D. N. Arden

S&EC TRAINING COURSE

Number Systems

Whirlwind I handles information in blocks called "words" which consist of 16 "bits". A "bit" is essentially the result of a choice between two alternatives, one of which must be true, but not both. For example, a yes or no answer to the question "Has Mr. G. smoked Camels in the past week?" is a "bit" of information. It is conceivable and even practical to use digital computers to work with data or information in this form. One important example of this type of application of digital computers is the design of switching circuits, where each bit gives information as to whether a particular relay was open or closed. Clearly, the number of "states" or "patterns" that can be represented by one bit is $2^1=2$ and the number of states representable by 16 bits is $2^{16} = 64,536$.

However, one of the most common uses of "words" or parts of words in Whirlwind is to represent numbers. There are many ways in which this could be accomplished and in fact there are several methods in common use by digital computers. All of these methods, and in fact all methods of storing information in words, depend upon a "positional" system. This means that the significance of a bit depends on its position from left to right in a given order. The bits in Whirlwind words are conventionally numbered from left to right with the integers 1, 1,...,15.

0	1	2	3	4	5	6	7	8	9	10	11	12	13	14	15

A very common example of a positional system is the decimal representation of numbers where the significance of each digit depends on its position. For example, in the number

.125

the power of 10 by which a given digit is understood to be multiplied is determined by the position of the digit. The first digit to the right of the decimal place is multiplied by $1/10 = 10^{-1}$, the second by $1/100 = 1/10^2 = 10^{-2}$, etc. That is .125 is shorthand for

$$1 \times 10^{-1} + 2 \times 10^{-2} + 5 \times 10^{-3}$$

Similarly 53 is shorthand for

$$+5 \times 10^1 + 3 \times 10^0$$

where $10^0 = 1$. The number 10 is called the "base" or "radix" of the representation. It is quite possible to use representations with other bases than 10, as long as it is clearly understood what base is to be used. If a base r were used instead of 10, where r could be any positive integer > 1, then only the digits 0, 1, ..., r-1 can occur in any position if the representation is to be useful as otherwise there would be too many ways to represent the same number. In fact, for representations with base r greater than 10, it is almost necessary to invent new single characters to replace 11, 12, ...r-1, or to agree that each position will be represented by two or more digits.

Since information in most digital computers is stored in the form of bits, the number 2 is a very convenient base for the representation of numbers, since in this case only 0 and 1 can occur in any position, and hence each position contains exactly one bit of information. Representation to the base two is called binary representation, and the binary representation of a number is sometimes loosely called a "binary number". As an example, the binary representation of 53.125 (decimal) is

$$110101.001 = 1(2^5) + 1(2^4) + 0(2^3) + 1(2^2) + 0(2^1)$$
$$+1(2^0) + 0(2^{-1}) + 0(2^{-2}) + 0(2^{-3})$$

Since such representations are frequently used by digital computers, it usually becomes necessary for programmers to change the representation of a number with a given base to a representation with a new base.

It will suffice to be able to perform this conversion for positive numbers only as a minus sign can be neglected for the duration of conversion and affixed to the final result. For this reason, only the conversion of positive numbers will be considered in the following. As a step in this direction, one observes that if either the new base or the old base is a power of the other, the problem is very easily solved. For example, if the old base is 2, the representation with this base is 110101.001 and the new base is $8 = 2^3$, one proceeds by beginning at the decimal point and factoring out the largest possible power of 8.

110101.001 with base 2

$$= 1(2^5) + 1(2^4) + 0(2^3) + 1(2^2) + 0(2^1) + 1(2^0) + 0(2^{-1}) + 0(2^{-2}) + 1(2^{-3})$$

$$= 1(2^2)(2^3) + 1(2)(2^3) + 0(2^3) + 1(2^2) + 0(2^1) + 1(2^0) + 0(2^2)(2^{-3}) + 0(2^1)(2^{-3}) + 1(2^{-3})$$

$$= \quad 1(2^2) + 1(2) + 0 \quad (2^3) + \quad 1(2^2) + 0(2^1) + 1(2^0) + \quad 0(2^2) + 0(2^1) + 1 \quad (2^{-3})$$

$$= \quad 4 + 2 + 0 \quad 8^1 + \quad 4 + 0 + 1 \quad 8^0 + \quad 0 + 0 + 1 \quad 8^{-1}$$

$$= 6(8^1) + 5(8^0) + 1(8^{-1})$$

= 65.1 with base 8

Conversely, to get the binary representation by using the binary representation of each digit, since 65.1 with the base 8

$$= 6(8^1) + 5(8^0) + 1(8^{-1})$$

$$= 6(2^3) + 5(2^0) + 1(2^{-3})$$

$$= \quad 1(2^2) + 1(2^1) + 0(2^0) \quad (2^3) + \quad 1(2^2) + 0(2^1) + 1(2^0) \quad (2^0) + \quad 0(2^2) + 0(2^1) + 1(2^0) \quad (2^{-3})$$

$$= 1(2^5) + 1(2^4) + 0(2^3) + 1(2^2) + 0(2^1) + 1(2^0) + 0(2^{-1}) + 0(2^{-2}) + 1(2^{-3})$$

= 110101.001 with base 2.

