper tape is output by STARTSEG, and STATUS and DUECHAIN are set to zero. Thus it is expected that a program segment will be entered at its first instruction, even though this will normally be an unconditional jump around procedure bodies. BEGINBLOCK is called to set up an outer block level for the program.

Label BEGINPSPEC:

This section is used to initiate the creation of a parameter specification record of a procedure specified in a communicator, or a procedure parameter declared within a program or Library segment. Calls are made on MAKEPSPEC and EXCPSPEC to accomplish this. If the procedure is a 'VALUE' procedure, a call is made on NEXTPARAM, and a jump made to NEXTPSET, to set up the first (context parameter) entry in the record.

Label BEHEADEXIT:

This section is used by a number of compiling actions to remove the top Arithmetic Operand record from the compiler's main stack, before returning to the syntax analyser. This record, which is no longer required, is removed by calling BEHEAD. When the record is removed, the operand pointers LH and RH are updated by SETLRH.

Label BITDECFAIL:

This section is used by NOBITS, NOAFTER, FIELDPOSW, and UNSFIELD if the values of the integers used in field specifications exceed the allowable range. This section prints a warning message, and returns control to the syntax analyser.

Label BITSIN:

This section is called where a 'BITS' operation is used in an expression, before the expression to which the operation is to be applied has been processed. This section packs the partword and type information, held in BITSPEC and IDTYPE, into BITSHIFT; as these variables may be used during the following Expression. The information packed in BITSHIFT will subsequently be used by RESBITS.

Label CALLPROC:

This section is called when the appropriate number of parameters of a procedure call have been processed. This number may be zero. Starting from the position on the stack given by FPP, there will be a five word Arithmetic Operand record for each parameter. The types of these will have been checked by ANYMORE, and they will all have zero indirect addresses. Thus the actual object to be passed will either be in an accumulator, or obtainable by means of a call of PICK, which may involve unpacking, but which will not require the use of a modifier.
Label CALLPROC:

Each record is examined in turn, (left to right) and the number of the accumulator required, obtained from the PAC field. If this accumulator is already holding another parameter, an instruction is output to exchange this with the accumulator in which the other parameter is required to be passed. This is repeated until the accumulator specified in the current record is either free, or contains the required parameter. A call of PICK is made to copy the parameter from store, if required, and to unpack if necessary.

After this scan has been completed, the required parameters have been placed in the required accumulators with the minimum number of instructions. The call of the procedure is output by OUT27, and ZEROACCS called to clear the accumulator record ACCS; as it is assumed that the contents of all the accumulators (except 7 if typed procedure) will be undefined on return from the procedure.

Label CHECKCV:

This section is called after the control variable of a for statement has been processed. If the variable is a partword, or is not of type integer or fixed, or is an unsubscripted array identifier, an error message is output, and its type changed to integer to allow compilation to continue. The exclusion of floating point control variables is an implementation restriction. In this case it is always faster to increment and test a fixed point variable, and subsequently float it each time, than it is to carry out the whole operation in floating point arithmetic. It also requires fewer instructions, and is often more accurate, as it avoids cumulative round-off and truncation errors.

After checking the type of the control variable, this type and scale is stored in EXPSCALE, and SCALEFIRM set, to ensure that the expressions of the for list are evaluated to the type and scale of the control variable. PREFACC is set to an arbitrary number, and ASSFUN set to function 10. This may subsequently be modified by assignment optimisation. A pointer to the Arithmetic Operand record of the control variable is placed in CV, and a call made of DIAG to output diagnostic information. This section returns to the syntax analyser via FORESTORE, to store the index expression of the control variable, if one has been evaluated.

Label CLEARTYPE:

This section is called to clear the variable IDTTYPE, prior to assembling type and scale information, in declarations, specifications, and typed primaries.

Label CODELAB:

This section is called after the occurrence of a label preceding an instruction in code. This action is performed by calling FIXLABEL with the result of a call of LOOKUPLAB as the parameter. This is similar in action to SETLAB, but does not generate any label trace instructions, even if 'LABEL' 'TRACE' is specified at this block level.
Label CODESHIFT:

This section is called to set up an anonymous Arithmetic Operand for a shift instruction in code. A call is made to NONAME to set this up, and the commentary of this record changed to "(SHIFT)". This section returns to the syntax analyser via INCTOS which deals with any integer constant specifying the number of places of shift.

Label COMOFF:

This section is called at the end of a 'COMMON' segment. A check is first made to ensure that the (hopefully) unique checksum for this common segment is not zero, this checksum then being output to the loader. The output is completed by a call of FINISHSEG, and the COMMON flag cleared. A scan is then made of the DECLIST to re-address the common entries, and to output a common-check tape. Whilst a 'COMMON' segment is being compiled, all addresses within it are allocated with respect to its Data and Special Data areas (tags D and S). When the size of these are known, the addresses may then be relocated relative to the start of the segment (tag C). N.B. The compiler assumes that the Special Data area of a common segment immediately follows the Data area.

Label COMON:

This section is called at the start of a 'COMMON' communicator segment. A check is first made to ensure that no other common specifications exist for this run of the compiler; program segments being only allowed to be compiled using one common segment. After this check, the COMMON flag is set, and the heading output by STARTSEG.

Label CONDADD:

This section is called after the second, and each subsequent, term of the expression on the right hand side of a comparator has been processed; if the term is preceded by a + symbol.

In the interests of efficiency, the expressions on either side of the comparator are evaluated as one expression and the result tested against zero. Provided that the comparator is neither > nor <=, the Terms on the right which are preceded by a + sign are subtracted, and those which are preceded by a - sign are added. Otherwise the left hand side is negated, and the expected add and subtract operations used.

This section jumps to either ADD or SUB depending on the comparator used.

Label CONDMINUS:

This section is called after the first Term of the expression on the right hand side of a comparator has been processed; if the Term is preceded by a - sign. In the interests of efficiency, the expressions on either side of a comparator are evaluated as one expression, and the result tested against zero.

If the comparator is neither > nor <=, this section jumps to ADD to output instructions to add this term to the expression.

(Addition being taken as being equivalent to negation and subtraction)

Otherwise the expression on the left is negated by means of a call of UNARYMINUS, the operand records interchanged by PERM, and a jump made to SUB.
Label CONDPLUS:

This section is called after the first Term of the expression on the right hand side of a comparator has been processed; if the Term is not preceded by a - sign. In the interests of efficiency, the expressions on either side of a comparator are evaluated as one expression, and the result tested against zero.

If the comparator is > or <= PERM is called, to exchange the Arithmetic Operand records of the left hand expression, and the right hand term. This changes the type of test required to one which is available in the machines instruction set. A jump to SUB is made, to output instructions to subtract the two operands.

Label CONDSUB:

This section is called after the second, and each subsequent, Term of the expression on the right hand side of a comparator has been processed; if the Term is preceded by a - sign.

(See CONDADD above)

Label CONSTANT:

This section is called where a numeric constant occurs in an expression, and also as an operand in code. It sets up an Arithmetic Operand record on the stack. The commentary is set to "(CONST)", and the TYPEBITS word set to the type of the constant, NUMBERSCALE, which has the sign bit already set. This denotes that the operand is a constant whose value is stored in the DIRADD field of the record.

Label DECSIZE:

This section is called after declarations of single word variables, to record the number of Data locations required.

Label DIVIDE:

This section is called to output the instructions for the division operation. The scale and type to which a division operation is carried out depends on the scales and types of the operands, and also the context in which the division occurs. The local procedure ENVTST is used for testing the "environment" of the division. If the division occurs as the last operator in a Term occurring in an expression which is required to be evaluated to a specified scale, or the proposed scale SCALE has fewer integer bits than the current value of the expression scale EXPSCALE this returns a result of 1, otherwise 0.

Before the calculation of the scale to which the division will be performed is carried out, a check is made on the operands. A call is made to ISINACC in case there is a copy of the left hand operand in an accumulator. If the right hand operand is a location, or requires unpacking, a call is made on PICK to copy it into an accumulator. The syntax of this implementation specifically prohibits multiplication or division of addresses, but this may still be performed, if the address is within a bracketed expression.
Label DIVIDE/

If either of the operands, or EXPSCALE, is of type floating, then the division is carried out in floating point arithmetic, by means of a call of FLOATOP, with a parameter of 4.

If the right hand operand is an integer constant which is a power of two, the division will not be performed using a divide instruction. If the left hand operand is of type integer, this will be performed by shifting, which will be combined with any shifts required for unpacking and rescaling. If ENVTEST returns a zero answer, the result is evaluated to integer scale, otherwise the integer is rescaled to the (fixed) scale given in EXPSCALE. If the left hand operand is of type fixed, the division is carried out by adjusting the scale field of its operand record. In this case the number of significant bits remains unchanged.

If both operands are of type integer, the division will be carried out as an integer division, unless ENVTEST indicates that there is an overriding scale requirement. The number of significant bits of the result will not be specified unless both the operands have the number of significant bits specified. In this case the number of significant bits is given by the number of significant bits of the left hand operand minus the number of significant bits of the right hand operand plus one to make allowance for the range of the result, plus one for the sign bit. If this is negative, it is set to zero (unspecified) and is otherwise subject to a limit of 24.

If the left hand operand is of type fixed, and the right hand operand of type integer, the division is evaluated to the scale of the left hand side, unless ENVTEST indicates that there is an overriding scale requirement or the integer is a constant. In this case the integer constant is converted to the equivalent fixed constant, and the division performed as a division of a fixed by a fixed. The use of the left hand scale is based on the assumption that the integer will be small. If this is not the case, the division should be bracketed and typed.

If the left hand operand is of type integer, and the right hand operand of type fixed, and there is no overriding scale requirement, the scale to which the division will be performed is based on the scale obtained if the left hand operand were to be converted to fixed, and the division treated as a division of a fixed by a fixed. The rescaling is not performed at this point.

The final case is the division of a fixed by a fixed. Unless ENVTEST indicates that EXPSCALE should be used, the scale of the result has the smaller number of significant bits of the two operands, and the number of integer bits is the number of integer bits of the left hand operand minus the number of integer bits of the right hand operand. In a simple case of division this will not give rise to rescaling.

