Base conversion mappings*

by DAVID W. MATULA
Washington University
St. Louis, Missouri

INTRODUCTION AND SUMMARY

Most computers utilize either binary, octal or hexadecimal representation for internal storage of numbers, whereas decimal format is desired for most input, output and for the representation of constants in a user language such as FORTRAN or ALGOL. The conversion of a normalized number in one base to a normalized number in another base is beset with difficulties because of the incompatibility of exact representations of a number when only a limited number of digits are available in each base.

Specifying conversion either by truncation (sometimes called chopping) or by rounding will define a function mapping the numbers representable with a fixed number of digits in one base into the numbers representable with some fixed number of digits in another base. Practical questions which immediately arise are:

1. Is the conversion mapping one to one, i.e., are distinct numbers always converted into distinct numbers?
2. Is the conversion mapping onto, i.e., is it possible to achieve every distinct number representable in the new base by the conversion of some number representable in the old base?

Excluding the case where both bases are powers of some other base (e.g., octal and hexadecimal), we show that a conversion mapping, whether by truncation or rounding, can never be both one-to-one and onto. Specifically, our major result, the Base Conversion Theorem, says: Given a conversion mapping of the space of normalized n digit numbers to the base β into the space of normalized m digit numbers to the base ν, where βi ≠ νj for any non-zero integers i, j, the conversion mapping via truncation or rounding is one-to-one if and only if

\[ ν^{m-1} ≥ β^n - 1 \]

The necessary condition for the existence of a one-to-one conversion intuitively appears overly severe, as the condition \( ν^m ≥ β^n \) is sufficient to assure a one-to-one conversion of all integers from 1 to \( β^n \). However, there are essential difficulties of numerical representation to different bases not related to the nature of the conversion method which indeed make this more restrictive bound necessary. From a practical viewpoint these difficulties will often occur for numbers of a reasonable order of magnitude, i.e., well within the limitations on exponent size required for the normalized number stored in a computer word. For example, conversion of normalized four digit decimal numbers into normalized four digit hexadecimal numbers is not one-to-one, and to see this consider the number 65,536 (i.e. 64K, which is certainly not considered too large by most computer users), and note that

\[ .1000_{10} × 16^4 = 65,536 < .6554 × 10^4 < .6555 × 10^4 < 65,552 = .1001_{16} × 16^4 \]

Thus the two distinct four digit decimal numbers .6554 × 10^4 and .6555 × 10^4 will be converted by truncation into the same four digit hexadecimal number .1001_{16} × 16^4. The necessity of having five hexadecimal digits to distinguish .6554 × 10^4 from .6555 × 10^4 is surprising when it is observed that the four digit decimal integers 1000 through 4095 may all be converted exactly utilizing just three hexadecimal digits.

The conditions of the Base Conversion Theorem assert the impossibility in general of a decimal to binary or binary to decimal mapping being both one-to-one and onto. Thus from a practical viewpoint, some compromise must be made when determining the number of significant digits allowable in the normalized decimal constants of a user language given a fixed size binary (or octal or hexadecimal) word available in some computer hardware. This problem is discussed in the final section.

*This research was performed under USPHS Research Contract No. ES-00139-01.
Significant spaces and normalized numbers

The notion of normalized number relates to how a number may be represented. Care must be taken in forming a precise definition of a space of normalized numbers so that a number is not confused with the representation of the number; i.e., in defining the space of three digit normalized decimal numbers, we do not wish to include ".135" and at the same time exclude ".1350," as they indeed represent the same real number. Hence we wish to think of the "space of n digit normalized numbers to the base $\beta$" not as a set of representations, but as a set of real numbers which may be given a specified form of representation.

Definition: For the integers $\beta \geq 2$, called the base, and $n \geq 1$, called the significance, let $S^n$ be the following set of real numbers

$$S^n = \left\{ x \mid x = \sum_{i=1}^{n} a_i \beta^{-i}, \quad a_i \text{ integers,} \quad 0 \leq a_i \leq \beta - 1 \text{ for } 1 \leq i \leq n \right\}$$

What we have called the significant space, $S^n$, may be interpreted as the space of normalized $\beta$-ary numbers of $n$ significant $\beta$-ary digits along with the number zero, since the elements of $S^n$ other than zero are exactly those numbers which may be given a normalized $n$ $\beta$-ary digit representation. We have tacitly assumed no bound on the exponent but the possible effects of bounding the exponent size are considered in the final section.

