



OAK RIDGE NATIONAL LABORATORY

operated by

UNION CARBIDE CORPORATION NUCLEAR DIVISION

for the

U.S. ATOMIC ENERGY COMMISSION



ORNL - TM - 1249

COPY NO. **140**

DATE - **NOV 21 1966**
RECEIVED
OFFICE OF GRANTS & RESEARCH CONTRACTS
8 56 AM '66

Neutron Physics Division

SIGNIFICANCE ARITHMETIC FOR FORTRAN

W. R. Burrus and B. W. Rust*

N67-12155

Abstract

The FORTRAN statement X # 193./71. - 2721./1001. has a true value of 0.0000281409..., but a single-precision FORTRAN computation (27 bit fraction) yields 0.28133392E-04 (using an E14.8 output FORMAT) due to loss of significance when subtracting. When the number of significant digits is questionable, the same problem may be run in both single and double precision. This is inconvenient because existing programs must be modified and because double precision requires twice as much storage space. We have implemented a simple alternative due to Max Goldstein for keeping track of the amount of significance in floating-point numbers by representing them in unnormalized form. Programs may be written in ordinary FORTRAN-IV, and the significance operations implemented by use of the ALTER feature of the loader. Some problems and applications of significance arithmetic are discussed.

Author

Computer Science Center, Union Carbide Nuclear Division, Oak Ridge.

GPO PRICE \$ _____
CFSTI PRICE(S) \$ _____
Hard copy (HC) *R. 00*
Microfiche (MF) *.50*
If 653 July 65

NOTE:
This Work Partially Supported by
NATIONAL AERONAUTICS AND SPACE ADMINISTRATION
Under Order R-104 *(01)*

NOTICE This document contains information of a preliminary nature and was prepared primarily for internal use at the Oak Ridge National Laboratory. It is subject to revision or correction and therefore does not represent a final report.

N67 12155
(ACCESSION NUMBER)
31
(PAGES)
800 95
(NASA CR OR TMX OR AD NUMBER)
(THRU) _____
(CODE) _____
(CATEGORY) *08*

LEGAL NOTICE

This report was prepared as an account of Government sponsored work. Neither the United States, nor the Commission, nor any person acting on behalf of the Commission:

- A. Makes any warranty or representation, expressed or implied, with respect to the accuracy, completeness, or usefulness of the information contained in this report, or that the use of any information, apparatus, method, or process disclosed in this report may not infringe privately owned rights; or
- B. Assumes any liabilities with respect to the use of, or for damages resulting from the use of any information, apparatus, method, or process disclosed in this report.

As used in the above, "person acting on behalf of the Commission" includes any employee or contractor of the Commission, or employee of such contractor, to the extent that such employee or contractor of the Commission, or employee of such contractor prepares, disseminates, or provides access to, any information pursuant to his employment or contract with the Commission, or his employment with such contractor.

INTRODUCTION

Meaningless digits come into a numerical calculation from four sources:

1. errors in the input data,
2. numbers going out of range (or spilling),
3. roundoff in the arithmetic operations,
4. loss of significance when subtracting numbers with similar value or adding numbers of similar magnitude and opposite sign.

The presence of meaningless digits in the final results is often not obvious from a superficial inspection, and many a spurious digit has found its way from a numerical calculation into the literature. Many schemes have been proposed to keep track of the number of significant digits in the results. (For explanation, we will refer to the number of accurate digits, even though the calculation is performed on a binary computer. It will be understood that one digit of accuracy implies about 3.4 bits of accuracy.) Some of these methods are mentioned briefly, but the scheme that was used with the IBM 7090/7094 FORTRAN-IV is considered in detail. A few results using a significance tracing arithmetic are given.

The simplest significance tracing scheme, in concept, is "interval arithmetic," in which each number is represented by a pair of numbers equal to the number's smallest and largest possible value. Every arithmetic operation or input/output operation transforms the lower and upper bound so that the resulting interval is sufficiently wide to include the true result. Interval arithmetic guarantees that all the digits which agree in the lower and upper bound are accurate. The primary disadvantage of interval arithmetic is that it requires two floating-point words to represent

one number, which necessitates more storage space; also, its use tends to give pessimistic results in a long computation, since for each arithmetic operation the assumption is made that the worst possible lower and upper bound will result.

Other schemes have been suggested to offset the pessimistic feature of interval arithmetic by considering tendencies for the accumulation of errors over many operations. Richtmeyer¹ suggests an arithmetic scheme in which some digits (or bits) explicitly carry the number of significant digits. Another type of arithmetic, discussed by Wadey,² computes the error at each step as if the errors of the operands were independently statistically distributed. This last method involves a square-root operation at each step.

The method that we selected to implement is based on an unnormalized floating-point number representation suggested by Ashenhurst and Metropolis³ in which numbers are carried as fractions that contain only significant digits preceded by leading zero digits. These fractions and a corresponding exponent are carried in a single IBM 7090/94 word.

Although significance arithmetics can handle all four types of error mentioned above, our emphasis is only on the loss of significance when two numbers of nearly equal size are subtracted or when two numbers of nearly equal magnitude and opposite sign are added. Our philosophy is to treat known errors in the input data explicitly in the program. Errors due to input truncation can similarly be taken into account by augmenting the input errors by a suitable amount. But we are still left with the not-so-easy-to-predict arithmetical errors which are inherent in a finite computer.

