

三生万物

万精油



道家说“道生一，一生二，二生三，三生万物”。儒家有“三人为众”，“三人行必有我师”。在中国传统文化里，“三”的地位是很高的。本文要从数学的角度说明，在所有的数字系统中，平衡三进制是最美丽，最优秀的。中国传统文化说“三”是大而广，数学上说“三”是小而精。“三”地位妙不可言。

“一个闷热夏夜，一群人注意到天上的星星……”，这是热门电视剧《生活大爆炸》里谢耳朵给彭妮讲物理的起源时的开场白，说的是两千六百年前古希腊的故事。如果他要讲数的进制的起源，那他得再往前推大约三千年。语言使人类

区别于动物；抽象数字、数字进位则是文明的更进一步。

数字进位使得我们能简易写、记大数。然而，多少数进一位呢？大自然没有给我们一个简易答案。理论上来说，任何进位都可以行得通。我们现在通用十进制，其根本原因是我们有十个手指头。英文“位”的单词 digit 其实就有指头的意思。也有一些文明用二十进制，大约是把脚趾头都算上了。美国土著用八进制，因为他们数的是指间缝隙，而不是指头本身。还有一些部落用 5 进制，大概他们只数一只手。

除了以指头为依据的进位制，还有其它的进位制在一些文明中出现。比如 12 进制（美国现在用的长度单位 1 英尺 = 12 英寸），16 进制（美国现在用的重量单位 1 英磅 = 16 盎司）以及 60 进位（一小时 = 60 分钟），24 进制，32 进制等等。每个进制的存在都有它的道理，我们就不一一细说了。

什么进制最优？前面说大自然没有给我们一个简易答案。这个说法不是完全准确。有些进制还是有先天优势的，比如二进制。二进制每位数只有两个不同的数，0 与 1，在电子线路上这两个数可以用“关”和“开”来表示，非常自然。超过两个值表达起来就没有这么简单了。便于电路表达是一个巨大优势，这也就是为什么我们现在用的计算机都是二进制。

二进制的优势是便于表达。如果除开这个“便于表达”的优势，单从数学上，或者从信息传递的角度来看，它是不是最好的呢？我们这篇文章的主题就是要搞清楚这个问题，比二进制更好的进制是三进制。三进制好在哪里，要回答这个问题，我们先要引进“基需”（Radix Economy）这个概念。

简单说起来，“基需”就是在一个固定基下表示一个数需要的开销。

比如，要表示 1000 以下的数，二进制需要 10 位数，八进制需要 4 位数，而十进



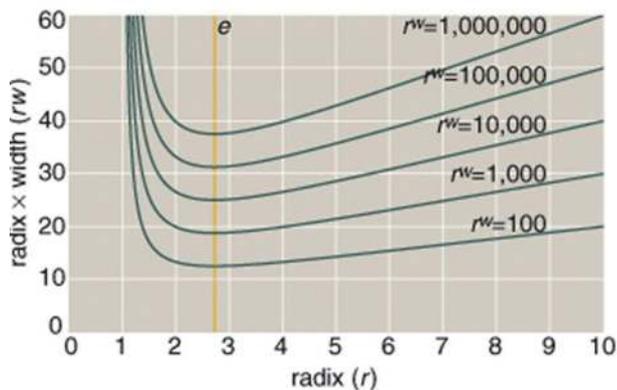
制只需要三位数。但是，位数短是有代价的。十进制每位数有十种不同的值，这比二进制的 0, 1 麻烦多了。值少位数多，位数短值就必须多。怎样把这两个量综合起来考虑呢？“基需”就是这样一个综合量。假设在基为 b 的时候储存每一位数的开销与 b 成正比（因为每位数有 b 个不同的值）， w 位数所需的开销就与 $w \times b$ 成正比。对于任意数 N ，在基 b 下表示数字 N 需要 $\log_b N + 1$ 位数。这样我们就有了基需的精确定义：

$$E(b, N) = b * (\log_b N + 1)$$

基需取整比较容易看得清楚。比如，表示 999 时，二进制下的基需是 $2 \times 10 = 20$ ，10 进制的基需是 $10 \times 3 = 30$ ，而八进制的基需是 $8 \times 4 = 32$ 。