Using the natural ordering of the real numbers every number in $S^n$ other than zero will have both a next smallest and next largest neighbor in $S^n$. These numbers play an important part in the discussion of conversion, so that we introduce the following notation.

Definition: For $x \in S^n$, $x \neq 0$, $x'$ is the successor of $x$ in $S^n$ if and only if

$$x' = \min \{ y \mid y > x, \quad y \in S^n \}$$

To avoid ambiguity we occasionally use $[x]'$ for $x'$.

Examples:

In $S^3_{10}$ we have

$$5123' = 5124$$

$$[-.18 \times 10^{15}]' = -.1799 \times 10^{15}$$

$$9.999' = 10$$

$$70700' = 70710$$

Corollary 1: In $S^n$ for any integer $i$

$$\left[ \beta^i (1-\beta^{-n}) \right]' = \beta^i$$

$$\left[ \beta^i \right]' = \beta^i (1+\beta^{-n})$$

Corollary 2: For $x \in S^n$ and any integer $i$

$$\beta^i \left[ x' \right]' = \left[ \beta^i x \right]$$

The relative difference between a number in $S^n$ and its successor is a measure of the accuracy with which arbitrarily chosen real numbers may be approximated in $S^n$. It is desirable to have this relative difference depend on the magnitude of the numbers involved and not on their sign of course they must both have the same sign).

Definition: The gap in $S^n$ at $y \in S^n$ for $y \neq 0$ is denoted by $\gamma^n(y)$ where $\gamma^n(y) = (y' - y)/y$ for $y > 0$

$$\gamma^n(y) = \gamma^n(-y) \quad \text{for } y < 0$$

When there is no ambiguity $\gamma(y)$ may be used for $\gamma^n(y)$. Although the difference $y' - y$ grows monotonically with $y \in S^n$, the gap function $\gamma^n(y)$ varies between fixed bounds for all $y \neq 0$.

$$\min \{ \gamma^n(y) \mid y \in S^n, \quad y \neq 0 \} = 1/(\beta^{n-1})$$

Lemma:

$$\max \{ \gamma^n(y) \mid y \in S^n, \quad y \neq 0 \} = 1/\beta^{n-1}$$

Proof: For any $y \in S^n, y > 0$, a $j$ can be chosen such that

$$y = \alpha_1 \beta^{i-1} + \alpha_2 \beta^{i-2} + \ldots + \alpha_n \beta^{i-n}$$

where $\alpha_i \neq 0, \quad 0 \leq \alpha_i \leq \beta - 1 \quad \text{for } 1 \leq i \leq n.$

Clearly $y' - y = \beta^{i-n}$, so that

$$\gamma^n(y) = (y' - y)/y = \beta^{i-n}(\alpha_1 \beta^{i-1} + \alpha_2 \beta^{i-2} + \ldots + \alpha_n \beta^{i-n})$$

$$= 1/(\alpha_1 \beta^{i-1} + \alpha_2 \beta^{i-2} + \ldots + \alpha_n)$$

The minimum on the right occurs when $\alpha_i = \beta - 1$ for all $i$, and the maximum occurs when $\alpha_1 = 1, \alpha_2 = \alpha_3 = \ldots \alpha_n = 0$.

Thus $1/\beta^{n-1} \leq \gamma^n(y) \leq 1/(\beta^{n-1})$ for all $y \in S^n, y > 0$, and since $\gamma^n(y) = \gamma^n(-y)$ the bounds hold for all $y \in S^n, y \neq 0$.

Note that

$$\gamma^n(\beta^{n-1}) = 1/(\beta^{n-1})$$

and

$$\gamma^n(\beta^{n-1}) = 1/\beta^{n-1}$$
Therefore both bounds are achieved and the lemma is proved.

Definition: The minimum gap \( g(S^o) \) and the maximum gap \( G(S^o) \) are given by
\[
 g(S^o) = \min \{ \gamma^o(y) \mid y \in S^o, y \neq 0 \}
\]
\[
 G(S^o) = \max \{ \gamma^o(y) \mid y \in S^o, y \neq 0 \}
\]

Thus from the Lemma
\[
g(S^o) = 1/(b^n-1)
\]
\[
G(S^o) = 1/b^{n-1}
\]

Observe that for \( y \in S^o, y \neq 0 \), the difference \( y'' - y' \) is either the same as \( y' - y \) or differs by a factor of \( b \), and that on a log-log scale all points of the graph of the gap function will lie on a saw tooth function. Figure 1 shows the saw tooth functions corresponding to \( \gamma^i(y), \gamma^o(y), \gamma^6(y) \), and \( \gamma^3(y) \). Note that there are much bigger jumps in the gap function with a larger base, so that a binary system comes closest to having uniform precision.

