ON BEYOND ZILLION

NYR INDICTOR
Chappaqua, New York

When Buddha undertook to court Princess Gopa, legend has it, her father had him compete with five other suitors in an array of physical and mental competitions. At one point Buddha's mental prowess was to be tested by the mathematician Arjuna. Arjuna asked him to name all the numbers above *koti*, the Sanskrit word for 10^7. Buddha responded with a long list of arcane number names ("a hundred kotis make an ayuta, a hundred ayutas make a niyuta, a hundred niyutas make a kankara...") up to *tallakshana* (10^{53}), and indicating a way to count to $10^{419} - 1$. The earliest known version of this tale is in the Lalitavistara, which was written in the third century BC, It is safe to say that no other culture was to come up with systematic number names of such high magnitude for another fifteen centuries. Buddhism attached great significance to large numbers, and those cultures which were influenced by Buddhism (Tibet, Burma, Thailand, etc.) generally have high number names.

The Romance languages had words for numbers up to 1000 (Fr. *mille*, Sp. *mil*, etc.) derived from Latin *mil-*. The Germanic tongues generally used cognates of thousand (cf. Ger. *Tausand*) which is ultimately derived from *thus-hand* (strong hundred). When, in the Middle Ages, the need arose for larger numbers, the word million was coined from the same root to mean a thousand thousands (10^6). Billion (a contraction of *bi(s)-million*) was used for a million million (10^{12}) and trillion for a million million millions (10^{18}). By analogy, we get quadrillion (10^{24}), quintillion (10^{30}), sextillion (10^{36}), septillion (10^{42}), octillion (10^{48}), nonillion (10^{54}), decillion (10^{60}). There are higher terms up to centillion (10^{600}), but these are much, much rarer. The word milliard is used occasionally to mean a thousand millions (10^9), but the analogous billiard, trilliard, etc. never caught on in English. (The use of the -illiard series is advocated by proponents of ther artificial language Interlingua.)

Unfortunately, in the US it was decided to count in thousands, rather than millions, so the following values were assigned: billion (10^9), trillion (10^{12}), quadrillion (10^{15}), quintillion (10^{18}), sextillion (10^{21}), septillion (10^{24}), octillion (10^{27}), nonillion (10^{30}), decillion (10^{33}). This slightly off-kilter system makes the centillion a very disappointing 10^{303}. Some countries vacillate between the US and the European system. In the 1960s the term *gillion* was coined for 10^9 in an attempt to develop an unambiguous nomenclature, but this term never caught on. The situation is so confusing that scientists discussing large numbers are much more likely to say "ten to the thirtieth" rather than try to remember whether they mean a nonillion or a quintillion. This inconsistency is not unique to the

-illion numbers: high number names in Sanskrit, Chinese and other languages have all been assigned different values by different users.

The -illion words, especially the word million, were borrowed freely by speakers of languages as disparate as Hawaiian and Hebrew. The more recent fanciful words *googol* (10^{100}) and *googolplex* (10^{googol}) are rarely used because paraphrases such as "ten to the hundredth" are generally so much clearer.

All numerical nomenclatures in natural languages suffer from two shortcomings: arbitrary periodicity and unbearable length. The first of these is avoidable, but at a cost; the second is inevitable. In fact these two problems are closely related,

On Arbitrary Periodicity

In describing counting systems, it is helpful to express periodicity as an array of the basic units from which compound numbers are formed. Consider the following number system of two Australian aboriginal language systems:

	Nickol Bay (period: 1,2,3..)	Port Darwin (period: 1,2)
1	koonjeree	kulagook
2	kootara	kalletillick
3	poorookoo	kalletillick kulagook

The first, which apparently has no numbers beyond three, has no compound forms. The second, which expresses all numbers beyond two as sums of 1 and 2, is a base two system, which has a period of 2. UK English counting has periods of 1..10, 10^2, 10^3, 10^6, while US English has periods of 1..10, 10^2, 10^3. Traditional Chinese has a periodicity like US English, but Classical Mayan had a periodicity of 1..10,20,20^2,20^3..20^6. The complex Sumerian system had periods of 1..10,60,600,3600,36000,216000. It is not uncommon for a counting system to exhibit traces of a periodicity no longer used; an example of this is the French quatre-*vingts* (four-twenties, i.e. eighty), a remnant of a vigesimal system that has been largely replaced by a decimal one.