When the scale to which the division will be performed has been decided, the number of shifts right prior to the division, SHIFTS, is calculated. If this is less than 23, an instruction is output to clear the auxiliary register. If the number is negative (shift left required), this is combined with any rescaling and unpacking required, by means of a call of PICK. Otherwise a call is made of GOODPICK to ensure that the left hand operand is in an accumulator, this being followed by a separate right shift instruction. In certain cases this may lead to slightly inefficient code.
Label DIVIDE/

The instruction to perform the actual division is then output. If the right hand operand is in the preferred accumulator the two operand records are now interchanged by PERM. The final instruction output is to copy the result of the division, which is in the auxiliary register, into the accumulator whose number is given in the left hand Arithmetic Operand record. This section returns to the syntax analyser via BEHEADEXIT, which removes the top Arithmetic Operand record from the stack.

Label DOCS:

This section is called after the occurrence of the symbol 'DO', which terminates the Forlist of a Forstatement. The value of FORSTATE, on entry, indicates the type of the last element and the complexity, of the Forlist. If FORSTATE is greater than three, the controlled statement is to be treated as an anonymous procedure, and instructions will be required to store, and exit on, the link.

This section outputs any instructions required for the testing of step-until elements, and for the calling of, or jumping round, the controlled statement. It also reserves any locations required for

1. The index of the control variable, if subscripted.
2. Holding step and limit values, if they have been evaluated.
3. The link, if the controlled statement is treated as an anonymous procedure.

Any other workspace, temporarily used in the evaluation of expressions in the Forlist, is relinquished. DATASTART, DATAPTR, and DATAMAX are updated as required. The locations reserved will be returned at the end of the controlled statement, and the above variables reset. If 'LOOP' 'TRACE' is required, a call of L9 is generated, with the value, type, and identifier string, of the control variable as parameters.

If the value of FORSTATE is greater than four, on entry, it will be set to four by FORCOM. The value of five is then used to indicate that the special case of a single step-until element, with three constants, has been detected. This will result in the test and incrementation of the control variable taking place at the end of the controlled statement, and consequently more efficient code being generated. A flowchart for this section is given, and the values of FORSTATE and the corresponding of Forlist is tabulated below:

<table>
<thead>
<tr>
<th>FORSTATE</th>
<th>Entry</th>
<th>Exit</th>
</tr>
</thead>
<tbody>
<tr>
<td>0</td>
<td>0</td>
<td>0</td>
</tr>
<tr>
<td>1</td>
<td>1</td>
<td>1</td>
</tr>
<tr>
<td>2</td>
<td>2 or 5</td>
<td>2</td>
</tr>
<tr>
<td>3</td>
<td>3</td>
<td>3</td>
</tr>
<tr>
<td>4</td>
<td>4</td>
<td>4</td>
</tr>
<tr>
<td>5</td>
<td>4</td>
<td>4</td>
</tr>
<tr>
<td>6</td>
<td>4</td>
<td>4</td>
</tr>
</tbody>
</table>

Expression
Expression 'WHILE'
Expression 'WHILE'
Expression, Expression 'WHILE'

(Excluding case 3)
Label DOSHIFT:

This section is called to apply shift operators ("SRA", "SLA", "SRL", and "SLC") using the top two Arithmetic Operands on the stack. Before this can be accomplished, a call has to be made to KILEXPR to remove the expression level set up for the optimisation of the expression giving the number of places of shift. The instructions for the shift operation are then output by OPERATE, which is passed the required function, held in BITSHIFT, as its parameter.

Label ELSEFI:

This section is called after both the consequence and the alternative expressions of a conditional expression. It ensures that each is evaluated to the scale given by EKSSCALE, in the accumulator given by PREFACC. A call is made to REVCHAIN so that the chain of the jump to skip over the expression will be set by the next action. (ELSEX or FIEX)

Label ELSES:

This section is called following the 'ELSE' in a conditional statement. A call is made on REVCHAIN, as the chain which holds the jumps to skip over the first part of the 'IF' statement is now due to be set. If STATUS is 4, then the statement between the 'THEN' and the 'ELSE' is void, and the condition is reversed using REVCJ.

If STATUS is 3, the statement is of the form:

  'IF' . . 'THEN' 'GOTO' label 'ELSE'

in this case a jump is not required to skip over the following statement; STATUS being set to 0 to indicate this.

If STATUS is 0, the statement preceding the 'ELSE' is a normal statement, and a skip is required over the following statement; STATUS being set to 3 to indicate this.

The remaining two cases, STATUS of 1 and 2, require no action at this point.

Label ELSEX:

This section is called following the occurrence of the symbol 'ELSE' in a conditional expression. It outputs a jump around the alternative expression, by means of a call of OUTUJ. The chain from the jump(s) around the consequence expression is set by SETDUE.

SCALEFIRM is set to ensure that the alternative expression is evaluated to the same scale as the consequence expression. This section returns to the syntax analyser via BEHEADEXIT, to remove the Arithmetic Operand record for the consequence expression from the stack.
Label ENDARRAY:

This section is called after the end of a set of arrays having common bounds. This may be followed by a further set. If the declaration is not overlayed, the total data requirement for the arrays, which is the product of the number of arrays and the sizes of all the dimensions, is given in ARRAYS, which is used to increment DATAMAX. If the declaration is overlayed, OVERBASE is incremented by ARRAYS.

Label ENDECS:

This Section is called after the end of the declarations at the head of a block. DATAPTR and DATASTART are set to the current data allocation address held in DATAMAX. These variables will be used for the allocation of anonymous workspace for temporarily dumping the contents of accumulators. DATAPTR will contain the next address to be allocated, and will be reset between statements; its maximum value being recorded in DATAMAX. After the declarations at the head of a block, normal presetting (if allowed at this level) is no longer allowed, and the flag PRESETOK is cleared.

Label ENDFOR:

This section is called after the completion of the controlled statement, to output the instructions required to complete the loop. The action required depends on the value of FORSTATE set up by DOCS, which indicates the type of the last element and the complexity of the Forlist.

If FORSTATE is one or three ('WHILE'), an unconditional jump back, to repeat the assignment and test the condition, is output. If however STATUS is non-zero, this jump is not required, and is suppressed. Any chain due to be set after the controlled statement is set to the repetition address, held in BJA.

If FORSTATE is two ('STEP' 'UNTIL'), the instructions are output to increment the control variable and to jump back to the test against the limit.

If FORSTATE is four the controlled statement is treated as an anonymous procedure, and an instruction to return using the stored link is output.

If FORSTATE is five, this indicates the special case of 'STEP' 'UNTIL' with three constants. If a copy of the control variable does not exist in an accumulator, then an instruction is output to copy it into accumulator 7. Instructions are then output to add the step value to the control variable and to store the result. An instruction is then output to subtract the limit value, if non-zero, and a test instruction output. The sense of this test instruction depends on the sign of the step. If the step is positive, a jump if zero or negative is required. As this is not available, the limit has one added to it, so that the zero case is included in the negative test. The final part, common to all cases, completes the housekeeping for this level of 'FOR' statement. The SKIPCHAIN, which is used by the final test, is copied into DUECHAIN, and SKIPCHAIN reset to its value before the statement, which is held in COPYSKIP. DATASTART is reset to release any anonymous locations used, and the previous values of any enclosing 'FOR' statement reset by means of a call of OFFSTACK. Finally, a call is made to ENDMABBLOCK, to remove the label block level which prevented illegal entry to the controlled statement.
Label ENDP:  

This section is called at the end of the body of a procedure declaration, in program and library segments.

EXCPSPEC is called to prepare for the completion of its parameter specification record, and STATUS is set to 1 to indicate that the following instruction is not directly entered. LOCALLIMIT is reset to point at the procedure identifier record, this value being set into DECLIST by the call of ENDBLOCK, to complete the block forming the procedure body. The values of the eight locations used by procedure declarations, and chained up by PROCCHAIN, are restored by means of a call of OFFSTACK, and finally the parameter specification record is completed by means of a call of FINISHPSPEC. It should be noted that once a procedure body has been compiled, only its parameter specification record remains, and not the identifier records of its parameters.

Label ENDP:  

This section is called before the final 'END' symbol of a program segment. If the program does not end with a jump, but "drops off the bottom", a call of Li is generated.

A scan is then made of all the communicator specification records. This detects entries required to be inserted in the 'COMMON' segment, and labels used which are external to this program segment. In this latter case, the label is set to the current program segment, and an indirect jump to the label output. ENDBLOCK is then called to perform the normal end of block actions. After this call, DATAMAX will hold the total data requirement for the segment, and LABDECLIST will point at a list of any unset labels. These are dealt with by the call of FINISHSEG, which also completes the paper tape output for this segment. A flowchart is given for this section.

Label ENDSPEC:  

This section is used to terminate the generation of a parameter specification record of a procedure specification. This is accomplished by means of calls of EXCPSPEC and FINISHPSPEC. This resets the values of IDTYPE and PARAMPTR, which will be required if the procedure (parameter) specification forms part of a procedure declaration.

Label ENDST:  

This section is called at the end of each statement compiled, to reset the variable DATAPTR, used for the allocation of anonymous workspace, to its "home" value, held in DATASTART. As anonymous workspace is not required to be carried from one statement to the next, it may be re-used in following statement.

Label EXITCHECK:  

This section is called at the end of compilation of a non-code procedure. If the procedure is untyped, i.e. does not deliver a result directly, a check is made to test whether the last instruction of the procedure was an unconditional jump, and that a return jump is not required. If this is the case no action is taken by this section; otherwise a call is made on SETDUE to set any chains. If, on the other hand, the procedure is typed, and is therefore required to deliver an answer; a check is made
Label EXITCHECK/

to ensure that the procedure contained at least one 'ANSWER' statement, and that the last effective statement of the procedure was either a 'GOTO' or an 'ANSWER'. If this is not the case, an error message is output. If the procedure is typed, and trace is required, a call is made on L6 with the result and its scale as parameters. In the case of a 'VALUE' procedure, the scale is obtained dynamically from the first (context) parameter, which must therefore not be altered by the procedure. If there is more than one 'ANSWER' statement, these will be chained up using EXITCH, which is set to the call of L6. Irrespective of whether the procedure is typed, if trace is required, a call on L5 is output, with the procedure identifier string as the parameter. It is assumed, in the case of typed procedures, that L5 will reset Q7 to the value supplied to L6. Finally, an exit instruction is output, using LINK as the address.