一个很自然的问题是，在什么基下，平均基需最小？如果把这个问题看成一个连续函数的极值问题，那么我们很容易得出 $b = e$ 时基需最低。这是因为 $E(b, N) = b * (\log_b N + 1) \approx b * \log_b N = (b / \ln(b)) * \ln(N)$ 。也就是说在基 b 下，任意数 N 的基需是 $\ln(N)$ 乘上一个固定数 $(b / \ln(b))$ 。而我们知道，函数 $(b / \ln(b))$ 在 $b = e$ 时值最小。最接近 e 的整数是 3，接下来是 2 与 4。

所以，我们从数学上证明了三进制是最经济的进制。下图是基需在不同基、不同位数下的曲线。可以清楚的看见，每条曲线都在 e 处取最小值。在所有整数中，3 处的值最小。



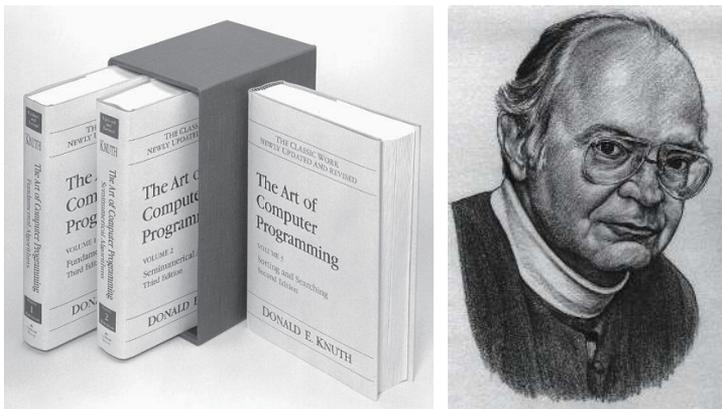
现在我们来再看一看三进制有什么特性。首先，从表达来说，三进制有两种表达方法。一种是非平衡表达法。每位数由 0, 1, 2 表示，到 3 就进位。这与别的进位制基本相同。比如， $100 = 81 + 0 \times 27 + 2 \times 9 + 0 \times 3 + 1 = 10201_3$ 。另一种叫平衡表达式。在介绍平衡表达式以前，我们先来看一个比较有趣天秤砝码问题。

如果我们要求你用一个天秤称出从 1 克到 40 克所有整数克的重量，你最少需要多少砝码，每个砝码是多重？显然，你需要一个重量为 1 克的砝码。如果用通常的 2 倍法，你可以有重量为 1, 2, 4, 8, 16, 32 克的 6 个砝码解决问题。1 到 40 内的所有整数都可以用这些砝码组合出来。能不能用少于 6 个砝码完成这个任务呢？我们可以用 3 代替 2。要称 2 克的东西，只需把 3 放一边，1 放另一边，在 1 那边能让天秤平衡的重量就是 2 克。有了 1, 3，我们可以称出 4 克。下一个砝码是多少呢？用同样的原理，我们可以把下一个砝码重量设为 9。用左右平衡的办法，我们用 1, 3, 9 可以称出 1 到 13 的所有数。以此类推，再下一个就是 27。用 1, 3, 9, 27 可以称出 1 到 40 的所有数。前面问题的答案是 4。4 个砝码就可以称出从 1 克到 40 克所有整数克的重量。这个问题可以推广到任

意数。从数学上来说， $1, 3, 9, 27, \dots, 3^k$ 的和是 $(3^{k+1} - 1)/2$ ，正好是 3^{k+1} 一半下面最大的整数， 3^{k+1} 一半以上的整数都可以用 3^{k+1} 减去前面 3 的幂数达到。所以用 3 的幂数为砝码重量能够称出所有整数，而且是最省砝码个数的方法。有了这个趣题做背景，我们现在可以回头来介绍三进制的平衡表示法。