GOLDSTEIN'S RULES

Goldstein⁴ has described a set of rules for performing unnormalized arithmetic on the IBM 7090/94. These rules have been incorporated as a special modification into the NYU Computing Center machine. A special operating mode was provided in which the ordinary floating-point arithmetic instructions were interpreted by the hardware as unnormalized instructions. Standard floating-point representation was used, except that each number had just enough leading zeros in the fraction that all the remaining bits were alleged to be significant. Addition and subtraction simply omitted the usual normalization step. Multiplication and division gave the same number of leading zeros as the less accurate operand, after which a floating-point round operation took place.

In unnormalized arithmetic the result cited in the abstract is obtained as

$$\begin{array}{r} 0.27183098E 01 \\ -0.27182817E 01 \\ \hline 0.00000281E 01 \end{array}$$

whereas the correctly normalized results is $0.28140873E-04$. Zero must be treated as a special case. The ordinary 7090/94 zero is an all-zero word, but it is possible to obtain a significance zero as the result of subtracting two equal numbers:

$$\begin{array}{r} 0.34567890E-13 \\ -0.34567890E-13 \\ \hline 0.00000000E-13 \end{array}$$

This "significance zero" implies that the true result is less in magnitude than $0.00000001E-13$, or 10^{-21} . Multiplication or division which involves one or two "significance zero" operands must return an appropriate result. Thus $(0.00000000E-13)**2$ should have a value which implies that the magnitude is less than 10^{-42} .

The price that one pays for using an unnormalized arithmetic is that more significance is lost than when using comparable normalized arithmetic. However, Goldstein has demonstrated that in most cases the loss of accuracy is not significantly worse than the loss that occurs in ordinary 7090/94 FORTRAN arithmetic, which ignores the rounding operation. Compared with interval arithmetic, unnormalized arithmetic does not guarantee that a certain number of digits are correct, but the results will usually be correct to within a few units in the last significant digit.

FORTRAN IMPLEMENTATION

The IBM 7090/94 FORTRAN's have built-in provision for three types of floating-point arithmetic: REAL, DOUBLE PRECISION, and COMPLEX. The normal FORTRAN floating-point arithmetic is biased or truncated, since a floating-point round operation is not called for. It occurred to us that it would be convenient to have many different types of arithmetic available. In addition to the three above, one might wish the following:

ROUNDED single precision arithmetic
 RANDOM ROUNDED arithmetic (using pseudo-randomly
 generated bits) in order to check statistical
 models of error propagation
 INTERVAL arithmetic, etc.

as well as unnormalized SIGNIFICANCE arithmetic. It would be convenient if the arithmetic were implemented in a package of subroutines so that the type of arithmetic could be changed without modification to the source FORTRAN program.

An easy solution for single-word arithmetic is to hand-code FORTRAN FUNCTION routines ADD, SUB, FMP, and DIV to perform the elementary operations. But this is somewhat inconvenient for the programmer, since an

expression of the form

$$X = A*(B-C)/D$$

becomes

$$X = \text{DIV}(\text{MPY}(A, \text{SUB}(B, C)), D) .$$

It would be much more preferable to use the ordinary symbols and have them interpreted in terms of the desired arithmetic.

Some FORTRAN compilers, such as the CDC FORTRAN-63, provide a TYPE OTHER declaration, so that transfers to a special subroutine are automatically provided. One may achieve the same effect for double-word arithmetic in FORTRAN by declaring the variables to be TYPE COMPLEX and by supplying hand-coded routines which have the same name as the COMPLEX routines but which implement the desired arithmetic. An additional problem in the implementation of SIGNIFICANCE arithmetic is that some way has to be provided for the number of significant digits to be ascertained from the output.

It has turned out that single-word auxiliary arithmetic can be implemented very simply in FORTRAN-IV for the 7090/94. In order to use the ordinary arithmetic operations and have them correctly reinterpreted, one may redefine the floating-point operations FAD, FSB, FDP, and FMP by means of MACRO definitions. It is first necessary to obtain an ASSEMBLY language listing of the compiled program (or a PREST deck) and to reassemble with the MACRO definitions included. In our version, we obtain a PREST deck from the FORTRAN compiler (IBFTC, version 9) and utilize the ALTER feature of the input editor (IEDIT) to insert the necessary MACRO's.

In the MACRO definition, we included a switch which could be tested. If SWITCH .EQ. 0.0, then the ordinary FORTRAN arithmetic was performed,

but if SWITCH .NE. 0.0, a transfer was made to a special routine. With the switch off, the execution time of the ordinary arithmetic was increased by seven cycles per operation on the 7090. To obtain easy access to SWITCH by all programs, it was placed in labeled COMMON. Thus for ordinary arithmetic

```
COMMON /SCOM/ SWITCH
```

```
  .
  .
  .
```

```
SWITCH # 0.0
C # A-B
```

```
  .
  .
```

and for significance arithmetic

```
  .
  .
```

```
SWITCH # 1.0
C # A-B
```

```
  .
  .
```

One could use additional values of SWITCH, if desired, to switch back and forth between several different types of arithmetic.