Label EXPRTYPE:

This section is called after the completion of a typed bracketed expression, to ensure that the expression is evaluated to the specified type and scale. This may not have already been performed if the expression did not contain addition or subtraction operations. The expression is evaluated to the required scale by means of a call of PICK. The section OFFSTEXPR is then entered to remove the expression level from the stack.

Label EXTADD:

This section is called, in 'EXTERNAL' communicators, after the integer constant(s) specifying the external relocator (and displacement) have been processed. The fourteen bit address is formed from the five bit relocator number, and the nine bit displacement to which is added 256. This address is masked, and the external tag(6) added. The address is then inserted into the direct address field of the identifier record being formed at the base of the stack.

Label FI:

This section is called at the end of a conditional statement, irrespective of whether it had an 'ELSE' part. If STATUS is 4, then the statement is void, and the test instruction is only output if FUNCTION is 24, the statement being of the form:

'IF' 'OVERFLOW' 'THEN' ;

which may be used to clear overflow. If STATUS is 3 or 4, the 'ELSE' part has been omitted, and STATUS is set to 0.

The SKIPCHAIN is joined to the DUECHAIN, if required, by means of a call of JOINCHAINS. The stacked value of SKIPCHAIN, chained up with IPCHAIN, is restored to its value prior to the 'IF'.

173
Label FIELDISP:

This section is called after the integer giving the Wordposition of a table field has been processed. The address of the table field is calculated and stored at the base of the stack in preparation for the identifier record for the field being moved onto the main stack declaration list by NEWNAME.

Label FIELDPOSN:

This section is called when the integer specifying the starting bit position (Bitposition) of a partword is processed. This integer, together with the result held in BITSPEC allows the two shift fields of the partword descriptor to be calculated. These fields are packed in BITSPEC, and are respectively:

\[ \begin{align*}
\text{MSS} & : \text{(most significant shift)} \\
& \text{shifts to align field to top of word} \\
\text{LSS} & : \text{(least significant shift)} \\
& \text{shifts to align field to bottom of word} \\
& \text{are calculated as follows:} \\
\text{MSS} & := \text{Bitposition (of most significant bit)} \\
\text{LSS} & := 24 - \text{Totalbits - Bitposition} \\
& \text{using Ferranti bit numbering, sign bit = bit 0.}
\end{align*} \]

If, after the above calculations are performed, LSS is negative, the partword specification is illegal, i.e. a field has been specified which "hangs off" the word. In this case the error is reported by BITDECFAIL.

Label FIEX:

This section is called at the end of a conditional expression. The instructions to evaluate both expressions in PREFACC will have been output by ELSEFI. The chain holding the jump around the alternative expression is set by SETDUE. The variables of any enclosing condition are restored by OFTSTACK, and DUECHAIN is reset.

An Arithmetic Operand record for the result of the conditional expression is set up by a call of GRABVAR, and the TYPERET field set to EXPSCALE. The commentary is set to "(IFEX)", and a call made to TOPACC, with PREFACC as the parameter, to update the DIRADD field and also ACCS.
Label FINISHPROC:

This section is called after the call of a procedure has been generated, or after the final call of a multiple call. It reduces the procedure call level by a call of OFFSTACK. The PARAMSPEC field of the arithmetic operand record for the procedure is cleared, to avoid its interpretation as PARTWORD. This record and ACCS are updated by a call of TOPACC with a parameter of 7, to indicate that the result (if any) of the procedure is in accumulator 7. The TYPEBITS of this record will have been set by SETUPPROC. If the procedure is untyped, this record will be nonsensical, but will be immediately deleted by BEHEAD.

Label FIRSTDIM:

This section is called to process the first dimension of multi-dimensional arrays. ADDRARRAY is called to insert the addresses of the first level Iliffe vectors in the identifier records on the stack. The number of records to be processed is given in ARRAYS. After the call of ADDRARRAY the number of entries required in the first level Iliffe vector is given in ARRAYS, which has been multiplied by NUMBER in DODIM.

Label FORSTORE:

This section is entered from STEPUNT and from CHECKCV. It arranges to store in anonymous workspace the index of the control variable, if subscripted, and the values of the step and limit expressions if they are required to be evaluated. These are stored by means of a call of DUMPACCS, which is a brute force method of storing any information held in the accumulators. If an anonymous location is used by DUMPACCS, DATAPTR will be incremented. This value is copied into DATASTART so that this location will not be overwritten if space is required temporarily during the evaluation of the next expression. After each expression, DATAPTR is reset to the current value of DATASTART to free any location so used. RTL (routine link) is set to the maximum value attained by DATASTART, in case the controlled statement is to be used as an anonymous procedure. The current program address is recorded in BJA (back jump address), for use as a point to jump back to, in order to perform the required test before each repetition of the controlled statement. As a jump may be made to this point, a call is made of ZERACCs.

Label FORTRACE:

This section sets 'BIT'[2]TRACE to the value of ACCUMULATOR. It is called in the case of a 'NO' 'LOOP' 'TRACE' directive, in which case the value of ACCUMULATOR will be zero; and also in the case of a 'LOOP' 'TRACE' directive, in which case the value of ACCUMULATOR will be 1.
Label GOTOL:

This section is called to process a 'GOTO' statement, where the destination is a label identifier and not a switch entry. The action depends on the value of STATUS, which is used as an index of the switch S. The actions taken in this section are designed to reduce as far as possible, the number of redundant unconditional jumps generated. These actions depend on the context in which the 'GOTO' statement occurs. (See 'PROCEDURE' STATUSCHECK). A flowchart is given for this section.

Label GOTOSK:

This section is called to output the jump instruction in the case of a 'GOTO' statement, where the Destination is a switch entry. This is performed by means of a call of INST with the RH arithmetic operand giving the address, and using function 27. Note that accumulator 4 is used, this is not necessary, but is in conformity with the other uses of function 27. The type checking of the switch identifier has previously been carried out by LABSK. The Arithmetic Operand Record created is treated in a similar manner to any other record, thus allowing the application of subscript optimisation. Before returning to the syntax analyser via BEHEADEXIT, to delete the record; STATUS is set to 1 to indicate that the following statement is not directly entered.

Label IFEX:

This section is called following the occurrence of the symbol 'IF' at the commencement of a conditional expression. Any partially evaluated results held only in accumulators are dumped by DUMPACCS. The chain DUECHAIN is stored in COPYDUE, and a call made of ONSTACK to stack:

IFCHAIN, SKIPCHAIN, COPYDUE and RELOP.

This is only required if the conditional expression occurs within another condition, but is carried out in all cases. DUECHAIN and SKIPCHAIN are then cleared in preparation for the following condition.

Where a conditional expression occurs in a context of defined scale, SCALEFIRM is non zero, and both the consequence and alternative expressions will be evaluated to this scale. Where the scale is not defined, the alternative expression will be evaluated to the final scale of the consequence expression. In either case, the expressions will be evaluated in the accumulator whose number is given in PREFACC.
Label IPS:

This section is called at the start of an 'IF' statement. A call is made on STATUSCHECK to complete any deferred actions, and to ensure that the statement can be entered. STATUS is then set to 4 to ensure completion of the following Condition.

The current SKIPCHAIN is stacked using IFCHAIN, and SKIPCHAIN set to zero. (DUECHAIN has been set to zero by STATUSCHECK). This clears the chains in readiness for the following Condition.

Label INCTOS:

This section is called to increment the direct address field of the Arithmetic Operand record of a code instruction by the integer constant held in NUMBER. This is performed by means of a call of ADDADD.

Label KILLPARAMS:

This section is called when it is discovered that a procedure body is a code statement. The allocation of the data space to hold the parameters and the link is cancelled, the parameter identifier records removed from DECLIST, and LINK and LABDECLIST (which hold records of 'LABEL' parameters) cleared.

Label LABL:

This section is called following the symbols 'LABEL'(Identifier) and after a jump instruction in code. It creates an Arithmetic Operand record of type integer, with the marker LBM set, to denote that the direct address contains a pointer to a label record.

Label LABSK:

This section is called after the occurrence of a subscripted identifier in a label context, to check that the identifier is the name of a switch. After the check, its type is changed to integer to allow its use in expressions by the use of 'LABEL'.

Label LABTRACE:

This section sets 'BIT'[0]TRACE to the value of ACCUMULATOR. It is called in the case of a 'NO' 'LABEL' 'TRACE' directive, in which case ACCUMULATOR will be zero; and also in the case of a 'LABEL' 'TRACE' directive, in which case the value of ACCUMULATOR will be 1.

Label LASTDIM:

This section is called to process the last dimension of multi-dimensional arrays. It uses DODIM to set up an Iliffe vector in Special Data, the number of entries of which is given by the current value of ARRAYS, which is the product of the number of arrays and the sizes of all previous dimensions. This Iliffe vector is required to be relocated with respect to the starting address of the Data area.
Label LESBITS:

This section is called where it is required to assign an expression to a partword field of a variable. A check is made to ensure that the variable is not itself a partword table field, and is not of type 'FLOATING'. If this is discovered, an error message is output. (In the first case the assignment is inefficient and ambiguous, and in the second case could give rise to unnormalised numbers. In the first case the required sub-field should be specified as a separate field, and in the second the overlay facility may be used, if this action must be carried out.)

The type of the variable is changed to the integer type set up in IDTYPE, and the required field specification inserted in its PARTWORD field.

Label LHSPROC:

This section is called where a statement commences with an identifier which is not followed by a colon (label) or a becomes symbol or subscript (assignment). It checks that the identifier is the identifier of a procedure (of any type). If not, a warning message is output, and the PARAMSPEC field set to point to the appropriate dummy parameter specification record at the base of the stack, in order to avoid a subsequent syntax failure.

SCALEFIRM is cleared in case the procedure is a 'VALUE' 'PROCEDURE'. Although the result is not required, in this case it will be evaluated to type floating for the sake of consistency.