It is reasonable to expect that a sufficient condition for a base conversion mapping to be one-to-one is that the minimum gap in the old base be at least as large as the maximum gap in the new base. Figure 1 suggests that if this condition does not hold and if the gap function of the old base does not share some common period with the gap function of the new base, then there will exist some range of numbers where the gap function in the new base is larger than the gap function in the old base. If this region is sufficiently long, a conversion mapping of the old base into the new base could not be one-to-one over this region. To confirm these observations we begin by proving a number theoretic theorem which is of independent interest.

A number theoretic result

Theorem 1:

If \( \beta, \nu > 1 \) are integers such that \( \beta^i \neq \nu^i \) for any positive integers \( i, j \), then there exist some range of numbers \( \beta^i/\nu^i \) where the gap function in the new base is larger than the gap function in the old base. If this region is sufficiently long, a conversion mapping of the old base into the new base could not be one-to-one over this region. To confirm these observations we begin by proving a number theoretic theorem which is of independent interest.

Proof:

\[
\alpha < \beta^i/\nu^i < \alpha + \epsilon
\]

It is sufficient to prove (*) for \( \alpha = 1 \), since interchanging the roles of \( \beta \) and \( \nu \) will then verify (*) for \( 0 < \alpha < 1 \). Furthermore it is now shown that the validity of (*) for \( \alpha = 1 \) implies the result for all \( \alpha > 1 \).

Given \( \alpha > 1, \epsilon > 0 \), let \( \epsilon' = \epsilon/\alpha \) and assume \( m, n > 0 \) can be chosen such that
\[
1 < \beta^m/\nu^n < 1 + \epsilon'
\]

Then \( k \geq 1 \) may be determined so that
\[
(\beta^m/\nu^n)^k > \alpha
\]

Therefore
\[
\alpha < (\beta^m/\nu^n)^k < (\beta^m/\nu^n)^k-1
\]

Hence with \( i = k \cdot m, j = k \cdot n \)
\[
\alpha < \beta^i/\nu^j < \alpha + \epsilon
\]

Thus to prove the theorem we need only show that
\[
(**) \text{ for any } \epsilon > 0 \text{ there exist } i, j > 0
\]
such that
\[
1 < \beta^i/\nu^j < 1 + \epsilon
\]

Let \( \delta = \inf \{ \beta^i/\nu^j \mid \beta^i \neq \nu^i \} \)

Assume \( \delta > 1 \). Then \( k \geq 1 \) may be determined to satisfy
\[
\delta^k \geq \nu
\]

Since then \( \delta^k < \nu \), the definition of \( \delta \) assures the existence of \( m, n > 0 \) and \( \delta' \geq \delta \) where
\[
\delta' = b^m/\nu^n
\]

Thus
\[
(\delta')^k < \nu
\]

But this contradicts the definition of \( \delta \). Therefore \( \delta = 1 \).

It is not possible to have \( i, j > 0 \) for which the limit \( \delta = 1 = \beta^i/\nu^j \) is achieved, since \( \beta^i \neq \nu^j \) for \( i, j > 0 \) by assumption in the theorem. Therefore (***) must hold, proving the theorem.

For completeness we include the following

Corollary:

If \( |\beta|, |\nu| > 1 \) are integers such that \( |\beta^i| \neq |\nu^i| \) for
any non-zero integers $i, j$, and if either $\beta$ or $\nu$ both are negative then the space of rational numbers of the form $\beta^i \nu^j$, where $i, j$ are positive integers, is dense in the real line.

**Proof:**

Apply theorem 1 to $b = \beta^i$ and $v = \nu^j$ to show density in the positive real line for the rationals $b/v$. Then either multiplication by $\beta$ (or division by $\nu$) of the fractions $b/v$ will show the density in the negative reals.

Note that the conditions $\beta^i \neq \nu^j$ for any non-zero integers $i, j$, is equivalent to saying that $\beta$ and $\nu$ are not both powers of some common base. This condition is clearly necessary for theorem 1 to hold as otherwise we would only be able to generate powers of this common base. The fact that this theorem can hold even when $\beta$ divides $\nu$ (or when $\nu$ divides $\beta$) separates this result from a major portion of number theoretic results where a g.c.d. of unity is an essential condition.