The following form was used for the redefinition MACRO's so that they could be used in conjunction with indexing and indirect addressing:

Assume that the FORTRAN statement is $A \# B + C$:

ZFAD	OPSYN	FAD	SO THAT ORDINARY FAD CAN BE USED
FAD	MACRO	X,Y	
	NOB	X,Y	REPLACES ORIGINAL FAD IN THE TEXT
			SO THAT IF X WERE ** CORRECT ADDRESS
			WOULD HAVE BEEN STORED IN PROPER PLACE
	NZT	SWITCH	
	TRA	*+5	TRANSFER FOR ORDINARY FAD
	SXA	*+2,4	PREPARE FOR TRANSFER TO ROUTINE
	TSX	SGADD,4	
R	AXT	** ,4	RESTORE INDEX REGISTER
	TRA	*+2	
	ZFAD*	*-7	NORMAL FAD PERFORMED WHEN SWITCH IS OFF
	ENDM	FAD	

The SGADD routine had the following form:

	ENTRY	SGADD	
SGADD	STO	B	
	SXA	RTN,4	STORES RETURN INFORMATION
	CLA	-4,4	PICKS UP NOP X,Y
	STA	FETCH	STORES ADDRESS
	STT	FETCH	AND STORES TAG
	XEC	1,4	REMOTELY EXECUTES INSTRUCTION AT RZ
			TO RESTORE XR4 IN CASE Y # 4
FETCH	CLA	**	
	STO	C	
	.		PERFORMS REQUIRED SIGNIFICANCE
	.		ADDITION; LEAVES ANSWER IN ACC.
	.		
RTN	AXT	** , 4	
	TRA	1,4	
	B		
	C		
	END		

A MAP listing of the ALTER deck and of the SIGNIFICANCE routines is given in the appendix. These routines follow Goldstein's rules except that we did not bother with a correct round operation after division. True zero (denoted by 0) and "significance zero" (denoted by 0.0) were treated as special cases. Figure 1 shows the results of the basic arithmetical operations when either operand or both operands are zeros. We make no claim for efficiency, and would appreciate being informed of faster methods.

Fortuitously, the outputting of the number of significant digits presents no problem in 7090/94 FORTRAN (II or IV). The input/output routines make the BCD conversion so that the number of leading zeros in the fraction is preserved if an "E" type FORMAT is used. For example, an unnormalized 1.0 might print out as 0.00100000E 03 with an E14.8 FORMAT.

EXAMPLES

Some of the features of unnormalized significance arithmetic are nicely brought out by attempting to invert the notoriously poorly conditioned Hilbert matrices of various orders. Table 1 shows a 3 x 3 Hilbert

		$A = B + C$		
		$B \neq 0$	0.0	0
$C \neq 0$	$\neq 0$ or 0.0	$\neq 0$	$\neq 0$	C
0.0	$\neq 0$	0.0	0.0	C
0	B	B	0	0

		$A = B - C$		
		$B \neq 0$	0.0	0
$C \neq 0$	$\neq 0$ or 0.0	$\neq 0$	$\neq 0$	$-C$
0.0	$\neq 0$	0.0	0.0	$-C$
0	B	B	0	0

		$A = B * C$		
		$B \neq 0$	0.0	0
$C \neq 0$	$\neq 0$	$\neq 0$	0.0	0
0.0	0.0	0.0	0.0	0
0	0	0	0	0

		$A = B / C$		
		$B \neq 0$	0.0	0
$C \neq 0$	$\neq 0$	$\neq 0$	0.0	0
0.0	$0.0\#$	$0.0\#$	$0.0\#$	$0\#$
0	$0\#$	$0\#$	$0\#$	$0\#$

#DIVIDE CHECK ON

Fig. 1. The Results of Significance Arithmetic for Various Combinations of Zero and Nonzero Arguments. 0.0 denotes a significance zero and 0 denotes a true zero.

Table 1. Three-by-Three Hilbert Matrix Example

MATRIX A

C.C9999999E 01	0.50000000E 00	0.33333333E-00
C.50000000E 00	0.33333333E-00	0.25000000E-00
C.33333333E-00	0.25000000E-00	0.20000000E-00

INVERSE OF A

C.CC090000E 04	-0.00360000E 04	0.CC300000E 04
-C.CC360000E 04	0.01920000E 04	-0.CC180000E 05
0.CC300000E 04	-0.01800000E 04	0.C1800000E 04

PRODUCT

C.CC009999E 04	-0.00000000E 04	0.00000000E 04
C.CC000000E 03	0.00010000E 04	0.CC000000E 04
-C.CC000000E 03	-0.00000000E 04	0.CC010000E 04

matrix, its computed inverse (using a Gaussian elimination method), and the product of the computed inverse times the original matrix. Note that the elements of the inverse matrix have lost several digits of significance and that the off-diagonal elements of the product are not all zero as they should be. It is clear from the significance arithmetic, however, that the off-diagonal elements of the product have no significance to speak of. The 4×4 case is shown in Table 2. About half the significance has been lost in the computed inverse. Because the elements of the correct Hilbert inverse have simple floating-point representation, the spurious digits are easy to recognize. The diagonal elements of the product have lost nearly all significance, so that we should expect the worse for the 5×5 case. Table 3 shows that this is about as far as we can go in single precision on the 7090/94.

Table 4 shows some elementary FORTRAN statements and a comparison of the true value with the values obtained by ordinary arithmetic and significance arithmetic. The first result, R1, is the example given in the abstract. The second is of interest because ordinary arithmetic yields a small negative number instead of zero. R4 and R6 show how this false zero can propagate into other results. The third result is a true zero. R5 illustrates the effect of attempting to divide by true zero. The eighth result, R8, is contrived to show an extremely adverse case of significance arithmetic, where three significant digits are indicated but none are correct. In this case, "0.1" had a small conversion error when converted to binary, and combined with the roundoff, the sum picked up a spurious bit once in about every five summations. One can see from this example that significance arithmetic is more reliable if only a few numbers enter into the result. The last result, R9, shows how the error in R8 propagates.