Label LIBADD:

This section is called in 'LIBRARY' declarations and communicators, after the integer constant specifying the Library reference number for the preceding identifier has been processed. This number is masked to ensure that it does not overflow into the tag field, which is set to 5. This address is inserted into the direct address field of the identifier record being formed at the base of the stack.

Label LOOKUPD:

This section creates an Arithmetic Operand record for a variable, and ensures that its address is placed in the DIRADD field. This section is called to obtain an address to be used either as an overlay base, or as an address part of a code instruction.

Label MIDDIM:

This section is called to process all the dimensions except the first and last of three or more dimensional arrays. It uses DODIM to set up an Iliffe vector in Special Data, the number of entries of which is given by ARAYS. This Iliffe vector is required to be relocated with respect to the starting address of the Special Data area.
Label MINUSSUB:

This section is called, in a subscript expression, after each Term preceded by a - symbol. This section calls SUBTERN, which performs subscript arithmetic and optimisation.

Label MOREFOR:

This section is called after the occurrence of each comma separating elements in a for list. The value of FORSTATE gives the type of the preceding element(s).

If FORSTATE is zero, this occurring if the preceding element consists simply of an expression, and is the first element; an unconditional jump instruction is output, its address recorded in ETA, and FORSTATE is set to 2. Otherwise, if FORSTATE is less than four, the controlled statement is to be treated as a procedure, and a call is made of SETRT.

A call is made to FORCOM to output a call of the controlled statement, and other associated instructions. This is followed by calls of REVCHAIN and SETDUE to set the SKIPCHAIN, which is used when the previous element is exhausted, to the current program address. The current program address is recorded in BJA, for use if the following element is a while element, and ZEROACCS is called. ASSFUN is reset to its initial value of function 10, and DATASTART reset to its starting value, making allowance only for the index of the control variable. If the control variable is subscripted, and the subscript has been evaluated and stored in an anonymous location, the "hole-marker" flag in its indirect address will be set.

Label MPY:

This section is called to output instructions for the multiplication of two operands. As the operation may have the operands in either order, an attempt is made to have them in the order which would generate the least number of instructions. If the left hand operand is a constant the order is reversed, because the operand will not require to be unpacked, and may be an integer constant which is a power of two. Otherwise SWOPOPT is called. An accumulator is then chosen, using GOODACC.

The type and scale to which the multiplication is carried out depends only on the types and scales of the operands, and not on the context. It is probably better, with fixed point operands which are held left justified, to multiply first and then rescale if required.

If either operand is of type floating, then the operation is carried out in floating point arithmetic, by means of a call of FLOATOP, with a parameter of 1.

If the right hand operand is an integer constant which is a power of two, the multiplication will not be performed using a multiply instruction. If the left hand operand is of type integer, this will be performed by shifting, which will be combined with any unpacking shifts required. The number of places of shift will be added to the number of significant bits of the operand if this is non zero. If the left hand operand is of type fixed, the multiplication is carried out by adjusting the scale field of its operand record. In this case the number of significant bits in the record remains unaltered.
Label MPY/

If both operands are of type integer, the multiplication will be carried out as an integer multiplication, delivering an integer result. This is unfortunately rather inefficient if overflow is to be correctly detected. If the number of significant bits of both operands is given, the number of significant bits of the result is given by the sum of their significant bits, with allowance being made for the sign bit, subject to the restrictions on the word length. Otherwise, the number of significant bits of the result is unspecified.

If one operand is of type integer, and the other of type fixed, the integer operand is converted to the equivalent fixed scale. This is performed by rescaling, if the operand is a constant, or by shifting if the operand is not. In this case the number of significant bits of the integer operand must be specified.

If both operands are, or have been converted to, type fixed, the scale of the result is calculated from their scales and numbers of significant bits. The number of significant bits of the result is taken to be the smaller number of the two operands, and the scale calculated has its number of integer bits equal to the sum of the integer bits of the two operands. As this may exceed the range allowed by the implementation, a check is made, and a warning message output if required.

After the type and scale of the result has been calculated, the actual instructions for the multiplication are then output. The left hand operand is copied into the chosen accumulator, and if the right hand operand is required to be unpacked, or is marked as being a location, this is copied into another accumulator. The instruction is then output to perform the multiplication. If the type of the result is integer, three instructions are output to shift the result back into the accumulator. This section returns to the syntax analyser via BEHEADEXIT, which removes the top Arithmetic Operand record from the stack.

Label MSK:

This section is called to apply the 'MASK' operation to the two top Arithmetic Operands on the main stack. This is performed by means of a call of OPERATE, with function 15 as the parameter. As the function is reversible, i.e.

\[ A \text{'MASK'} B = B \text{'MASK'} A \]

for all values of A and B, a call is first made on SWOPOPT to reverse the order of the operands, if this would result in less code being generated.
Label MULTICALL:

This section is called after each call of a multiple call of a procedure, except the final call. It resets the STACKPOINTER, to delete the previous set of parameters, and resets the FSP pointer to the parameter specification record.

If the procedure is typed, the result of the previous call will be used as the first (explicit) parameter of the next; and the type of this parameter must exactly be the same type as the type of the procedure, which must have more than one parameter. If the procedure is not a 'VALUE' procedure, the record for this parameter is set up by a single call of MAKEPARAM, with a parameter of zero. In the case of a 'VALUE' procedure, three calls are made to MAKEPARAM with parameters of 1, 0, and 1. This sets up three parameters which are, the type required for the result, the result of the previous call, and its type. Note that when compiling a call on a 'VALUE' procedure, a marker is set in PNP by SETUPPROC.

Label NEGNUM:

This section negates the variable NUMBER. It is called, in cases where a signed constant not occurring in an expression is preceded by a minus sign, after the number has been assembled.

Label NEQ:

This section is called to apply the 'DIFFER' operation to the two top Arithmetic Operands on the main stack. This is performed by means of a call of SWPOPT, for optimisation purposes, followed by a call of OPERATE, with function 16 as the parameter.

Label NEWNAME:

This section is called after the occurrence of an identifier in a declaration. It creates an Identifier Specification Record on the stack. A check is made to ensure that this is the only occurrence of the identifier at this block level, and that it is not the (illegal) seventh parameter of a procedure. The record is then placed on the stack by ONSTACK, the relevant information being held in the "declaration breeding ground" at the base of the main stack. If the identifier is an indirectly addressed parameter of a procedure, the direct and indirect addresses in the record are exchanged, and the location marker deleted.

If the identifier is a parameter, a call is made to NEXTPARAM, to increment the Parameter Specification record for the procedure; and if the parameter is a 'LABEL', a label record is created by LOOKUPLAB, its reference being temporarily stored in the INDADD field of the parameter identifier record.
Label NEXTPSET:

This section is called between sets of parameters in procedure declarations, between parameters in parameter specifications, and after the context parameter in declarations of 'VALUE' procedures. This section clears the type and other fields of IDTYPE, leaving only the parameter accumulator field PAC, and the sign bit marker, unchanged. This allows the type of the following parameter to be built up in the usual manner.

Label NISACC:

This section is called where the operand of an instruction in code is an accumulator. An Arithmetic Operand record is set up by NONAME, and the accumulator number (with AMARK added) is inserted in the DIRADD field.

Label NISMOD:

This section is called where the operand of an instruction in code contains a modifier. After checking that the accumulator specified is the number of a modifier, this number (with AMARK added) is stored in the INDADD field of the Arithmetic Operand record of the instruction.

Label NOAFTER:

This section is called after the integer specifying the number of bits after the binary point has been processed. This occurs in 'FIXED' numbertypes, and also in table fields. The scale is calculated and inserted into the PVL field of IDTYPE, the TYP field being set to the value for 'FIXED' variables (1). The scale (or P value + 32) is given by:

\[
\text{Total bits} - \text{Number after point} + 31
\]

a check is made to ensure that this lies within the allowable range for this implementation. (0 to 63)

Label NOBITS:

This section is called after the integer specifying the number of bits has been processed. In the case of normal declarations this specifies the number of significant bits required to hold the number (although a whole word will be used), and in the case of table fields and 'BITS' operation this specifies the number of bits in the partword. After checking that the number of bits is within the allowable range, the number is inserted into the SGB field of IDTYPE.

Label NOSIG:

This section is used in the formation of a partword descriptor held in BITSPEC. It is called after the occurrence of the integer specifying the number of bits required by a table field and also by a 'BITS' operation. The number of bits is subtracted from the number of bits in the computer word, and stored in BITSPEC. No check is made at this stage. The number held in BITSPEC will be subsequently used by FIELDPOSN.
Label OFFSTEXPR:

This section is called at the end of an untyped bracketed expression, and is also entered at the end of a typed bracketed expression. It removes one stack level for expressions, resetting the five variables starting with EXPRCHAIN. The top Arithmetic Operand record is moved down, and a call made on SETLRRH to reset the LH and RH pointers. STACKPOINTER is updated, and a call made on ACCUPDATE to update ACCS.

Label ONEBIT:

This section is called where the number of bits specified in a 'BITS' operation has been omitted, and is assumed to be one. The action of this section is identical to the actions carried out by NOBITS,NOSIG and PARTINT where NUMBER is one. BITSPEC is set to 23, the SGB field of IDTYPE set to one, and the PVL field set to integer scale.

Label ONEDIM:

This section is called to process the sole dimension of single dimensional arrays. ADDRARRAY is called to insert the addresses of the Data space allocated to the arrays, in the identifier records on the stack. The number of records to be processed is given in ARRAYS.

Label ORACT:

This section is called following the symbol 'OR' in a Condition. It outputs the test instruction, reversing the test, with the address part chained on the DUECHAIN; and sets the SKIPCHAIN. REVCHAIN is called (twice) to ensure that the procedures OUTCJ and SETDUE operate on the required chains.

Label ORP:

This section is called to apply the 'UNION' operation to the two top Arithmetic Operands on the main stack. This is performed by means of a call of SWOPOPT, for optimisation purposes, followed by a call of OPERATE, with function 17 as the parameter.

Label OUTCODE:

This section is called at the end of each instruction in code to output the completed instruction. The accumulator number is held in ACCUMULATOR, and the function number in FUNCTION. The top (RH) Arithmetic Operand record on the stack contains the information required for the N (address) and M (modifier) fields. The instruction is output by INST after the LCM field of the record has been cleared, as this may have been set by LABL or STRINGEX. This section returns to the syntax analyser via BEHEADEXIT which removes the top Arithmetic Operand record from the stack.