**Conversion mappings**

Since the two spaces $\mathbb{S}_\beta$ and $\mathbb{S}_\nu$ are both sets of real numbers, elements of these spaces may be compared with each other by the natural ordering on the real numbers. Utilizing the successor function we now define the conversion mappings.

**Definition:** The truncation conversion map of $\mathbb{S}_\beta$ into $\mathbb{S}_\nu$, $T: \mathbb{S}_\beta \rightarrow \mathbb{S}_\nu$, is defined as follows:

for $x \in \mathbb{S}_\beta$, $y \in \mathbb{S}_\nu$

$T(x) = y \text{ for } x > 0, y \leq x < y'$

$T(x) = y' \text{ for } x < 0, y < x \leq y'$

$T(0) = 0$

**Definition:** The rounding conversion map of $\mathbb{S}_\beta$ into $\mathbb{S}_\nu$, $R: \mathbb{S}_\beta \rightarrow \mathbb{S}_\nu$, is defined by:

for $x \in \mathbb{S}_\beta$, $y \in \mathbb{S}_\nu$

$R(x) = y' \text{ for } x > 0(y + y')/2 \leq x < (y' + y'')/2$

$R(x) = y' \text{ for } x < 0(y + y')/2 < x \leq (y' + y'')/2$

$R(0) = 0$

**Note:** The above definitions are for the usual truncation and rounding by magnitude with sign appended, so that $T(-x) = -T(x)$ and $R(-x) = -R(x)$ for all $x \in \mathbb{S}_\beta$. Assuming a truncation or rounding by algebraic size (for which the above symmetry does not hold) would not change the main body of our results, but would require more tedious proofs, i.e., conversion of negative numbers would have to be treated separately in the proofs.

With these preliminaries we are now ready to state and prove the major result.

**Base Conversion Theorem**

The truncation (rounding) conversion map

$T: \mathbb{S}_\beta \rightarrow \mathbb{S}_\nu \quad (R: \mathbb{S}_\beta \rightarrow \mathbb{S}_\nu)$, where $\beta^i \neq \nu^j$

for any non-zero integers $i, j$,

(1) is one-to-one if and only if $\nu^{m+1} > \beta^m - 1$

(2) is onto if and only if $\beta^{m+1} > \nu^m - 1$

**Proof:**

The theorem will be divided into eight parts.

(i) $\nu^{m+1} > \beta^m - 1 \Rightarrow T: \mathbb{S}_\beta \rightarrow \mathbb{S}_\nu$ is one to one

(ii) $\nu^{m+1} < \beta^m - 1 \Rightarrow T: \mathbb{S}_\beta \rightarrow \mathbb{S}_\nu$ is not one to one

(iii) $\nu^{m+1} > \beta^m - 1 \Rightarrow R: \mathbb{S}_\beta \rightarrow \mathbb{S}_\nu$ is one to one

(iv) $\nu^{m+1} < \beta^m - 1 \Rightarrow R: \mathbb{S}_\beta \rightarrow \mathbb{S}_\nu$ is not one to one

(v) $\beta^{m+1} > \nu^m - 1 \Rightarrow T: \mathbb{S}_\beta \rightarrow \mathbb{S}_\nu$ is onto

(vi) $\beta^{m+1} < \nu^m - 1 \Rightarrow T: \mathbb{S}_\beta \rightarrow \mathbb{S}_\nu$ is not onto

(vii) $\beta^{m+1} > \nu^m - 1 \Rightarrow R: \mathbb{S}_\beta \rightarrow \mathbb{S}_\nu$ is onto

(viii) $\beta^{m+1} < \nu^m - 1 \Rightarrow R: \mathbb{S}_\beta \rightarrow \mathbb{S}_\nu$ is not onto

For brevity we shall prove here only parts i, ii, and vi which are illustrative of the methods used. The complete proof is available in reference (1).

(i) $\nu^{m+1} > \beta^m - 1 \Rightarrow T: \mathbb{S}_\beta \rightarrow \mathbb{S}_\nu$ is one-to-one:

Assume that $\nu^{m+1} > \beta^m - 1$.

Note that this is equivalent to saying that the minimum gap in $\mathbb{S}_\beta$ is greater than or equal to the maximum gap $\mathbb{S}_\nu$.