The problem is that all periodicity is arbitrary, except for the following: 1 one, 2 one-one, 3 one-one-one, ... 100 one-one-...-one-one. This system expresses all numbers uniquely with only one word, but it becomes unwieldy before you get to 10. The first non-trivial system is a binary system, such as the Port Darwin numbers mentioned above. But if we allow the period to increase as in the following example (period 1,2,4,8) the system becomes much more efficient:

1 one	5 four-one	9 eight-one	13 eight-four-one
2 two	6 four-two	10 eight-two	14 eight-four-two
3 two-one	7 four-two-one	11 eight-two-one	15 eight-four-two-one
4 four	8 eight	12 eight-four	

Such systems are actually found in natural language (for example, Archaic Japanese), though no purely binary counting exists. For 16, we must create a new word. We can make the counting system more efficient by making the periodicity increase at a faster rate, as in this example (period 1,2,2x2=4, 4x4=16): 1 one, 2 two, 3 two-one (higher numbers followed by lower numbers are added), 4 four, 5 four-one, 6 four-two, 7 four-two-one, 8 two-four (lower numbers followed by higher numbers are multiplied), 9 two-four-one, .. 12 two-one-four, .. 16 sixteen, .. 47 two-sixteen-two-one-four-two-one, .. 255 two-one-four-sixteen-two-one-sixteen-two-one-four-two-one. Showing the details of the calculation, 255 = [(2+1)x4x16]+[(2+1)x16]+[(2+1)x4]+2+1. Note that, using only the four words one, two, four and sixteen, we can count to 255.

Is there a more efficient system, where N wotrds can be used to name more than 2 raised to the power of 2 N times numbers? Yes, but since we need to express higher order functions than addition and multiplication, we must introduce function words: plus, times, power. Using these, 16 four-power-two, 255 two-one-two-one-power-four-two-one-two-power-four-two-one-four-two-one, 256 four-power-four. Numerically, 255 = $(2+1)x4^{(2+1)}+(2+1)x4^2+(2+1)x4+2+1$.

Amazingly, the next word that would be needed would be the word for 256^{256}! This is counting to well over a googol, and with only six words. The down side, of course, is the tremendous length and complexity of most of the expressions for such terms.

Can we improve on this so that we may count higher with even fewer simple forms? The answer is, again, yes. As long as we keep adding function words for higher order operations, we can keep increasing the functionality of the value words.

Although such a system seems quite alien to everyday usage, a base two periodicity is common in the computer world , where the standard prefixes have been reassigned as follows:

	standard meaning	computer meaning
kilo-	10^3	2^{10} = 1024
mega-	10^4	2^{20} = 1048576
giga-	10^4	2^{30} = 1073741824
tera-	10^5	2^{40} = 1099511627776
peta-	10^6	2^{50} = 1125899906842624

Thus, although a megawatt is 1,000,000 watts, a megabyte is usually understood to mean 1,048,576 bytes.

On Unbearable Length

Let us assume we have created a system for naming every number in English, no matter how large. We know that, in such a system, for any number name, there is always another number that has a longer name. How do we know this? Assume that some number name was the longest number name, and that it had X letters in it. We know that the number of words that can be formed of X or fewer letters is $26^{(X+1)}-1$. Since there are more than this number of number names, there is at least one that is longer than X.

In a series of Word Ways articles (August 1975, February 1976, May 1983, November 1986), John Candelaria developed a nomenclature for very large numbers. Without going into the details of his articles, we can note that in his first nomenclature system the highest numbers were milli-millillion, milli-milli-millillion, milli-milli-milli-millillion, etc. Even these simple forms (a complex form would be three milli-millillion and seven) are destined to become infinitely long. Later, Candelaria came up with ways that allow one to reach higher numbers more efficiently, but he is unable to resolve the problem of the endless name. Since we cannot do away with number names of infinite length, we are left to consider which numbers should have long names and which short. We saw above how a simple round number in English may translate into a complex one in Mayan or Sumerian; this is because number nomenclature systems implicitly direct our attention to the simple forms. We perceive the number one thousand as being "simpler" than two hundred fifty six, though in a binary system the reverse would be true. We may say that, within a number nomenclature system, the shorter the name is, the more important the number. To the Sumerians, the number 216,000 was clearly more important than 1,000,000.