Label OUTPRES:

This section is called after each numeric value or void in a Presetlist. The constant is rescaled, if required, by SCALECON, and output by OUTPRESSTD.
Label OVEROFF:

This section is called at the end of an overlay declaration, and clears the OVERLAY flag.

Label OVERON:

This section is called after the symbol 'WITH', in an overlay declaration. The address of the start of the area to be overlayed is held in an Arithmetic Operand on the top of the stack, and is obtained using TOSADD, which deletes the operand. This address is held in OVERBASE and is inserted into the required location at the base of the stack, and the OVERLAY flag set.

Label OVRTEST:

This section is called following the symbol 'OVERFLOW' in a condition. FUNCTION is set to 24, the overflow test function; ACCUMULATOR being previously set, depending on the presence (0), or absence (1), of the symbol 'NO'.

Label PARAMTAB:

This section is called before, and after, the specification of the fields of a table parameter. It stores, and later restores, the current values of IDTYPE and DIRADD[0]. The information held in these (parameter accumulator, parameter address) is required for any subsequent parameters, and would be destroyed by the field specifications if not preserved. The variables IDTYPE2 and PARAMPTR2 are used for this purpose. (See EXCPSPEC)

Label PARTINT:

This section is called after the specification of a partword integer. It inserts the scale for integer (P=23) into the PVL field of IDTYPE.

Label PLUSSUB:

This section is called, in a subscript expression, after each Term preceded by a + symbol, and also after the first Term if unsigned. This section calls SUBTERM, which performs subscript arithmetic and optimisation.

Label PRESETSTRING:

This section is called after each quoted string in a Presetlist. The string is output, destined for the Special Constants area, by SPECSTRING, and the address of the string output, destined for the Data area and relocated, by OUTPRESETD.

As it is allowed to assign (the addresses of) strings to variables, the extension has been made to allow this assignment to be preset.
Label PROCENTRY:

This section is called when the entry point of a non-code procedure is reached.

The local procedure SCAN applies its procedure parameter PROC to each parameter identifier record in turn (if any), using the pointer P, which is non-local to the procedures.

The procedure LABPARAM is designed to be a parameter of SCAN. If the parameter record is that of a label parameter, an indirect Jump instruction, with the Data address of the parameter as its address, is output, and a label with the same identifier as the parameter, set to this instruction.

The other procedure, DUMPPARAM, is also designed to be a parameter of SCAN. This procedure outputs an instruction to store a parameter, of the procedure being compiled, in the Data location allocated for this parameter. The number of the accumulator is obtained from the parameter identifier record, as is the address. No action is taken if the accumulator number is zero, as in this case the identifier record will be that of a field of a table parameter.

PROCENTRY starts with a call of SCAN, with LABPARAM as a parameter. This outputs one instruction for each label parameter, and creates a label identifier record. This enables (conditional) jumps to be made to the label parameter in the normal manner, using the STATUS optimisations. It is expected that each label parameter will be used at least once in the procedure body, and most probably in a conditional jump. (Error exits etc.) A call is then made on ENTERPROC, which causes the loader to set the contents of the location containing the entry point to the current program address. An instruction is then output to store the link; this being followed by a call of SCAN with DUMPPARAM as a parameter, to output the instructions to store the parameters. If procedure trace is required, a call on L4, with the procedure identifier string as a parameter, is then output.

Label PROCTRACE:

This section sets 'BIT'[1]TRACE to the value of ACCUMULATOR. It is called in the case of a 'NO' 'PROCEDURE' 'TRACE' directive, in which case ACCUMULATOR will be zero; and also in the case of a 'PROCEDURE' 'TRACE' directive, in which case the value of ACCUMULATOR will be 1.

Label RAISE:

This section is called to apply the exponentiation operation (t) to its operands. If the right hand operand is not of type integer, the operation is carried out in floating point arithmetic by means of a call of FLOATOP with a parameter of 5. If the right hand operand is not the integer constant 2 or the left hand operand is of type floating, the operation is carried out by (fast) floating point arithmetic, by means of a call of FLOATOP with a parameter of 6.

The remaining case, squaring a fixed point number, is carried out by fixed point multiplication. The left hand Arithmetic Operand record is duplicated, replacing the right hand record (of the constant 2); if the operand exists only in an accumulator, an instruction is output to dump a copy of this operand into a temporary workspace. The section MPY is then entered.
Label RELADD:

This section is called where the operand of a jump instruction in code is a relative address, using 'SELF' or *. An Arithmetic Operand record is created by NOHAMS, and its commentary set to "(SELF)". The relative address is obtained by adding the current program address PTA, and the displacement held in NUMBER. This is performed using ADDADD, to avoid the possibility of overflow changing the tag field. The result is stored in the DIRADD field of the record.

Label RELATION:

This section is called after the second of the two expressions in a Comparison have been processed. It calculates the jump instruction required, which is stored in FUNCTION; and ensures that the result is in an accumulator, whose number is stored in ACCUMULATOR.

Where the test involves a single variable, its address, where possible, is used to update the accumulator record ACCS. If this variable is a partword, it will be unpacked without shifting where possible.

Label RHSBITS:

This section is called where a 'BITS' operation is used in an expression, after the expression to which the operation is to be applied has been processed. If the expression is an address, or is a partword field, or a constant; an instruction is output to copy it into an accumulator. This will provide redundant instructions, but will enable the operation to be carried out in an unambiguous manner. The partword specification and type information, which has been packed into BITSPEC by BITSIN is unpacked and inserted into the TYPEBITS and PARTWORD fields of the Arithmetic Operand. The actual bits operation being later carried out by PICK.

Label SCALETERM:

This section is called after each Term of an expression. If SCALEFIRM is set, the scale required for the expression has been defined, and the action of this section is confined to changing the type of the Terms to or from 'FLOATING' as required. If the required type is floating, a call is made on FLOATIT to ensure that the Term is of, or will be converted to, type floating. Otherwise, if the type of the Term is floating, a call to L21 is output, to convert the Term to the required fixed point scale.

On the other hand, if SCALEFIRM is not set, the scale of the expression is determined Term by Term. EXPSCALE is initially set to zero, and as each Term is processed, it is progressively modified according to the types and scales of the Terms. The type of the expression progressively goes through the sequence 'INTEGER' to 'FIXED' to 'FLOATING' according to the "strongest" type found. A comparison between the current value of EXPSCALE and the SCALE of all Terms except the first, is made by the use of SCALESTEST to switch to the appropriate label. In the case of the first Term, EXPSCALE is taken as the SCALE of the Term. Using the notation given in the description of SCALESTEST, the action taken for all Terms after the first is summarised below.
Label SCALETERM/

1 Io, Io & Io, Is & Is, Io

The scale is taken to be integer, with unspecified significance.

2 Is, Is

The scale is taken to be integer, with the larger number of significant bits.

3 & 4 Is, Fx & Fx, Is

The scale of the integer type is converted to the equivalent fixed scale (provided that the number of significant bits is known) and the scale determined as in section 5.

Note however, that when an integer is treated as a fixed, it is regarded as having an infinite number of (zero) fraction bits.

5 Fx, Fx

A new fixed scale is calculated, which has the larger number of integer bits, and the smaller number of fraction bits, of the two scales. If this would exceed the wordlength, fraction bits are dropped.

6 Fl, Any & Any, Fl

In this case the type is taken as floating.

Label SETANS:

This section is called after the occurrence of the 'ANSWER' symbol. It checks that an answer statement is allowed, and sets EXPSCALE and SCALEFIRM for the following expression. If the procedure is a 'VALUE' procedure, these are both cleared, thus allowing the type to be determined by the following expression, it being assumed that the context parameter is used to dynamically determine this. On the other hand, if the procedure is typed, the expression following the 'ANSWER' symbol is evaluated to the type of the procedure. In this case EXPSCALE is set to the type and scale of the procedure, and the flag SCALEFIRM is set. An answer statement is not allowed:

A. In an untyped procedure
B. In a 'CODE' procedure
C. Outside a procedure.

Note that this section starts with the declaration of DUMMYBLOCK, which makes this section a block, and allows the use of FAULT as a local label.

Label SETLAB:

This section is called after the occurrence of an identifier setting a label. STATUSCODE is called, to complete any deferred action associated with STATUS, to set it to zero, and to set any chain held by DUECHAIN. SETLABEL with LOOKUPLAB as a parameter, looks up and creates a label record if required, inserts the current program address into the record, and sets any chains of previous usage of the label.
Label SETLB:

This section is called after the integer constant giving the lower bound to a dimension of an array has been processed. The number is stored negatively in OFFSET.

Label SETLEVEL:

This section sets the value of the variable LEVEL, to the value of NUMBER. This section is called after a 'LEVEL' directive, to set the diagnostic level of the compiler output. At present, as implemented there are four diagnostic levels, 0 to 3, the default value being 1. The effects are as follows:

<table>
<thead>
<tr>
<th>Level</th>
<th>Diagnostics</th>
</tr>
</thead>
<tbody>
<tr>
<td>0</td>
<td>None</td>
</tr>
<tr>
<td>1</td>
<td>Page titles, labels, procedures and for statements printed at compile time.</td>
</tr>
<tr>
<td>2</td>
<td>As above, plus output of label, procedure, and for statement information to the loader.</td>
</tr>
<tr>
<td>3</td>
<td>As above, plus detailed commentary on each instruction output to the loader.</td>
</tr>
</tbody>
</table>

Label SETLIBSEG:

This section is called at the start of a Library procedure segment. The segment header is output by STARTSEG, and the housekeeping procedure, PROCSTACK, called. The DUECHAIN is set to the procedure address (containing the reference number), and marked for the convenience of the loader. The first Data location is allocated for the link, subsequent ones being used for the parameters and local workspace. TRACE is cleared, as procedures requiring trace should not be placed in the Library.

Label SETLIBVAR:

This section is called for the (trivial) compilation of Library variables. A check is first made to ensure that the preset value of this variable is greater than (in integer convention) \(-8372225 (-2^{23}+2^{14}-2^{0})\). This allows all floating point numbers, and most fixed and integers.