$g(\mathbb{S}_\beta) \geq g(\mathbb{S}_\nu)$

Given $x, y \in \mathbb{S}_\beta, x \neq y$, we must show $T(x) \neq T(y)$. If either $x$ or $y$ is zero or if $x$ and $y$ are of opposite sign the result is immediate. Furthermore if both $x$ and $y$ are negative and $T(x) = T(y)$, then also $T(-x) = T(-y)$.

Hence it is sufficient to consider the case $x > y > 0$.

$x = y + [(x - y)/y] y$

$\Rightarrow T(x) = g(\mathbb{S}_\beta) T(y)$

$\Rightarrow T(y) + G(\mathbb{S}_\beta) T(y)$

$\Rightarrow T(y) + \left\{[(T(y)]' - T(y)]/T(y) \right\} T(y)$

$= [T(y)]'$

Thus

$T(x) \neq [T(y)]'$ and $T(x) \neq T(y)$. 
This proves the sufficiency of the condition $\nu^{m-1} \geq \beta^n - 1$ for the mapping $T$ to be one-to-one.

(ii) $\nu^{m-1} < \beta^n - 1$ $\Rightarrow$ $T: S^m_n \rightarrow S^n_\nu$ is not one-to-one.

Let $T: S^m_\nu \rightarrow S^n_\nu$ and assume $\nu^{m-1} < \beta^n - 1$.

Multiplying by $(\nu^i - m \beta^{-n})$

$0 < \nu^i \beta^{-n} - \beta^{-m} \nu^{l-m}$

or

$1 < (1-\beta^{-n}) (1+\nu^{l-m})$

so that

$1/(1-\beta^{-n}) > (1+\nu^{l-m})$

By theorem 1 there exists integers $i, j > 0$, such that

$1/(1-\beta^{-n}) < \beta^i \nu^j < 1 + \nu^{l-m}$

Thus

$\nu^i < \beta^i (1-\beta^{-n}) < \beta^i (1-\beta^{-n})$

and using corollary 1 of the definition of successor

$\nu^i < \beta^i (1-\beta^{-n}) = \nu^i (1+\nu^{l-m})$

The inequalities above imply $T(\beta^i (1-\beta^{-n})) = T(\beta^i)$, so that $T: S^m_\nu \rightarrow S^n_\nu$ is not one-to-one.

(vi) $\beta^n - 1 < \nu^{m-1} \Rightarrow$ $T: S^m_\nu \rightarrow S^n_\nu$ is not onto:

Now if $y > x$, then $T: S^m_\nu \rightarrow S^n_\nu$ is not onto.

If $y = x$ by theorem 1 we choose $i, j > 0$ such that

$x'/y' > \nu^i/\beta^i > 1$

and observing that $\beta^i [x] = [\beta^i x]$,

$\nu^i [y] = [\nu^i y]$

$\beta^i x < \nu^i y < [\nu^i y] < [\beta^i x]$.

Hence $\nu^i y \in S^n_\nu$ is not covered and $T$ is not onto.

Observe that the one to one and onto conditions of the Base Conversion Theorem yield immediately the following:

**Corollary**: The truncation (rounding) conversion map of $S^m_\nu$ into $S^n_\nu$ is onto if and only if the truncation (rounding) conversion map of $S^n_\nu$ into $S^n_\nu$ is one-to-one.

**CONCLUSIONS AND APPLICATIONS**

Base conversion introduces a form of error purely of numerical origin and not due to any uncertainty of measurement; the reality of this error, however, cannot be ignored. Just as it is preferable in reporting experimentally determined values to give a number of significant digits which reflect the accuracy of the measurement (no fewer digits, since useful information may be made unavailable to other researchers; no more digits, since the implication of additional precision is unfounded and may lead to spurious deductions), it is similarly desirable, after conversion of a number, to have the number of “significant” digits in the new base reflect the precision previously available with the number of digits used in the old base. Therefore, it is reasonable to require that the new base should have no fewer significant digits than the number necessary to allow the conversion mapping to be onto, nor more significant digits than is necessary for the conversion to be one-to-one. The Base Conversion Theorem has shown, in general, that these bounds on the number of digits must be different, and we now consider what the bounds are and how big are their respective differences.

For the conversion of $S^m_\nu$ into a space with base $\nu$, we define both an onto number and a one to one number.