A useful fact in the problem of conversion is that regardless of the base of a representation, the digits to the left of the decimal point are multiplied by positive or zero powers of the base and the digits to the right are multiplied by negative powers of the base. In other words, the decimal point separates the representation into the representation of an integer or a whole number and the representation of a proper fraction. It is easy to see that if two numbers written as the sum of an integer and a proper fraction are equal, then the integers must be equal and the fractions must be equal

($a+b = c+d$ a, c, integers ≥ 1; $0 \leq b, d < 1$ then $(a+b) - (c+d) = 0$
and $|a-c| = |d-b|$
but if $|a-c| \neq 0$, it is ≥ 1 but $|d-b| < 1$
hence $|a-c| = 0$, and $|d-b| = 0$)

A mathematician would say that this property of the decimal point is invariant under change of base.

Since the decimal point has an invariant meaning, "decimal point" is perhaps bad terminology since it implies restriction to the base 10. "Radix point" is sometimes used to avoid this prejudice, but we will continue to use "decimal point".

It is important to observe that if a number is multiplied by r then the decimal point in its representation with base r is shifted one position to the right. Similarly, division by r shifts the point one position to the left. These properties of the decimal point will be used shortly in converting from a representation with one base to the representation with a different base.

In general, when changing from an old base to a new base, there are three methods available. Two of these involve arithmetic in the old base and the other involves arithmetic in the new. Usually, either the new or old representation will use a decimal base and advantage may be taken of familiarity with decimal arithmetic in either case. The three methods usually involve:

1) A table of powers of the old base represented in new base, addition in new base arithmetic

2) A table of powers of the new base represented in old base, subtraction in old base arithmetic

3) No tables, but multiplication or division in arithmetic of old base.

In order to represent decimally the octal (base 8) number 372.73, method 1) would usually be used as the others involve arithmetic in the octal notation. With this method, one obtains from a table (perhaps memorized or calculated) the values

$8^2 = 64$

$8^1 = 8$

$8^0 = 1$

$8^{-1} = .125$

$8^{-2} = .015625$

and calculates

$$
\begin{aligned}
3(8^2) &= 192. \\
7(8^1) &= 56. \\
2(8^0) &= 2. \\
7(8^{-1}) &= .875 \\
3(8^{-2}) &= \underline{.046875} \\
& \quad 250.921875
\end{aligned}
$$

as the decimal representation.

In order to represent octally the decimal number 133.375, either method 2) or 3) could be used in the absence of facility with octal arithmetic. With method 2), a table of powers of the new base (8) as decimal numbers is required as before.

$8^3 = 512$

$8^2 = 64$

$8^1 = 8$

$8^0 = 1$

$8^{-1} = .125$

and the procedure is to subtract the largest possible power of 8 as many times as possible without getting a negative remainder. The number of subtractions possible with a given power of 8 is the corresponding digit in the octal representation. Thus,

133.375 -512 neg.	0
133.375 -64 69.375 -64 5.375 -64 neg.	2
5.375 -8 neg.	0
5.375 -1 4.375 -1 3.375 -1 2.375 -1 1.375 -1 .375 -1 neg.	5

(cont. on next page)

```
 .375
-.125
 .250          3
-.125
 .125
-.125
 .000
-.125
  neg.
```

Hence the result is

0205.3 or 205.3 octal

With method 3) no tables would be necessary.

Let 133.375 (dec.) $d_1d_2d_3d_4 \,.\, d_5d_6d_7$ (octal) but from the invariance of the decimal point

$$133(\text{dec}) = d_1d_2d_3d_4(\text{oct})$$

$$.375(\text{dec}) = .d_5d_6d_7(\text{oct})$$

where the d_i represent the digits in the octal representation which are to be found. Then

$$\frac{133}{8} = \frac{d_1(8^3) + d_2(8^2) + d_3(8^1) + d_4(8^0)}{8}$$

$$16\tfrac{5}{8} = d_1(8^2) + d_2(8^1) + d_3(8^0) + d_4(8^{-1}) \simeq d_1d_2d_3\,.d_4$$

and by equating fractional parts

$$\frac{5}{8} = d_4(8^{-1}) \text{ or } \frac{5}{8}(\text{dec}) = .d_4(\text{oct}) \quad 5 = d_4$$

and by equating integral parts

$$16(\text{dec}) = d_1d_2d_3.(\text{oct})$$

dividing by 8 again gives

$$2(\text{dec}) = d_1d_2(\text{oct}) \quad 0(\text{dec}) = .d_3(\text{oct}) \quad 0 = d_3$$

and again

$$0(\text{dec}) = d_1.(\text{oct}) \ \frac{2}{8}(\text{dec}) = .d_2(\text{oct}) \quad 2 = d_2$$

$$d_1 = 0, \text{ hence } 133(\text{dec}) = 205(\text{octal})$$

The fractional part is done similarly by using the invariance of the decimal point (or equating integral and fractional parts) and successive multiplication

$$.375(dec) = .d_5 d_6 d_7(oct)$$

$$8 \times .375 = 3.000(dec) = d_5.\ d_6 d_7(oct)$$

$$3.(d) = d_5(o) \quad d_5 = 2$$

$$.00(d) = .d_6 d_7(o) \quad d_6 = 0 \quad d_7 = 0$$