Table 2. Four-by-Four Hilbert Matrix Example

MATRIX A

C.09999999E 01	0.50000000E 00	0.33333333E-0C	0.25000000E-00
C.50000000E 00	0.33333333E-00	0.25000000E-0C	0.20000000E-00
C.33333333E-00	0.25000000E-00	0.20000000E-00	0.16666666E-00
C.25000000E-00	0.20000000E-00	0.16666666E-00	C.14285714E-00

INVERSE OF A

0.00001600E 06	-0.00001200E 07	0.00002400E 07	-0.00001400E 07
-C.00012000E 06	0.00012000E 07	-0.0002700E 08	0.00016800E 07
C.00024000E 06	-0.00027000E 07	0.00064801E 07	-0.00042001E 07
-C.00014000E 06	0.00016800E 07	-0.00042001E 07	C.00028000E 07

PRODUCT

C.0000100E 06	0.00000000E 07	-0.00000000E 07	-0.00000000E 07
C.00000000E 06	0.0000010E 07	-0.00000000E 07	0.00000000E 07
-C.00000000E 05	0.00000000E 07	0.00000010E 07	-0.00000000E 07
C.00000001E 05	0.00000001E 06	-0.00000000E 07	C.00000010E 07

Table 3. Five-by-Five Hilbert Matrix Example

MATRIX A

C.09999999E 01	0.50000000E 00	0.33333333E-00	0.25000000E-00	0.20000000E-00
C.50000000E 00	0.33333333E-00	0.25000000E-00	0.20000000E-00	0.16666666E-00
C.33333333E-00	0.25000000E-00	0.20000000E-00	0.16666666E-00	0.14285714E-00
C.25000000E-00	0.20000000E-00	0.16666666E-00	0.14285714E-00	0.12500000E-00
C.20000000E-00	0.16666666E-00	0.14285714E-00	0.12500000E-00	0.11111110E-00

INVERSE CF A

C.00000025E 08	-0.00000003E 10	0.C0000009E 10	-C.C000013E 10	0.00000005E 10
-C.00000030E 09	0.00000048E 10	-0.C0000019E 11	0.C0000266E 10	-0.00000124E 10
C.00000104E 09	-0.00000189E 10	0.C0000079E 11	-0.C000175E 10	0.00000567E 10
-C.00001400E 08	0.00000268E 10	-0.C000176E 10	C.C0001792E 10	-0.00000881E 10
C.00000630E 08	-0.00000126E 10	0.C0000567E 10	-C.C000881E 10	0.00000440E 10

PRODUCT

C.00000001E 08	0.00000000E 10	-0.C0000000E 10	-0.C0000000E 10	-0.00000000E 10
C.00000000E 08	0.00000000E 10	-0.C0000000E 10	0.C0000000E 10	-0.00000000E 10
C.00000000E 08	0.00000000E 10	-0.C0000000E 10	C.C0000000E 10	-0.00000000E 10
C.00000000E 08	0.00000000E 09	-0.C0000000E 10	-0.C0000000E 10	-0.00000000E 10
C.00000000E 08	0.00000001E 09	0.C0000000E 10	-0.C0000000E 10	-0.00000000E 10

Table 4. Comparison of True Value with Computed Value Using Ordinary and Significance Arithmetic for Some Simple FORTRAN Expressions.

FORTRAN Statement	True Values	Computed Value	
		Ordinary Arithmetic	Significance Arithmetic
R1 # 193./71.-2721./1001.	0.28140873E-04	0.28133392E-04	0.00000281E 01
R2 # 0.1*10. - 1.0	0.	-0.75405806E-08	0.00000000E 00
R3 # 0.0	0.	0.	0.
R4 # 1.0/R2	∞	-0.13421772E 09	0.00000000E 17*
R5 # 1.0/R3	∞	0.*	0.*
R6 # 100. + R4	∞	-0.13421763E 09	0.00000000E 17
R7 # 1.0 + 100. - 100.	0.10000000E 01	0.09999999E 01	0.01000000E 02
DO 2 I # 1, 10000 2 R8 # R8 + 0.1	0.10000000E 04	.99997382E 03	0.10000094E 04
R9 # R8 - 1000.	0.	-0.26130676E-01	0.00000947E 03

*Divide check.

PRECAUTIONS

It is possible that system MACRO's such as an explicit or implied FLOAT will fail to operate properly because of the redefined floating-point operations. It is safer to switch the significance arithmetic off while performing a mode conversion, etc., as shown below:

```

      .
      .
      SWTCH # 0.0
      X # I
      SWTCH # 1.0
      .
      .
  
```

If SWTCH is not set and is undefined, the FORTRAN-IV loader will store an STR instruction in its location, and significance arithmetic will result.

It is conceivable that one might want to perform normal floating-point operations on unnormalized numbers and obtain normalized results. The present method of implementing significance arithmetic will not allow this because a floating-point multiplication of two unnormalized numbers may or may not give a normalized result and a floating point division of two unnormalized numbers will give the wrong answer if the divisor has more leading zeros than the dividend. For this reason care should be taken not to attempt to perform these normal floating-point operations on unnormalized numbers. One way of avoiding this difficulty would be to normalize the operands before carrying out the operations. Instead of testing the status of the switch in the MACRO, one could just transfer to the subroutine that carries out the arithmetic. Then in the arithmetic subroutine the switch could be checked. If significance arithmetic were desired, the operation could proceed in the same manner as it does now,

but if normal arithmetic were desired the operands could be normalized and the normal floating-point operation carried out. In such a scheme the floating divide MACRO might have the form

FDP	MACRO	X, Y	
	NOP	X, Y	REPLACES ORIGINAL FDP IN THE TEXT SO THAT IF X WERE ** CORRECT ADDRESS WOULD HAVE BEEN STORED IN PROPER PLACE
	SXA	*+2, 4	PREPARE FOR TRANSFER TO ROUTINE
	TSX	SGDVP, 4	
R	AXT	** , 4	RESTORE INDEX REGISTER