**Definition**: The onto number, $O(S^m_\nu, \nu)$, and one-to-one number, $I(S^m_\nu, \nu)$, for the conversion mapping of $S^m_\nu$ into a space with base $\nu$, are given by

$O(S^m_\nu, \nu) = \max \{m \mid T: S^m_\nu \rightarrow S^n_\nu \text{ is onto}\}$

$I(S^m_\nu, \nu) = \min \{m \mid T: S^m_\nu \rightarrow S^n_\nu \text{ is one-to-one}\}$

Note that $R$ (rounding conversion) could have been used instead of $T$ (truncation conversion) in the definition, and the same result would be obtained.

From the Base Conversion Theorem we have:

**Corollary 1**: If $\beta^i \neq \nu^j$ for non-zero integers $i, j$, then

(i) $O(S^m_\nu, \nu) = \max \{m \mid \beta^{n-1} \geq \nu^{m-1}, m \text{ an integer}\}$

(ii) $I(S^m_\nu, \nu) = \min \{m \mid \nu^{m-1} \geq \beta^{n-1}, m \text{ an integer}\}$

(iii) $I(S^m_\nu, \nu) > O(S^m_\nu, \nu)$

For computer applications the interesting conversions are between decimal and some variant of binary. Thus for “input” conversions, let $\beta = 10$, $\nu = 2^i$, where $i$ is 1 for binary, 3 for octal, and 4 for hexadecimal. Observe that the condition $10^{n-1} \geq 2^{m-1}$ cannot be an equality for any $n \geq 2$, since the left hand side is then even and greater than zero whereas the right hand side is either odd or zero.

From the collection of the Computer History Museum (www.computerhistory.org)
Thus for \( n \geq 2 \) and \( 10^{n-1} \geq 2^m \) we have

\[
m \leq \frac{(n-1)}{i} \log_2 10
\]

Using the number theoretic function \( \lfloor x \rfloor \) for the greatest integer in \( x \) yields

**Corollary 2:** For \( n \geq 2, i \geq 1, \)

\[
O(S_{10}^n, 2^i) = \lfloor n \log_2 10 \rfloor / i - \lfloor (\log_2 10) / i \rfloor
\]

By reasoning as above with \( n \geq 1, 2^{i(m-1)} \geq 10^n - 1 \) implies that \( 2^{i(m-1)} > 10^n \) so that:

**Corollary 3:** For \( n \geq 1, i \geq 1, \)

\[
I(S_{10}^n, 2^i) = \lfloor n \log_2 10 \rfloor / i + 2
\]

Note that both \( O(S_{10}^n, 2^i) \) and \( I(S_{10}^n, 2^i) \) are, except for the fact that they must take on integral values, linear functions of \( n \) with the same slope \( (\log_2 10) / i \). Thus the difference between \( O(S_{10}^n, 2^i) \) and \( I(S_{10}^n, 2^i) \) is essentially constant and does not grow with \( n \), being equal to either \( \lfloor (\log_2 10) / i \rfloor + 2 \) or \( \lfloor (\log_2 10) / i \rfloor + 3 \).

Table I lists the cases of interest for conversion of decimal input to the commonly used variants of binary representation in computer hardware.

For completeness note that the "output" conversions, \( \beta = 2^i \) and \( \nu = 10 \), can be derived by the above approach.

**Corollary 4:**

\[
O(S_{10}^n, 10) = \lfloor n(i \log 2) - i(\log 2) \rfloor \text{ for } n \geq 1, i \geq 1
\]

\[
I(S_{10}^n, 10) = \lfloor n(i \log 2) + 2 \rfloor \text{ for } n \geq 2, i \geq 1
\]

The criteria for a conversion map to be one-to-one, is strictly speaking, the necessary condition for the whole space \( S^n \) to be mapped into \( S^m \) in a