A tape is then produced giving the following information to the loader:

- The Library Number of the variable
- Its Preset value
- Its Identifier string
- A directive tag (4) and checksum.

This information is repeated on the printer.
Label SETNEZ:

This section is called where the comparator and second expression have been omitted in a comparison. In this case it is assumed that the comparison is against zero, with a zero value counting as "false", and a non-zero value counting as "true". This allows conditions such as:

'IF' 'BIT'[9]A 'THEN' . . .

RELOP is set to zero, which is the value it would have received if the comparator < > had been read.

Label SETNO:

This section sets the value of the variable ACCUMULATOR, to 0. It is called in cases where the symbol 'NO' is optional, and is present. Examples of such cases are:

'IF' 'NO' 'OVERFLOW' 'THEN' . .

'NO' 'PROCEDURE' 'TRACE';

Label SETPARAMS:

This section is called when it is discovered that a procedure body is not a code statement. This confirms the allocation of the data space required to hold the parameters and the link, and records the position on the stack of the identifier record of the last parameter (if any) in PARAMDECS.

Label SETPREF:

This section is called to select a suitable accumulator for the evaluation of bracketed expressions. An accumulator number is supplied by FINDPREF, and stored in PREFACC. If the preferred accumulator of the previous expression level is free, this will be chosen.

Label SETSHIFT:

This section is called after the occurrence of a shift operator in an expression, to prepare for the optimisation of the expression giving the number of shifts required. The required shift function is copied into BITSHIFT, and an anonymous Arithmetic Operand record created. The section SETUPSUB is then entered.

Label SETTEN:

This section copies the constant held in the variable NUMBER into TOPTEN. It is called where two integer constants may occur in close succession. The value of the first is held in TOPTEN until the second has been processed. This section is called in the case of packed address constants in 'SPECIAL' 'ARRAY's, and in the case of the base and displacement of 'EXTERNAL' specifications.

Label SETTEST:

This section sets the value of the variable TEST, to 1. It is called in cases where the symbol 'TEST' is optional, and is present. This allows test compilations to be performed. The effect of setting TEST is to inhibit the paper tape output on the punch. TEST is reset to zero at the end of each segment.
THE CORAL 66 COMPILER FOR FERRANTI ARGUS 500 COMPUTER (U)

UNCLASSIFIED

RSRE-TN-799

DRIC-BR-67199

NL
Label SETTEST/

'TEST' may also be used where a 'COMMON' communicator has been previously compiled, and it is required to set up the compiler for the compilation of a segment referring to it, without producing an output tape.

Label SETSUB:

This section is called after the integer constant giving the upper bound to a dimension of an array has been processed. The size of this dimension is calculated in NUMBER, and if less than 1, an error message is output, and NUMBER set to one.

Label SETUPSUB:

This section is called after the occurrence of the starting symbol of a subscript ([]), and is also entered from SETSHIFT.

A check is first made that the Arithmetic Operand on the top of the stack may be legally subscripted, if not an error message is output. A new expression level is created by a call of STACKEXPR, and this level set to indicate that an integer expression is mandatory. A suitable modifier is chosen as the preferred accumulator.

Label SETYES:

This section sets the value of the variable ACCUMULATOR, to 1. It is called in cases where the symbol 'NO' is optional, and has been omitted. Examples of such cases are:

'IF' ('NO') 'OVERFLOW' 'THEN' . .
('NO') 'PROCEDURE' 'TRACE';

This section is also called at the start of the declaration of an untyped procedure. Until the symbol following 'PROCEDURE' has been processed, it is not possible to determine whether a procedure declaration or a trace directive is to be processed. The call in these circumstances is harmless. As a consequence of this, the section TYPEPROC is also called.

Label SIMPLEASS:

This section is called where an assignment takes the form of:

Variable Becomes (±) Term

to optimise the cases of:

Variable Becomes (±) Variable.

If the variable on the right is the same as the variable on the left, and is unsigned or preceded by a + symbol, ASSFUN will have been set to function 12. In this case ASSFUN is cleared, and no instructions will be output. This case may occur as follows:

'FOR' A ± A 'WHILE' . . .

If the Term on the right is preceded by a - symbol, and is not the same as the variable on the left, and neither is of type floating; function 11 will be used for the assignment where appropriate, otherwise a call is made of UNARYMINUS, and function 10 used. This section is not called if assignment trace is required.
Label SKIPDTA:

This section is called following the Becomes symbol in a data declaration. Provided presetting is allowed, a directive is output, if required, to the loader, to skip the Data transfer address forwards, over unpreset areas of data.

Label SPECCON:

This section is called after the occurrence of an integer constant or identifier, preceded by a / symbol, in the presetlist of a 'SPECIAL' 'ARRAY'. An Arithmetic Operand record will have been created on the top of the stack, this is deleted and its direct address field obtained by TOSADD. This address is packed and output by OUT1014CS.

Label SPECLAB:

This section is called after the occurrence of an identifier, preceded by a : symbol, in the presetlist of a 'SPECIAL' 'ARRAY'. This identifier is treated as a label identifier, and its treatment differs in program and common segments. This section is also used to process entries in switch lists.

In a program segment, the label record is accessed using LOCKUPLAB, and this reference to the label chained up. It should be noted that the label cannot have been set at this block level before its use.

In a common segment the label may or may not have previously been explicitly specified, or used in a special array. A scan is made of the identifier records of the common segment. If a record does not already exist, a new one is created, as if the label were being explicitly specified. Otherwise, after checking that the record is that of a common label, this reference to the label is chained up to the previous reference.

In either case, the chain address of the label is output using OUT1014CS.

Label SPECNUM:

This section is called after the occurrence of a numeric constant in the presetlist of a 'SPECIAL' 'ARRAY'. The constant is rescaled to the scale held in IDTPE and output by OUT24CS. It should be noted that, although the scale of a 'SPECIAL' 'ARRAY' is integer, preset numeric constants may be stored to other scales if specified.

Label SPECONE:

This section is called after each identifier of a label, switch, or procedure, declared in a common segment. A location in special data, whose address will have been obtained using ADDRSPEC, is reserved, and set to an initial value of zero. The loader will subsequently insert the requisite address into this location.

Label SPECREL:

This section is called after the occurrence of a relative reference ( 'SELF' or * ) in the presetlist of a 'SPECIAL' 'ARRAY'. The displacement, held in NUMBER, is added to the current Special Data address STA, and output by OUT1014CS, which inserts an integer into the top ten bits if required.
Label STARTDEC:

This section is called before all declarations, to initialise the "declaration breeding ground" at the base of the main stack.

Label STARTFOR:

This section is called after the occurrence of the symbol 'FOR'. A call is made of STATUSCHECK to complete any deferred actions, and to ensure that the statement can be entered. A call is made of BEGLABBLOCK to ensure that the controlled statement cannot be entered by means of a jump to a label from a point outside. The values of the working variables used for the compilation of for statements:

FORCHAIN,FORSTATE,CV,COPYSKIP,DATAADD,BJA,RTA,RTL, and CCC

are stored by a call of ONSTACK, in case the for statement occurs within another. COPYSKIP is set to the current value of SKIPCHAIN, which is then cleared, as this is required for use in Conditions following 'WHILE' and also for the jump made when a for element is exhausted. The current value of DATASTART is recorded in DATAADD, so that it may be reset at the end of the for level. This value is also recorded in RTL.

FORSTATE, which will be used to determine the complexity of the list of for elements, and the type of each, is set to its initial value of zero.

Label STEPFUNT:

This section is called after the expressions following the 'STEP' and 'UNTIL' symbols in a 'FOR' statement have been processed. CCC is used to detect the case where all three expressions are constants, and is calculated by a logical and ('MASK') of the constant markers of the three expressions. The type of the expression is changed, if required, to the type of the control variable. If the expression is a constant, this is performed by means of a call of SCALENUM, otherwise a call is made of PICK. One is added (twice per Forelement) to FORSTATE, to indicate that the for element is of the form step-until. This section returns to the syntax analyser via FORSTORE, which stores the expression in an anonymous workspace if it has been evaluated.

Label STORE:

This section is called at the end of an assignment statement, to output the instructions for the assignment. This is performed by means of a call of STOREAWAY with the assignment trace flag, stored in TRACE, as the parameter.

Label STOREZERO:

This section is called when it is discovered that an assignment takes the form of:

Variable Becomes Zero.

If the variable is a Wordreference, a marker is set in ASSFUN to indicate that accumulator 0 is to be used with the assignment function (10). If the variable is a partword, then the optimisation will be carried out using FFLAG in STOREAWAY. This causes the field to be set to zero by masking the rest of the word. This section is not called if assignment trace is required.
Label STRINGEX:

This section is called after the occurrence of a quoted string as an expression, or in code, to create an Arithmetic Operand of type integer for the string. The string is output by SPECSTRING, and its address placed in the DIRADD of the record.

Label SUB:

This section is called to output the instructions for the subtraction of one Term from another. This is performed by means of a call of ADDSUB with function 03 as the parameter.

Label SUBA:

This section is called after the first Term of the expression on the right hand side of an assignment symbol; if the Term is preceded by a - symbol. If the Term has already been evaluated in an accumulator, a call is made of UNARYMINUS to output an instruction to negate it. Otherwise, if the Term is a variable, and is the same variable as the one on the left of the assignment symbol, and is not a partword or of type floating, ASSFUN is set to function 13, so that a "negate-store-and-add-to-store" assignment will be used. A mark is then set in ASSFUN, to indicate to SIMPLEASS or SELOPTA that the term was preceded by a minus sign. This section is not called if assignment trace is required.

Label SUBCOMMA:

This section is called after the occurrence of a comma in a subscript. A local pointer(LE) to the Arithmetic operand of the variable being subscripted, is calculated. If the indirect address of this variable is zero, this only occurring after the first comma where the previous expression consisted of an integer constant, the direct address is transferred to the indirect address of the record, and the direct address cleared. Otherwise, a suitable modifier is chosen, and an instruction is output, to copy the contents of the address so far calculated, into it. (This instruction may itself be modified.) The modifier number and accumulator marker (AMARK) are placed in the indirect address field of the Arithmetic Operand record, and the direct address field cleared. Finally a call is made on ACCUPDATE to update the accumulator record, ACCS.