The subroutine itself might have the form

	ENTRY	SGDVP	
SGDVP	STO	DVND	
	SXA	RIN, 4	STORE RETURN INFORMATION
	CLA	-2, 4	PICKS UP NOP X, Y
	STA	FETCH	STORE ADDRESS
	STI	FETCH	AND TAG
	XEC	2, 4	REMOTELY EXECUTES INSTRUCTION AT R TO RESTORE XR4 IN CASE Y = 4
FETCH	CLA	**	
	STO	DVSR	
	CLA	SWTCH	CHECK THE SWITCH AND TRANSFER TO SG IF SIGNIFICANCE ARITHMETIC DESIRED
	TNZ	SG	
	CLA	DVSR	OTHERWISE,
	FAD	ZERO	NORMALIZE THE DIVISOR
	STO	DVSR	
	CLA	DVND	
	FAD	ZERO	THEN NORMALIZE THE DIVIDEND
	FDP	DVSR	AND CARRY OUT THE DIVISION
	TRA	RIN	
SG	.		PERFORM THE
	.		DIVISION IN
	.		SIGNIFICANCE ARITHMETIC
RIN	AXT	** , 4	
	TRA	2 4	
SCOM	CONTROL	SWTCH, DVND	LABELED COMMON BLOCK SCOM
SWTCH	BSS	1	SWTCH IN LABELED COMMON
DVND	HTR	0	
DVSR	HTR	0	
ZERO	DEC	0.0	

Note that such an implementation has a disadvantage in that it requires a transfer to a subroutine and the execution of several instructions to carry out ordinary arithmetic. This would increase the running time considerably

if the program used mostly ordinary arithmetic performed on numbers that would never have leading zeros. But it has an advantage in that it decreases the number of instructions in the MACRO itself and hence the number of instructions that would replace each floating-point operation in the main program. This might be advantageous for large programs which almost fill up the memory.

Since FORTRAN library routines return normalized results, the amount of significance will be lost if they are called with unnormalized arguments unless special precautions are taken. Goldstein⁴ and Ashenurst⁵ discuss the evaluation of FUNCTIONS which preserve significance. One possible scheme to determine the amount of significance in the function value corresponding to an unnormalized argument is to evaluate the function at both ends of the significance interval corresponding to the argument. For example, if it were required to find the SIN of the number $0.00012321E+4$, then the number $0.00012321E+4$ could be replaced by two numbers - $0.00012321E+4$ itself and the number obtained by adding a binary one to the last bit of it. The result would be the significance interval corresponding to $0.00012321E+4$. Then the SIN could be called for both of these values, getting a normalized result each time. The amount of significance that the answer should have would be the number of digits that are identical in the two results, and the correct answer could be obtained by denormalizing one of them until it had just this many significant digits.

CONCLUSIONS

Unnormalized floating-point significance arithmetic is easy to implement in FORTRAN-IV; however, some precautions must be exercised by the programmer. It is easy to extend the method used to obtain a variety

of different types of arithmetic, as well as significance arithmetic. We have found significance arithmetic to be extremely valuable when running a new problem for the first time or when running the problem to obtain final results to be published.

In addition, there are some instances where significance arithmetic can be used routinely to good advantage. A common problem is testing a floating-point number against zero to branch in an algorithm. Testing against zero is hazardous in ordinary arithmetic because of the accumulation of roundoff errors and the loss of significance. But significance arithmetic provides a simple solution by ignoring the bits of lowest significance, as in

```
IF (AND(CRIT,MASK) .EQ. 0.0) .... ,
```

with MASK defined by a DATA statement to be an octal constant which masks out the last few bits of the word.

ACKNOWLEDGMENT

We wish to express our appreciation to Jack Zeigler and Buford Carter of the Computer Sciences Center, Union Carbide Nuclear Division, Oak Ridge, for their advice and encouragement in trying to "beat the system."

REFERENCES

1. R. D. Richtmeyer, The Estimation of Significance, NYO-9083 (1960).
2. W. G. Wadey, J. Assoc. Computing Mach. 7, 129 (1960).
3. R. L. Ashenurst and N. Metropolis, J. Assoc. Computing Mach. 6, 415 (1959).
4. M. Goldstein, Commun. Assoc. Computing Mach. 6, 111 (1963).
5. R. L. Ashenurst, J. Assoc. Computing Mach. 11, 168 (1964).

APPENDIX

Listing of ALTER DECK

and

SIGNIFICANCE Routines

LISTING OF ALTER DECK.

	*ALTER	I
	NOCRS	
MAKE	MACRO	A, B
ZZA	OPSYN	A
A	MACRO	X, Y
	PMC	ON
	NOP	X, Y
	NZT	SWTCH
	TRA	++5
	SXA	++2, 4
	TSX	B, 4
	AXT	+, 4
	TRA	++2
	ZZA*	-7
	PMC	OFF
	ENDM	A
	ENDM	MAKE
MAKE	FAD, SGADD	
MAKE	FSB, SGSUB	
MAKE	FMP, SGMPY	
MAKE	FDP, SGDVP	
	*ENDAL	

SIQADD
ASSEMBLED TEXT.