<table>
<thead>
<tr>
<th>NUMBER OF DECIMAL DIGITS</th>
<th>Maximum for Onto Conversion</th>
<th>Minimum for One-to-One Conversion</th>
<th>Maximum for Onto Conversion</th>
<th>Minimum for One-to-One Conversion</th>
<th>Maximum for Onto Conversion</th>
<th>Minimum for One-to-One Conversion</th>
</tr>
</thead>
<tbody>
<tr>
<td>1</td>
<td>1</td>
<td>5</td>
<td>0</td>
<td>3</td>
<td>0</td>
<td>2</td>
</tr>
<tr>
<td>2</td>
<td>3</td>
<td>8</td>
<td>1</td>
<td>4</td>
<td>0</td>
<td>3</td>
</tr>
<tr>
<td>3</td>
<td>6</td>
<td>11</td>
<td>2</td>
<td>5</td>
<td>1</td>
<td>4</td>
</tr>
<tr>
<td>4</td>
<td>9</td>
<td>15</td>
<td>3</td>
<td>6</td>
<td>2</td>
<td>5</td>
</tr>
<tr>
<td>5</td>
<td>13</td>
<td>18</td>
<td>4</td>
<td>7</td>
<td>3</td>
<td>6</td>
</tr>
<tr>
<td>6</td>
<td>16</td>
<td>21</td>
<td>5</td>
<td>8</td>
<td>4</td>
<td>6</td>
</tr>
<tr>
<td>7</td>
<td>19</td>
<td>25</td>
<td>6</td>
<td>9</td>
<td>4</td>
<td>7</td>
</tr>
<tr>
<td>8</td>
<td>23</td>
<td>28</td>
<td>7</td>
<td>10</td>
<td>5</td>
<td>8</td>
</tr>
<tr>
<td>9</td>
<td>26</td>
<td>31</td>
<td>8</td>
<td>11</td>
<td>6</td>
<td>9</td>
</tr>
<tr>
<td>10</td>
<td>29</td>
<td>35</td>
<td>9</td>
<td>13</td>
<td>7</td>
<td>10</td>
</tr>
<tr>
<td>11</td>
<td>33</td>
<td>38</td>
<td>11</td>
<td>14</td>
<td>8</td>
<td>11</td>
</tr>
<tr>
<td>12</td>
<td>36</td>
<td>41</td>
<td>12</td>
<td>15</td>
<td>9</td>
<td>11</td>
</tr>
<tr>
<td>13</td>
<td>39</td>
<td>45</td>
<td>13</td>
<td>16</td>
<td>9</td>
<td>12</td>
</tr>
<tr>
<td>14</td>
<td>43</td>
<td>48</td>
<td>14</td>
<td>17</td>
<td>10</td>
<td>13</td>
</tr>
<tr>
<td>15</td>
<td>46</td>
<td>51</td>
<td>15</td>
<td>18</td>
<td>11</td>
<td>14</td>
</tr>
<tr>
<td>16</td>
<td>49</td>
<td>55</td>
<td>16</td>
<td>19</td>
<td>12</td>
<td>15</td>
</tr>
<tr>
<td>17</td>
<td>53</td>
<td>58</td>
<td>17</td>
<td>20</td>
<td>13</td>
<td>16</td>
</tr>
<tr>
<td>18</td>
<td>56</td>
<td>61</td>
<td>18</td>
<td>21</td>
<td>14</td>
<td>16</td>
</tr>
<tr>
<td>19</td>
<td>59</td>
<td>65</td>
<td>19</td>
<td>23</td>
<td>14</td>
<td>17</td>
</tr>
<tr>
<td>20</td>
<td>63</td>
<td>68</td>
<td>21</td>
<td>24</td>
<td>15</td>
<td>18</td>
</tr>
<tr>
<td>21</td>
<td>66</td>
<td>71</td>
<td>22</td>
<td>25</td>
<td>16</td>
<td>19</td>
</tr>
<tr>
<td>22</td>
<td>69</td>
<td>75</td>
<td>23</td>
<td>26</td>
<td>17</td>
<td>20</td>
</tr>
<tr>
<td>23</td>
<td>73</td>
<td>78</td>
<td>24</td>
<td>27</td>
<td>18</td>
<td>21</td>
</tr>
<tr>
<td>24</td>
<td>76</td>
<td>81</td>
<td>25</td>
<td>28</td>
<td>19</td>
<td>21</td>
</tr>
<tr>
<td>25</td>
<td>79</td>
<td>85</td>
<td>26</td>
<td>29</td>
<td>19</td>
<td>22</td>
</tr>
</tbody>
</table>

**TABLE I:** Bounds for Onto and One to One Conversions of Decimal to Binary, Octal, or Hexadecimal Systems
one-to-one fashion. Naturally certain ranges of values in \( S^n \) may be mapped in a one-to-one fashion even though the criteria \( \beta^{n-1} \geq \beta^n - 1 \) is not satisfied.