平衡表示法每位数由 $-1, 0, 1$ 表示。为了进一步表现平衡，那个负号通常放在上面，也就是 $\bar{1}$ 。从上面的天秤砝码问题可以看出，任何整数都可以通过加减 3 的幂数来达到。所以，任何数都可以在三进制下用 $\bar{1}, 0, 1$ 来表示。比如 $100 = 81 + 27 - 9 + 1 = 11\bar{1}01_3$ 。这个平衡表示法的另一个优点是，负数也可以用这样表达，而不需要在前面加负号。比如， $-50 = -81 + 27 + 3 + 1 = \bar{1}0111_3$ 。

平衡表示法在运算中有很大优势。正负数不需要特殊符号，加法与减法基本上是一回事，乘法表也出奇的简单（因为只有 $-1, 0, 1$ ）。归整运算（就是我们平常说的四舍五入）也比二进制简单。平衡三进制是如此优美、自然，以至于计算机大师高德纳（Donald Knuth）在他的名著《计算机编程艺术》中说，最美的数字体系就是平衡三进制（"Perhaps the prettiest number system of all is the balanced ternary notation"）。甚至预言未来世界的计算机应该是三进制的。著名科幻作家罗伯特·海莱因（Robert Heinline）的一篇讲外星人的小说里，外星人把 341640 表述成 122100122100，隐含外星先进文明的计算机是三进制的。



高德纳及其著作《计算机编程艺术》

在二进制下，每一位数叫一个比特（bit），三进制下每一位数叫一个粹特（trit）。在二值逻辑下，1 对应于真，0 对应于假。它们之间的逻辑结合关系见下表：

x	y	$x \wedge y$	$x \vee y$	x	$\neg x$
0	0	0	0	0	1
1	0	0	1	1	0
0	1	0	1		
1	1	1	1		

三值逻辑就没有这么简单了。在三值逻辑下，不仅有真、假，还有一个叫“不确定”。

它们之间的逻辑关系见下表：

(F: FALSE, U:UNKNOWN, T: TRUE)

NOT(A)		AND(A, B)				OR(A, B)						
A	$\neg A$	$A \wedge B$		B			$A \vee B$		B			
F	T	A	F	F	U	T	F	U	T	F	U	T
U	U		U	F	U	U	U	U	U	U	U	U
T	F		T	F	U	T	T	T	T	T	T	T

(- 1: FALSE, 0:UNKNOWN, + 1: TRUE)

NEG(A)		MIN(A, B)				MAX(A, B)						
A	$\neg A$	$A \wedge B$		B			$A \vee B$		B			
-1	+1	A	-1	-1	0	+1	-1	0	+1	-1	0	+1
0	0		0	-1	0	0	0	0	0	+1	+1	+1
+1	-1		+1	-1	0	+1	+1	+1	+1	+1	+1	+1

看起来比较复杂，但复杂的东西搞清楚了就有优势。比如，二进制下，NaN (Not a Number) 必须特殊处理。而三值逻辑就可以用“不确定”来表示（数据库软件 SQL 就是用三值逻辑来处理 NaN）。另外，二进制下比较两个数，要考虑三种情况，>, =, <, 一般要两步完成，而在三值逻辑中，只需一次就到位。其它还有很多有趣的优点。

理论上可以证明，三进制计算机比二进制计算机效率高，有明显优势。这么多优势，为什么我们现在的计算机却都是二进制的呢？主要原因是二值逻辑一开一关很容易有对应的电子表达。三值粹特除了一开一关，还有另一个状态，很难表达。很难用电子线路表达一个半开半关的状态。所以，开发三进制计算机的打不过二进制计算机，以至于现在几乎全军覆没。



Setun 三进制计算机

在开发三进制计算机的努力中，有一个特别值得一提。那就是由前苏联莫斯科大学的数学家谢尔盖·索伯列夫 (Sergei Sobolev) 与计算机科学家尼古拉·布鲁森佐夫 (Nikolay Brusentsov) 在 1958 年领导制造的 Setun 三进制计算机，比当时的二进制计算机有明显的优点，比如造价低，耗能少