Label TABADD:

This section is called before the fields of a table are specified. It stores the direct and indirect addresses of the start of the table in TABLED and TABLEI, for subsequent use by FIELDDISP. If the direct address of the table is null (Table parameters) it has a marker set to ensure that all table fields are treated in a consistent manner by LOOKUPD.

Label TABLESIZE:

This section is called after the occurrence of the Integer constant giving the number of locations required for a table. Provided that the table is not being overlayed, the number of locations is added to DATAMAX.
Label THENEX:

This section is called following the occurrence of the symbol 'THEN' in a conditional expression. The final test instruction is output by OUTCJ, and any chain resulting from the use of 'OR' set by SETDUE.

Label TYPEARRAY:

This section is called to set the AYM of IDTYPE, to denote that the following identifiers are identifiers of an array. This occurs in the case of arrays declared within a program or common segment, in the case of arrays specified in external and absolute communicators, and also for every table field.

Label TYPEFLOAT:

This section is called, following the symbol 'FLOATING', to insert the type number for floating point (2), into the TYP field of IDTYPE.

Label TYPEIARRAY:

This section is called in the case of array and table parameters. It sets the AYM and LCM bits of IDTYPE to denote that the location of an array or table is required to be passed as a parameter.

Label TYPEINT:

This section is called after the occurrence of the symbol 'INTEGER' in the specification of an integer of unspecified significance. It is also used in 'TABLE' and 'SPECIAL' 'ARRAY' declarations, both of which are treated as integer arrays.

This section inserts the scale for integer (P=23) into the PVL field of IDTYPE.

Label TYPEISWITCH:

This section is called following the symbol 'SWITCH' in common communicators, and in procedure declarations and specifications. It sets the TYP field of IDTYPE to the type number for labels (3), and sets the AYM and LCM flags. The LCM flag indicates that the switch is to be indirectly addressed.

Label TYPELAB:

This section is called, following the symbol 'LABEL' in declarations, to insert the type number for labels (3), into the TYP field of IDTYPE.

Label TYPELOC:

This section is called following the symbol 'LOCATION' in procedure specifications and declarations. It sets the marker LCM in IDTYPE, to denote that a location is required as the actual parameter. The marker LMZ is also set, this being used for non-standard parameters.
Label TYPEPROC:

This section is called, preceding the symbol 'PROCEDURE', to insert the type number for untyped procedures (4), into the TIP field of IDTYPE.

Label TYPESPEC:

This section is called in the case of non-standard (untyped) parameters, in a procedure declaration or specification. In this case the symbols 'VALUE' and 'LOCATION' are not followed by a Numbertype. The SPB field of IDTYPE is set to 1, to denote the first parameter of the pair (the second being the type/scale), and the PVL field is set to the default scale. (Integer)

Label TYPESWITCH:

This section is called following the symbol 'SWITCH' in switch declarations, and 'EXTERNAL' and 'ABSOLUTE' switch specifications. It inserts the type number for labels (3) into the TIP field, and sets the array marker AYM, of IDTYPE. Thus a switch is treated as an array of labels.

Label TYPETPROC:

This section is called, following a Numbertype, and preceding the symbol 'PROCEDURE', to convert the TIP field of IDTYPE, from a type number for a variable, to a type number for a typed procedure.

<table>
<thead>
<tr>
<th>Number Type</th>
<th>Variable type number</th>
<th>Procedure type number</th>
</tr>
</thead>
<tbody>
<tr>
<td>'INTEGER'</td>
<td>0</td>
<td>5</td>
</tr>
<tr>
<td>'FIXED'</td>
<td>1</td>
<td>6</td>
</tr>
<tr>
<td>'FLOATING'</td>
<td>2</td>
<td>7</td>
</tr>
</tbody>
</table>

Label TYPEVPROC:

This section is called between the symbols 'VALUE' and 'PROCEDURE' to insert the type information into the TIP (5) and SPB (3) fields of IDTYPE, for a variable-type procedure.

Label TYPEEXPR:

This section is called following a Numbertype preceding a bracketed expression. In this case the expression is required to be evaluated to the type and scale stated, even if the context demands a differing type or scale. (In this case the value will be rescaled to the type demanded by the context after the evaluation of the bracketed expression.) A new expression level is created by a call of STACKEXPR, and the EXPSCALE of this new level set to the required scale. SCALEFIRM is set to indicate that this scale is mandatory. The section SSTEPREF is then entered.
Label UNSFIELD:

This section is called to complete the partword descriptor BITSPEC, and to make appropriate adjustments to IDTYPE, in the case of 'UNSIGNED' table fields, and 'BITS' operations. A check is made to ensure that a 24 bit unsigned field is not being specified; and the UNS marker of BITSPEC is then set. One is added to the SGB field of IDTYPE, and if the scale is 'FIXED', one is added to the PVL (scale). This makes allowance for the sign bit, which although not present, is allowed for when unpacking and aligning the field.

Label VARCHCK:

This section is called to check that an identifier, used as a variable, is an arithmetic variable; i.e. is of type integer, fixed, or floating. If it is not, a warning message is output, and its type changed to integer to allow compilation to continue. This section returns to the syntax analyser via ATMCHECK to ensure that a subscript is not required, or has already been supplied.

Label WHILEL:

This section is called after the Condition following a 'WHILE' symbol in a 'FOR' statement has been processed. The final jump-if-false is output by a call of OUTCJ, and the destination of any jump-if-true instructions, set by means of a call of SETDUE. One is added to FORSTATE to indicate that the for element is of the while form.

Label ZEROARRAYS:

This section is called at the start of a set of arrays having common bounds. The location, ARRAYS, which keeps count of the number of arrays is cleared. TOPTEN is also cleared so that the top ten bits of any addressing vectors generated will be clear.

Label ZEROCOMP:

This section is called where the expression on the right of a comparator is zero. In this case two expressions do not require comparison. If the comparator is > or =, the left hand expression requires negation, as the instruction set of the machine does not contain the appropriate jump instructions.

Label ZERONUM:

This section clears the variable NUMBER. It is called in cases where a constant is optional, and has been omitted. In this case its value is assumed to be zero.
ARGUS 500 CORAL COMPILER

List of Errors

Page: ARITHMETIC
Procedure: FLOATOP
Line: 3
After: ACCUPDATE(LH);

Page: ASSIGNMENT AND FOR
Procedure: STOREAWAY
Line: 4
After: PARTWORD[RH] = 0
Insert: 'AND' PARTWORD[LH] = 0

Page: TABLES AND SPECIAL
Label: SPECLAB
Line: 3
Before: 'COTO' NEWNAME
Insert: OUT1014CS(ADD);

Page: CALLS AND CODE
Label: CALLPROC
Line: 6
Replace: PICK(LH, A, TYPEBITS[LH], 0)
By: PICK(LH, A, STB[LH], 0)
PAGE 1

\[ \text{EXTERNAL \{ 'INT', 'EXT1' \}} \]

( TYPES )

\[ \text{DEFINE 'INTEGER 'I';} \]
\[ \text{DEFINE 'FIXED 'J';} \]
\[ \text{DEFINE 'FLOATING 'K';} \]
\[ \text{DEFINE 'LABEL 'L';} \]
\[ \text{DEFINE 'PROCEDURE 'M';} \]
\[ \text{DEFINE 'INTROPROC 'N';} \]
\[ \text{DEFINE 'PROC' \{ 'FIMPROC = 0', 'FLOATPROC = 7' \}} \]

( CONSTANTS )

\[ \text{DEFINE 'MARK 'O00000000';} \]
\[ \text{DEFINE 'MARK 'O77777777';} \]

( FORMULAE )

\[ \text{DEFINE 'CYCLE\{AAA, BBB\} 'C\{AAA, ELC\{BBB\}\};} \]
\[ \text{DEFINE 'CHECKSUM\{AAA, BBB\} 'AAA\{AAA\} ELC\{BBB\} 'DIFFER' \{BBB\};} \]

( ARRAYS )

\[ \text{DEFINE 'INSTART 'LOC\{EXT1\};} \]
\[ \text{DEFINE 'STACKBEGIN 'LOC\{EXT1\};} \]
\[ \text{DEFINE 'STACKFINISH 'LOC\{EXT1\};} \]

PAGE 2

\[ \text{DEFINE 'STACK';} \]