The computer hardware representation of a number places a limit on the size of the exponent as well as on the number of digits allowed. Thus only a given range of numbers in \( S^n \), say, between \( 10^{-40} \) and \( 10^{40} \), need be mapped one-to-one in order to effect a one-to-one conversion. However, unless the decision criteria for a one-to-one conversion, \( \beta^{n-1} \geq \beta^n - 1 \), is very close to being satisfied, there will almost certainly be a value \( x \) well within the range \( 10^{-40} \leq x \leq 10^{40} \) such that \( x \) and \( x' \) are converted to the same number. For example, it may be verified that for decimal to hexadecimal conversion the number .0625 and its successor in \( S_{10} \) will both be truncated to the same value in \( S_{16} \) for all \( n \) from 3 to 25 except \( n=5, 11, 17, 23 \). With decimal to octal conversion, 8 and its successor in \( S_{10} \) will both be truncated to the same value in \( S_8 \), \( j=I(S_{10}, 8)-1 \), for \( n \) from 3 to 25 except \( n=10, 19 \), and an even more critical region for decimal to octal conversion is the neighborhood of \( .991 \times 10^{30} \). Our point is that the necessary bounds reported in Table I are not required because of peculiarities with pathologically large or small numbers, but are generally applicable for the range of magnitude of floating point numbers handled in computer hardware.

For the computer programmer utilizing decimal input, a one-to-one conversion into machine representation is desirable from the point of view of maintaining distinctness between two unequal decimal numbers. On the other hand, any conversion map allows the programmer to achieve every distinct numerical machine word and, therefore, more fully utilize the machine hardware capabilities. Unfortunately decimal-binary conversion cannot be both one-to-one and onto, so some compromise must be made.

There is one case, in this author’s opinion, where the decision in favor of a one-to-one conversion should be mandatory. That is in a more sophisticated implementation of a user language where the number of digits in a decimal constant is used to signal either “single precision” or “double precision” machine representation. Then the allowable constants which specify single precision should be in a range where conversion into machine representation is by a one-to-one mapping. This is necessary to prevent the anomaly of having two distinct decimal constants become identical in the single precision machine representation when these same two decimal constants written with trailing zeros (forcing double precision representation) would receive distinct machine representations in double precision.

For example in an implementation of FORTRAN where normalized decimal constants of 7 or less digits are converted by truncation (or by rounding for that matter) to 6 hexadecimal digit normalized numbers and where 8-16 digit decimal constants are converted to 14 hexadecimal digit normalized numbers, the logical expression

\[
512.0625 \cdot \text{EQ.} \cdot 512.0626
\]

would be true whereas

\[
512.06250 \cdot \text{EQ.} \cdot 512.06260
\]

would be false. This problem could and should be eliminated by specifying, according to our theorem, single precision only for constants of 6 or less decimal digits, and double precision for constants of 7 or more decimal digits. Alternatively a more refined test could be made by first determining if the number \( y \in S_{10} \) falls in a range where \( \gamma_{10}^6 (y) \geq \gamma_{10}^7 (y) \), and if so, then prescribe double precision representation for \( y \).

Figure 2 shows that those intervals where \( \gamma_{10}^6 (y) \) exceeds \( \gamma_{10}^7 (y) \) are erratically spaced and cover a substantial portion, amounting to about 40% on the
log scale, of the interval $1 \leq y \leq 10^{20}$. This agrees very well with the limiting frequency of the range where $\gamma_{16}^6(y)$ exceeds $\gamma_{16}^7(y)$, which is given by $\left[\log(16^{-5}/10^{-7})\right]/(2 \log 16) = 40.68\%$.

Utilizing the comparison of gap functions we can verify that a simple test allowing a one-to-one conversion mapping of $S_{10}^7$ into either $S_{16}^6$ or $S_{16}^4$ is to map a given number of $S_{10}^7$ into $S_{16}^4$ if the leading non-zero hexadecimal digit of the converted number is strictly greater than the leading non-zero digit of the original decimal number, and otherwise map into $S_{16}^4$.

In general, the practical relevance of the Base Conversion Theorem to digital computer data handling can be summarized by observing that if we were to postulate the goals of Table II there is an essential incompatibility between the uniqueness of conversion goal and the optimal utilization of storage goal. Thus when a "systems" decision on the number of digits allowable for decimal constants in the implementation of a user language is made, the user should become aware of where and why the compromise was made to avoid certain "unexpected" results.

REFERENCES

1. D W Matula
   Base Conversion Mappings
   Report No AM-66-1 School of Engineering and Applied Science Washington University St Louis Mo