				STEXT	SIQADD			
						ENTRY	SGADD	
BINARY CARD ID. SIQA3002								
00003	3631	03	0	00345	10001	SGADD	STO	B
00031	3634	03	4	00343	10001		SXA	RTN,4
00032	3530	03	4	77774	10000		CLA	-4,4
00033	3621	03	0	00006	10001		STA	FETCH
00034	3625	03	0	00006	10001		STT	FETCH
00035	3522	03	4	00301	10000		XEC	1,4
00036	3530	03	0	00300	10000	FETCH	CLA	**
00037	3631	03	0	00346	10001		STO	C
00013	3530	03	0	00345	10001		CLA	B
00011	3130	03	0	00342	10001		TZE	USEC
00012	4320	03	0	00347	10001		AVA	MASKI
00013	3130	03	0	00323	10001		TZE	TESTC
00014	3530	03	0	00345	10001		CLA	B
00015	4330	03	0	30346	10001		UFA	C
00016	3760	03	0	00311	10000		F3N	
00017	4100	03	0	00043	10001		TNZ	RTN
00023	3530	03	0	00346	10001		CLA	C
00021	4320	03	0	00350	10001		AVA	#077700000000
00022	3020	03	0	00343	10001		TRA	RTN
BINARY CARD ID. SIQA3003								
00023	3530	03	0	00346	10001	TESTC	CLA	C
00024	3130	03	0	00340	10001		TZE	USEB
00025	4320	03	0	00347	10001		AVA	MASKI
00026	3130	03	0	00035	10001		TZE	BIGXP
00027	3530	03	0	00346	10001		CLA	C
00033	4330	03	0	00345	10001		UFA	B
00031	3760	03	0	00311	10000		F3N	
00032	4130	03	0	00043	10001		TNZ	RTN
00033	3530	03	0	00345	10001		CLA	B
00034	3020	03	0	00343	10001		TRA	RTN
00035	4530	03	0	00345	10001	BIGXP	CAL	B
00036	4430	03	0	00046	10001		SBM	C
00037	4120	03	0	00342	10001		TMI	USEC
00043	3530	03	0	00345	10001	USEB	CLA	B
00041	3020	03	0	00343	10001		TRA	RTN
00042	3530	03	0	00346	10001	USEC	CLA	C
00043	3774	03	4	00000	10000	RTN	AXT	**4
00044	3020	03	4	00301	10000		TRA	1,4
00045	3030	03	0	30300	10000	B	HTR	0
BINARY CARD ID. SIQA3004								
00046	3033	03	0	00300	10000	C	HTR	0
00047	300777777777				10000	MASKI	OCT	000777777777
00053	777030300000				10000		*LORG	
				00300	01111		END	

SIGSUB
ASSEMBLED TEXT.

		\$TEXT		SIGSUB				ENTRY	SGSUB
BINARY CARD ID. SIGS3032									
00003	3631	03	0	00346	10001	SGSUB	STO	B	
00031	3634	03	4	00344	10001		SXA	RTN,4	
00032	3530	03	4	77774	10000		CLA	-4,4	
00033	3621	03	3	00306	10001		STA	FETCH	
00034	3625	03	0	00006	10001		STT	FETCH	
00035	3522	03	4	00301	10000		XEC	1,4	
00036	3530	03	0	00300	10000	FETCH	CLA	**	
00037	3730	03	0	00302	10000		CHS		
00013	3631	03	0	00347	10001		STO	C	
00011	3530	03	0	00346	10001		CLA	B	
00012	3130	03	0	00343	10001		TZE	USEC	
00013	4320	03	0	00350	10001		ANA	MASKI	
00014	3130	03	0	00324	10001		TZE	TESTC	
00015	3530	03	0	00346	10001		CLA	B	
00016	4330	03	0	00347	10001		UFA	C	
00017	3730	03	0	00311	10000		FRN		
00023	4130	03	0	00344	10001		TNZ	RTN	
00021	3530	03	0	00347	10001		CLA	C	
00022	4320	03	0	00351	10001		ANA	#077700000000	
BINARY CARD ID. SIGS3033									
00023	3020	03	0	00344	10001		TRA	RTN	
00024	3530	03	0	00347	10001	TESTC	CLA	C	
00025	3130	03	0	00341	10001		TZE	USEB	
00026	4320	03	0	00350	10001		ANA	MASKI	
00027	3130	03	0	00336	10001		TZE	BIGXP	
00033	3530	03	0	00347	10001		CLA	C	
00031	4330	03	0	00346	10001		UFA	B	
00032	3730	03	0	00311	10000		FRN		
00033	4130	03	0	00344	10001		TNZ	RTN	
00034	3530	03	0	00346	10001		CLA	B	
00035	3020	03	3	00344	10001		TRA	RTN	
00036	4530	03	3	00346	10001	BIGXP	CAL	B	
00037	4430	03	0	00347	10001		SBM	C	
00043	4120	03	3	00343	10001		TMI	USEC	
00041	3530	03	0	00346	10001	USEB	CLA	B	
00042	3020	03	3	00344	10001		TRA	RTN	
00043	3530	03	0	00347	10001	USEC	CLA	C	
00044	3774	03	4	00300	10000	RTN	AXT	** ,4	
00045	3020	03	4	00301	10000		TRA	1,4	
BINARY CARD ID. SIGS3034									
00046	3030	03	0	00000	10000	B	HTR	0	
00047	3030	03	0	00300	10000	C	HTR	0	
00053	303777777777				10000	MASKI	OCT	000777777777	
30051	777000303030				10000		*LORG		
				00000	01111		END		

SIGMPY
ASSEMBLED TEXT.