In an ideal system, which numbers does one want to be short, and which long? One answer is to let periodicity determine length, as in the examples above. Other systems with merit have been proposed. George Dalgarno (*Ars Signorum*, London, 1661) in his pioneering proposal for an artificial language created an ingenious system. Each of the ten digits was assigned a vowel and a consonant as follows: 1 a,m, 2 η,n, 3 e,f, 4 o,b, 5 υ,d, 6 u,g, 7 ai,p, 8 ei,t, 9 oi,k, 0 i,l. All numbers were preceded by "v" to distinguish them from other words. Thus, vel = 30, vado = 154 etc.

In this system, every number, no matter how large, has a unique and pronounceable name. Vowels alternate regularly with consonants and the names increase in length with the value of the numbert. An obvious problem is that one decillion would be valilililililililililililililililililililililili-lilililililililili, an impractical number of repetitions for spoken use, certainly. Another problem is that, since the vocalic and consonantal representations of a number are unrelated, the similarities of certain numbers are obscured: 54321 vυbena, 4321 vofηm.

It is possible to improve on Dalgarno's nomenclature and many have tried. Many of these improved systems are discussed by L. Couturat and L. Leau in *Histoire de la langue universelle* (Paris, 1903). One obvious revision is to arrange the consonants and vowels alphabetically, so that, say, b=0 and z=9. Another is to incorporate some device to indicate a superscript. If "w" were used to indicate "raised to the power of", then Dalgarno's system could express a decillion concisely as valwef (10^{33}) or valilwam (1000^{11}).

By far the most ambitious counting system is the one created for the artificial language Lojban. In this language, accounts of which are readily available over the Internet, any mathematical expression, no matter how complex, may be expressed in simple words by means of a small number of rules dictating how these words are used.

To summarize, no set of number names will satisfy both of the following criteria (1) a finite number of value words (simple numbers) and functions (addition, multiplication, ..) are employed, (2) complex numbers are finite in length. Notwithstanding the above, there are an infinite number of practical partial solutions to the number naming problem (and an even larger number of impractical ones). The ideal characteristics of a number naming system depend on the way in which such a system is to be used; in particular, it is important to resolve the related matters of what periodicity is desirable and what numbers are to be singled out by the brevity and simplicity of the names.

Table 1: The US and UK English Counting Systems

	US 10^{3+3x}	UK 10^{6x}		US 10^{3+3x}	UK 10^{6x}
million	10^6	10^6	quattuordecillion (14)	10^{45}	10^{84}
billion (2)	10^9	10^{12}	quindecillion (15)	10^{48}	10^{90}
trillion (3)	10^{12}	10^{18}	sexdecillion (16)	10^{51}	10^{96}
quadrillion (4)	10^{15}	10^{24}	septendecillion (17)	10^{54}	10^{102}
quintillion (5)	10^{18}	10^{30}	octodecillion (18)	10^{57}	10^{108}
sextillion (6)	10^{21}	10^{36}	novemdecillion (19)	10^{60}	10^{114}
septillion (7)	10^{24}	10^{42}	vigintillion (20)	10^{63}	10^{120}
octillion (8)	10^{27}	10^{48}	*trigintillion (30)	10^{93}	10^{180}
nonillion (9)	10^{30}	10^{54}	*quadragintillion (40)	10^{123}	10^{240}
decillion (10)	10^{33}	10^{60}	*septuagintillion (70)	10^{213}	10^{420}
undecillion (11)	10^{36}	10^{66}	*octagintaseptillion (87)	10^{264}	10^{534}
duodecillion (12)	10^{39}	10^{72}	centillion (100)	10^{303}	10^{600}
tredecillion (13)	10^{42}	10^{78}	*nongentillion (900)	10^{2703}	10^{5400}

Notice that the traditional UK system is a more rational one than the innovative US one, inasmuch as the powers expressed are multiples of 6. The US system, which has gained wider acceptance, is more cumbersome because the powers are shifted by 3. Numbers in parentheses indicate the meaning of the Latin prefix (bi-, tri-, etc.). Although the naming system is virtually open-ended, it is very rarely used above *decillion*. Asterisks (*) indicate names that do not exist outside of the literature on number nomenclature and are not standardized. One problem with the system is that round numbers like 10^{100} (US *ten *trigintaduillion*, UK *ten thousand sexdecillion*) require long expressions whereas single words like *tredecillion* refer to quantities of no particular distinction. (Ref.: *Webster's Third New International Dictionary*. Springfield, MA: G. &. C. Merriam Co., 1968.)