\[ \text{COMPLEX2D} \]
\[ \text{DEFINE 'TABLE' STACK2D;} \]
\[ \text{CHAIN 'INT' 64 CHAIN OF STACKED ITEMS;} \]
\[ \text{STRING 'INT' 64 STRING FOR COMMENTARY;} \]
\[ \text{TYPEBITS 'INT' 64 DETAILS OF ITEM ON STACK;} \]
\[ \text{CON 'INT' 64 CONSTANT MARKER;} \]
\[ \text{ARY 'INT' 64 ARRAY MARKER;} \]
\[ \text{LCH 'INT' 64 LOCATION MARKER;} \]
\[ \text{PAC 'INT' 64 PARAMETER ACUMULATION;} \]
\[ \text{LVP 'INT' 64 LOCATION;} \]
\[ \text{LCP 'INT' 64 'LCP' 64 INDEX OF SCALE AND TYPE;} \]
\[ \text{DIFF 'INT' 64 'DIFF' 16 SIGNIFICANT BITS;} \]
\[ \text{PVL 'INT' 64 'PVL' 16 SCALE FACTOR;} \]
\[ \text{PARAMSPEC 'INT' 256 POINTER TO PROC PARAMETER SPEC;} \]
\[ \text{PARTWORD 'INT' 256 PART WORD SPECIFICATION;} \]
\[ \text{UNS 'INT' 256 MARKER FOR UNSIGNED FIELD;} \]
\[ \text{LES 'INT' 256 MARKER FOR LAZY FIELD;} \]
\[ \text{LES1 'INT' 128 MARKER FOR LAZY FIELD;} \]
\[ \text{LES2 'INT' 128 MARKER FOR LAZY FIELD;} \]
\[ \text{TSR 'INT' 256 'TSR' 256 MARKER FOR SHIFT TO ALIGN AT BOTTOM;} \]
\[ \text{TSR2 'INT' 256 'TSR2' 256 MARKER FOR SHIFT TO ALIGN AT TOP;} \]
\[ \text{STRADD 'INT' 64 DIRECT ADDRESS;} \]
\[ \text{ACC 'INT' 64 'ACC' 64 MARKER FOR DIRECT ADDRESS IS ACC;} \]
\[ \text{HDP 'INT' 64 'HDP' 64 'HDP' 64 MARKER FOR FIELD;} \]
\[ \text{PBY 'INT' 64 'PBY' 64 MARKER FOR FIELD;} \]
\[ \text{PAR 'INT' 64 'PAR' 64 'PAR' 64 MARKER FOR FIELD;} \]
\[ \text{PPL 'INT' 64 'PPL' 64 'PPL' 64 MARKER FOR FIELD;} \]
\[ \text{TRR 'INT' 64 'TRR' 64 'TRR' 64 MARKER FOR FIELD;} \]
\[ \text{STRADD 'INT' 64 INDIRECT ADDRESS;} \]
\[ \text{ACI 'INT' 64 'ACI' 64 MARKER FOR INDIRECT ADDRESS IS ACC;} \]
\[ \text{HCP 'INT' 64 'HCP' 64 'HCP' 64 MARKER FOR FIELD;} \]
\[ \text{TPR 'INT' 64 'TPR' 64 'TPR' 64 MARKER FOR FIELD;} \]
\[ \text{PBL 'INT' 64 'PBL' 64 'PBL' 64 MARKER FOR FIELD;} \]
\[ \text{TRR 'INT' 64 'TRR' 64 'TRR' 64 MARKER FOR FIELD;} \]
\[ \text{STRING 'INT' 64 IDENTIFIER USED AS STRING;} \]
\[ \text{STR 'INT' 64 'STR' 64 NUMBER OF WORDS USED BY STRING;} \]
\[ \text{CHN 'INT' 64 'CHN' 64 NUMBER OF CHAR IN STRING;} \]
\[ \text{COMMENT 'LABEL DECLARATION LISTS;} \]
\[ \text{RPA 'INT' 8 ( PROGRAM CHAIN;} \]
\[ \text{RCP 'INT' 8 ( DATA CHAIN;} \]
\[ \text{LEN 'INT' 8 ( LENGTH OF LABEL STRING (+1;} \]
\[ \text{MARK 'INT' 8 ( LABEL IDENTIFIER;} \]
\[ \text{O&M 'INT' 8 ( 'INT', 'DESC', 'DTYPE', 'SITYPE', 'ARRAY';} \]
\[ \text{OVERLAY 'STACK' WITH 'INT' 'NAME';} \]

\[ \text{OVERLAY 'STACK3' WITH 'INT' 'NAME';} \]
PAGE 4\1

ENDEELINED WORDS' (O.TIEIEN.ITEIs, 166066f FR60I6 O 10 IV1 61.

Q61,11' 11171[1)CL(4190 '0616' '00T0' CNIED1ZI'60T0' SPi

IF- LIVELds-O THIN '60656' NE6LINSEM"PA6Et: 'EN0'

ELSEV 11*0;

Clis506' TZ*READ 'WHILE' 12(68 ,64.64 '1O' 'IS' 11.63 '1666' OUTCHAE(TI)HSINECON;'60T0'

( TEST FOR 4X

TEST FOR 4 AND 4.1)

else READTI4(60

ENDIF T2-5.

READTZIA.*Z&SOOEO0j(0.62S1 6.4 II'ENS'

IF, 126.10

THEN' l50105 FTI'POR' 12.12:-Hil

INTERNUTE DSCALI6A021?LDXN66EI 'EN0',1.i 'GOTO'

CEXPONENT PADT ONLY

ADNU66ED*1i4bS6AC*0IAHSEESCALE*-10I'6O0l 60011

8 OCTAL NERBERS I

OCeNADTI:

00,915'

THEN' 'SOTO'

INTS

ELSE'

NEHDTI:N0NOCT:'I0'

124601) 'THEN'

'6010' 51

'ELSE'

GNOCTI

LIII'READT245

THEN' lOEGIN' NODS*TAIEEADT2

'END 'ELSE' NOOD*IINUEEHEIIS

Ti(6St

' THEN' '6010'

FYI

FOE

NODSa.NOOSs.1

1

1

tDE'

READ-$C)(EDE,6.EAI*S

-600

THEN'

'DOTO' IN1TY

'ELSE' '6010'

FI

ENISH

r41

I CHRACTEH

FASLYS I

P,

1

IYEA#('0COAHACTfil

FHULIHIl

c

SHIFT

OPENGTODS I

SOMEON 0051

TO

CLENS

1511

CRITO"Il

....

IS

Ti...

....

.\1

PAGES ARE BEST QUALITY PRACTICAL.
THIS PAGE IS BEST QUALITY FACSIMILE
FROM COPY PUBLISHED TO BOOK
IDENTIFIER OCCURRENCES IN COMPILER 20

STACK(1, STACK AND LABELS(5), TABLES AND SPECIAL(2, 8))

DO

READER 4(3), READ 4(2, 3)

DECLARE STACK(1), STACK AND LABELS(5), TABLES AND SPECIAL(2, 3), PROCEDURES AND ARRAYS(1, 9)

READER 4(3), READ 4(2, 3)

IDENTIFIER OCCURRENCES IN COMPILER 20

STACK(1), STACK AND LABELS(5), TABLES AND SPECIAL(2, 8)

DO

READER 4(3), READ 4(2, 3)

DECLARE STACK(1), STACK AND LABELS(5), TABLES AND SPECIAL(2, 3), PROCEDURES AND ARRAYS(1, 9)

READER 4(3), READ 4(2, 3)

FUNCTIONS

READER 4(3), READ 4(2, 3)

SUBROUTINE 4(3, 2)

STACK(1, STACK AND LABELS(5), TABLES AND SPECIAL(2, 8))

DO

READER 4(3), READ 4(2, 3)

DECLARE STACK(1), STACK AND LABELS(5), TABLES AND SPECIAL(2, 3), PROCEDURES AND ARRAYS(1, 9)

READER 4(3), READ 4(2, 3)
BASE RULES 2

BASE RULES 3
PAGE RULES 7

COLLATION(ITEMT, 02)

TEXT(SEC.CST)

SEC(ITEMT, BITS IN PR, FINEST)

SET(ITEMT, CLEARTYPE, DB, INT, DEDUCTIBLE, FIELD POSITION, UNSFIELD, CB)

DBDeductible=هجوم, DBIg, PARTANT, COMMA, INT

DB

PR(ITEMT)

NAME(ITEMT, CONSTANT)

DB, STACKED EXP, SETPREFIX, EXP-CR, OFFSET-EXP

TFR(ISITEMT, CONSTANT)

DB, LOOKUP(CA)

DB(ITEMT, TYPE, CLEARTYPE, TYPE, EXP-EXP, DB-CR, DB-EXP-EXP)

NAME, SINGLE SUBS

LOCATION, DB, LCVR, LOCATIONS, CR

LABEL, DB, DB T, RS, CR

TFR(SUBSCRIPT, VARCHECK)

ITEMT(ITEMT, ASSEMBLED)

DB(ITEMT, BITSHFT, DB(ITEMT, DBSHFT, 01))

DB(ITEMT, SEC, PLUS SUB)

DB, STACKED EXP, PLUS SUB

DB, SUBCR, CB

DB

NAME, TEXT, RS, 02

DB

UNION, COLLATION, OFF, 03

DB

DIV, UNITS, 66, 04

DB

FAULT, FACTOR, MS, 05

DIV, FACTOR, 61, 09

DB

DB(ITEMT, PLUS, TERM, MADD, 06)

NAME, TERM, SUB 02

DB

DB(ITEMT, GUIDE, 03)

DB

NAME, COMPLEMENT, COMPLEMENT

DB(ITEMT, PLUS, TERM, MADD, 06)

NAME, TERM, PLUS

DB(ITEMT, TERM, COMPLEMENT)

DB(ITEMT, PLUS, TERM, MADD, 06)

NAME, TERM, SUB 02

DB

DB(ITEMT, PLUS, TERM, MADD, 06)

NAME, TERM, SUB 02

DB

DB(ITEMT, GUIDE, 03)

DB

NAME, COMPLEMENT, COMPLEMENT

DB(ITEMT, PLUS, TERM, MADD, 06)

NAME, TERM, PLUS

DB(ITEMT, TERM, COMPLEMENT)
IDENTIFIER OCCURRENCES IN SYNTAX 13

WHILE: DECELS, TS 21(1), ASSIGNMENT(1), 2;
WITH: DECELS, TS 21(2), RULES 31(3), 3;
YESNO: DECELS, RULES 1(1), RULES 32(3), 4;
ZERO: DECELS, TS 1(1), RULES 4(4), RULES 7(7), ASSIGNMENT(1), 7;
EXCOND: LABELS, RULES 17(17), 2;
SYNUM: LABELS, RULES 21(21), RULES 22(22), 3;
INT: DECELS, RULES 5(5), RULES 5(5), EOB(5), 7;
Overall security classification of sheet: UNCLASSIFIED

(As far as possible this sheet should contain only unclassified information. If it is necessary to enter classified information, the box concerned must be marked to indicate the classification eg (R) (C) or (S)

<table>
<thead>
<tr>
<th></th>
<th></th>
<th></th>
<th></th>
</tr>
</thead>
<tbody>
<tr>
<td>TN 799</td>
<td></td>
<td></td>
<td>UNCLASSIFIED</td>
</tr>
</tbody>
</table>

5. Originator's Code (if known) 6. Originator (Corporate Author) Name and Location
RSRE (MALVERN)

5a. Sponsoring Agency's Code (if known) 6a. Sponsoring Agency (Contract Authority) Name and Location

7. Title The Coral 66 Compiler for the Ferranti Argus 500 Computer

7a. Title in Foreign Language (in the case of translations)

7b. Presented at (for conference papers) Title, place and date of conference

8. Author 1 Surname, initials 9(a) Author 2 9(b) Authors 3,4... 10. Date pp. ref.
GORMAN, B


15. Distribution statement

Descriptors (or keywords)
Computer Languages
Coral 66
Compiler

Abstract