		\$TEXT		SIGMPY		ENTRY	SGMPY
BINARY CARD ID. SIGMJOB2							
00000	3634	03	4	00362	10001	SGMPY	SXA
00001	3530	03	4	77774	10000		CLM
00002	3621	03	0	00005	10001		STA
00003	3625	03	0	00005	10001		STT
00004	3522	03	4	00001	10000		XEC
00005	3530	03	0	00000	10000	INST	CLA
00006	3631	03	0	00364	10001		STO
00007	4630	03	0	00365	10001		STQ
00008	3630	03	0	00370	10001		STZ
00009	3530	03	0	00365	10001		CLA
00010	3130	03	0	00335	10001		TZE
00011	4320	03	0	00371	10001		AVA
00012	4130	03	0	00322	10001		TNZ
00013	4530	03	0	00365	10001		CAL
00014	3432	03	0	00374	10001		SUB
00015	3430	03	0	00373	10001		ADD
00016	3632	03	0	00365	10001		SLW
00017	3631	03	0	00370	10001		STO
00018	3530	03	0	00364	10001	TESTY	CLA
BINARY CARD ID. SIGMJOB3							
00019	3130	03	0	00334	10001		TZE
00020	4320	03	0	00371	10001		AVA
00021	4130	03	0	00335	10001		TNZ
00022	4530	03	0	00364	10001		CAL
00023	3432	03	0	00374	10001		SUB
00024	3430	03	0	00373	10001		ADD
00025	3632	03	0	00364	10001		SLW
00026	3631	03	0	00370	10001		STO
00027	3020	03	0	00335	10001		TRA
00028	3630	03	0	00370	10001		STZ
00029	3074	03	4	03000	10011	BLAP	TSX
00030	3030	03	0	00365	10001		PZE
00031	3030	03	0	00367	10001		PZE
00032	3074	03	4	03000	10011		TSX
00033	3030	03	0	00364	10001		PZE
00034	3030	03	0	00366	10001		PZE
00035	3560	03	0	00365	10001		LDQ
00036	3260	03	0	00364	10001		FMP
00037	3760	03	0	00311	10000		FRN
BINARY CARD ID. SIGMJOB4							
00038	3631	03	0	00364	10001		STO
00039	3074	03	4	04000	10011		TSX
00040	3030	03	0	00364	10001		PZE
00041	3030	03	0	00367	10001		PZE
00042	3030	03	0	00366	10001		PZE
00043	3530	03	0	00370	10001		CLA
00044	3130	03	0	00361	10001		TZE
00045	4530	03	0	00364	10001		CAL
00046	4320	03	0	00372	10001		AVA

SIGMPY
ASSEMBLED TEXT.

00057	1430 03 0	00074	10001		ADD	ONEX
00060	1632 03 0	00064	10001		SLW	YSTO
00061	1530 03 0	00064	10001	AVS	CLA	YSTO
00062	1774 03 4	00000	10000	RTN	AXT	00,4
00063	1020 03 4	00001	10000		TRA	1,4
00064	1110 03 0	00000	10000	YSTO	HTR	0
00065	1110 03 0	00000	10000	QSTO	HTR	0
00066	1110 03 0	00000	10000	YNO	HTR	0
00067	1110 03 0	00000	10000	QNO	HTR	0
00073	1030 03 0	00000	10000	FLAG	HTR	0

BINARY CARD ID. SIGM0005

00071	000777777777	10000	MASK1	OCT	000777777777
00072	777000000000	10000	MASK2	OCT	777000000000
00073	000000000001	10000	ONE	OCT	000000000001
00074	001000000000	10000	ONEX	OCT	001000000000
	00300 01111			END	

SIGDVP
ASSEMBLED TEXT.

		\$TEXT		SIGDVP				
						ENTRY	SGDVP	
BINARY CARD ID. SIGD002								
00000	0631	03	0	00063	10001	SGDVP	STO	DVND
00001	0634	03	4	00061	10001		SXA	RTN,4
00002	0500	03	4	77774	10000		CLA	-4,4
00003	0621	03	0	00006	10001		STA	INST
00004	0625	03	0	00006	10001		STT	INST
00005	0522	03	4	00001	10000		XEC	1,4
00006	0500	03	0	00000	10000	INST	CLA	**
00007	0631	03	0	00064	10001		STO	DVSR
00008	0600	03	0	00067	10001		STZ	FLAG
00009	0500	03	0	00063	10001		CLA	DVND
00010	0100	03	0	00035	10001		TZE	BLAP
00011	4320	03	0	00070	10001		ANA	MASKI
00012	4100	03	0	00022	10001		TVZ	TESTY
00013	4500	03	0	00063	10001		CAL	DVND
00014	0402	03	0	00073	10001		SUB	ONEX
00015	0400	03	0	00072	10001		ADD	ONE
00016	0602	03	0	00063	10001		SLW	DVND
00017	0601	03	0	00067	10001		STO	FLAG
00018	0500	03	0	00064	10001	TESTY	CLA	DVSR
BINARY CARD ID. SIGD003								
00019	0100	03	0	00034	10001		TZE	BLAP-1
00020	4320	03	0	00070	10001		ANA	MASKI
00021	4100	03	0	00035	10001		TVZ	BLAP
00022	4500	03	0	00064	10001		CAL	DVSR
00023	0402	03	0	00073	10001		SUB	ONEX
00024	0400	03	0	00072	10001		ADD	ONE
00025	0602	03	0	00064	10001		SLW	DVSR
00026	0601	03	0	00067	10001		STO	FLAG
00027	0020	03	0	00035	10001		TRA	BLAP
00028	0600	03	0	00067	10001		STZ	FLAG
00029	0074	03	4	03000	10011	BLAP	TSX	NRMCT,4
00030	0000	03	0	00063	10001		PZE	DVND
00031	0000	03	0	00065	10001		PZE	DVNNO
00032	0074	03	4	03000	10011		TSX	NRMCT,4
00033	0000	03	0	00064	10001		PZE	DVSR
00034	0000	03	0	00066	10001		PZE	DVSNO
00035	0500	03	0	00063	10001		CLA	DVND
00036	0241	03	0	00064	10001		FDP	DVSR
00037	4600	03	0	00064	10001		STQ	DVSR
BINARY CARD ID. SIGD004								
00038	0074	03	4	04000	10011		TSX	DNORM,4
00039	0000	03	0	00064	10001		PZE	DVSR
00040	0000	03	0	00065	10001		PZE	DVNNO
00041	0000	03	0	00066	10001		PZE	DVSNO
00042	0500	03	0	00067	10001		CLA	FLAG
00043	0100	03	0	00060	10001		TZE	ANS
00044	4500	03	0	00064	10001		CAL	DVSR
00045	4320	03	0	00071	10001		ANA	MASK2
00046	0400	03	0	00073	10001		ADD	ONEX

SIGDVP
ASSEMBLED TEXT.

00057	3632 03 0 30364	10001		SLW	DVSR
00063	3550 03 0 30364	10001	ANS	LDQ	DVSR
00061	3774 03 4 30300	10000	RTN	AXT	**4
00062	3020 03 4 00001	10000		TRA	1,4
00063	3030 03 0 00000	10000	DVND	HTR	C
00064	3030 03 0 00000	10000	DVSR	HTR	0
00065	3030 03 0 00000	10000	DVNVO	HTR	0
00066	3030 03 0 00000	10000	DVSVO	HTR	0
00067	3030 03 0 00000	10000	FLAG	HTR	0
00073	303777777777	10000	MASK1	OCT	000777777777

BINARY CARD ID. SIGD3035

00071	777000000000	10000	MASK2	OCT	777000000000
00072	000000000001	10000	ONE	OCT	000000000001
00073	001000000000	10000	OVEX	OCT	001000000000
	30300 01111			END	

DNORML
ASSEMBLED TEXT.

				\$TEXT	DNORML			
						ENTRY	DNORM	
BINARY CARD ID. DNORJ002								
00000	0500	00	4	00002	10000	DNORM	CLA*	2,4
00001	0550	00	4	00003	10000		LDQ*	3,4
00002	0640	00	0	00002	10011		TLQ	*+2
00003	0131	00	0	00000	10000		XCA	
00004	0771	00	0	00033	10000		ARS	27
00005	0621	00	0	00015	10001		STA	BACK
00006	0601	00	0	00023	10001		STO	NCT
00007	0500	00	0	00024	10001		CLA	#27
00010	0402	00	0	00023	10001		SUB	NCT
00011	0621	00	0	00016	10001		STA	BAKMO
00012	0500	00	4	00001	10000		CLA*	1,4
00013	4765	00	0	00033	10000		LGR	27
00014	0351	00	0	00023	10001		ACL	NCT
00015	0757	00	0	00000	10000	BACK	ALS	**
00016	4763	00	0	00000	10000	BAKMO	LGL	**
00017	4773	00	0	00033	10000		R2L	27
00020	0750	00	0	00011	10000		FRN	
00021	0601	00	4	00001	10000		STO*	1,4
00022	0020	00	4	00004	10000		TRA	4,4
BINARY CARD ID. DNORJ003								
00023	0000	00	0	00000	10000	NCT	HTR	0
00024	0000000000	00	33	00000	01111		*LORG	
				00000	01111		END	

NRMCT
ASSEMBLED TEXT.

		\$TEXT		NRMCT		ENTRY	NRMCT
BINARY CARD ID. NRMCT002							
00000	0500	00	4	00001	10000	NRMCT CLA*	1,4
00001	0601	00	0	00015	10001	STO	DSTO
00002	0300	00	0	00017	10001	FAD	ZRO
00003	0601	00	0	00016	10001	STO	DNSTO
00004	0601	00	4	00001	10000	STO*	1,4
00005	0500	00	0	00016	10001	CAL	DNSTO
00006	0320	00	0	00020	10001	AVA	MASK
00007	0602	00	0	00016	10001	SLW	DNSTO
00010	0500	00	0	00015	10001	CAL	DSTO
00011	0320	00	0	00020	10001	AVA	MASK
00012	0602	00	0	00016	10001	SUB	DNSTO
00013	0602	00	4	00002	10000	SLW*	2,4
00014	0320	00	4	00003	10000	TRA	3,4
00015	0000	00	0	00000	10000	DSTO	HTR
00016	0000	00	0	00000	10000	DNSTO	HTR
00017	000000000000			10000		ZRO	DEC
00020	377000000000			10000		MASK	OCT
				00000	01111		END

Internal Distribution

- | | | | |
|--------|-------------------|---------|-------------------------------|
| 1-3. | L. S. Abbott | 34. | W. Zobel |
| 4-13. | W. R. Burrus | 35. | G. Dessauer (Consultant) |
| 14. | V. R. Cain | 36. | B. C. Diven (Consultant) |
| 15. | G. T. Chapman | 37. | M. L. Goldberger (Consultant) |
| 16. | C. E. Clifford | 38. | M. H. Kalos (Consultant) |
| 17. | D. C. Irving | 39. | L. V. Spencer (Consultant) |
| 18. | F. C. Maienschein | 40-41. | Central Research Library |
| 19. | R. W. Peelle | 42. | Document Reference Section |
| 20-30. | B. W. Rust | 43-223. | Laboratory Records Department |
| 31. | R. T. Santoro | 224. | Laboratory Records ORNL R.C. |
| 32. | D. K. Trubey | 225. | ORNL Patent Office |
| 33. | J. W. Wachter | | |

External Distribution

- 226-240. Division of Technical Information Extension (DTIE)
241. Division of Research and Development (ORO)