Fibonacci coding

From Wikipedia, the free encyclopedia

In mathematics and computing, **Fibonacci coding** is a universal code which encodes positive integers into binary code words. It is one example of representations of integers based on Fibonacci numbers. Each code word ends with "11" and contains no other instances of "11" before the end.

The Fibonacci code is closely related to the *Zeckendorf representation*, a positional numeral system that uses Zeckendorf's theorem and has the property that no number has a representation with consecutive 1s. The Fibonacci code word for a particular integer is exactly the integer's Zeckendorf representation with the order of its digits reversed and an additional "1" appended to the end.

Numeral systems

Hindu–Arabic
Eastern Arabic · Western Arabic
Bengali · Indian · Sinhala · Tamil
Burmese · Khmer · Lao · Mongolian · Thai
Balinese · Javanese · Dzongkha

East Asian
Chinese (Suzhou) · Japanese · Korean · Vietnamese
Counting rods

Alphabetic
Abjad · Armenian · Āryabhaṭa · Cyrillic · Ge'ez · Georgian · Greek · Hebrew · Roman

Former
Aegean · Attic · Babylonian · Brahmi · Egyptian · Etruscan · Inuit · Kharosthi · Mayan · Quipu
Prehistoric

Positional systems by base
2 · 3 · 4 · 5 · 6 · 8 · 10 · 12 · 16 · 20 · 60

Non-standard positional numeral systems
Bijective numeration (1) · Signed-digit representation (Balanced ternary) · factorial · negative · Complex base systems (2*i*) · Non-integer representation (φ) · mixed

List of numeral systems

v · t · e

Contents [hide]

Definition [edit]

For a number N, if $d(0), d(1), \ldots, d(k-1), d(k)$ represent the digits of the code word representing N then we have:

$$N = \sum_{i=0}^{k-1} d(i)F(i+2), \text{ and } d(k-1) = d(k) = 1.$$

where $F(i)$ is the ith Fibonacci number, and so $F(i+2)$ is the ith distinct Fibonacci number starting with $1, 2, 3, 5, 8, 13, \ldots$. The last bit $d(k)$ is always an appended bit of 1 and does not carry place value.

It can be shown that such a coding is unique, and the only occurrence of "11" in any code word is at the end i.e. *d*(*k*−1) and *d*(*k*). Note that the penultimate bit is the most significant bit and the first bit is the least significant bit. Note also that leading zeros cannot be omitted as they can in e.g. decimal numbers.

The first few Fibonacci codes are shown below, and also the so-called *implied distribution*, the distribution of values for which Fibonacci coding gives a minimum-size code.

Symbol	Fibonacci representation	Fibonacci code word	implied distribution
1	$F(2)$	11	1/4
2	$F(3)$	011	1/8
3	$F(4)$	0011	1/16

4	$F(2)+F(4)$	1011	1/16
5	$F(5)$	00011	1/32
6	$F(2)+F(5)$	10011	1/32
7	$F(3)+F(5)$	01011	1/32
8	$F(6)$	000011	1/64
9	$F(2)+F(6)$	100011	1/64
10	$F(3)+F(6)$	010011	1/64
11	$F(4)+F(6)$	001011	1/64
12	$F(2)+F(4)+F(6)$	101011	1/64
13	$F(7)$	0000011	1/128
14	$F(2)+F(7)$	1000011	1/128

To encode an integer *N*:

1. Find the largest Fibonacci number equal to or less than *N*; subtract this number from *N*, keeping track of the remainder.
2. If the number subtracted was the *i*th Fibonacci number *F*(*i*), put a 1 in place *i*−2 in the code word (counting the left most digit as place 0).
3. Repeat the previous steps, substituting the remainder for *N*, until a remainder of 0 is reached.
4. Place an additional 1 after the rightmost digit in the code word.

To decode a code word, remove the final "1", assign the remaining the values 1,2,3,5,8,13... (the Fibonacci numbers) to the bits in the code word, and sum the values of the "1" bits.

Comparison with other universal codes [edit]

Fibonacci coding has a useful property that sometimes makes it attractive in comparison to other universal codes: it is an example of a self-synchronizing code, making it easier to recover data from a damaged stream. With most other universal codes, if a single bit is altered, none of the data that comes after it will be correctly read. With Fibonacci coding, on the other hand, a changed bit may cause one token to be read as two, or cause two tokens to be read incorrectly as one, but reading a "0" from the stream will stop the errors from propagating further. Since the only stream that has no "0" in it is a stream of "11" tokens, the total edit distance between a stream damaged by a single bit error and the original stream is at most three.

This approach - encoding using sequence of symbols, in which some patterns (like "11") are forbidden, can be freely generalized [1].

Example [edit]

The following table shows that the number 65 is represented in Fibonacci coding as 0100100011, since 65 = 2 + 8 + 55. The first two Fibonacci numbers (0 and 1) are not used, and an additional 1 is always appended.

0	1	1	2	3	5	8	13	21	34	55	–
$F(0)$	$F(1)$	$F(2)$	$F(3)$	$F(4)$	$F(5)$	$F(6)$	$F(7)$	$F(8)$	$F(9)$	$F(10)$	additional
–	–	0	1	0	0	1	0	0	0	1	1

Generalizations [edit]

The Fibonacci encodings for the positive integers are binary strings that end with "11" and contain no other instances of "11". This can be generalized to binary strings that end with *N* consecutive 1's and contain no other instances of *N* consecutive 1's. For instance, for *N* = 3 the positive integers are encoded as 111, 0111, 00111, 10111, 000111, 100111, 010111, 110111, 0000111, 1000111, 0100111, …. In this case, the number of encodings as a function of string length is given by the sequence of Tribonacci numbers.

See also [edit]

- Golden ratio base
- Ostrowski numeration

- Universal code
- Varicode, a practical application
- Zeckendorf's theorem

References [edit]

- Allouche, Jean-Paul; Shallit, Jeffrey (2003). *Automatic Sequences: Theory, Applications, Generalizations*. Cambridge University Press. p. 105. ISBN 978-0-521-82332-6. Zbl 1086.11015.
- Fraenkel, Aviezri S.; Klein, Shmuel T. (1996). "Robust universal complete codes for transmission and compression". *Discrete Applied Mathematics* **64** (1): 31–55. doi:10.1016/0166-218X(93)00116-H. ISSN 0166-218X. Zbl 0874.94026.

Further reading [edit]

- Stakhov, A. P. (2009). *The Mathematics of Harmony: From Euclid to Contemporary Mathematics and Computer Science*. Singapore: World Scientific Publishing.

Lossless	Entropy type	Unary · Arithmetic · Golomb · Huffman (Adaptive · Canonical · Modified) · Range · Shannon · Shannon–Fano · Shannon–Fano–Elias · Tunstall · Universal (Exp-Golomb · **Fibonacci** · Gamma · Levenshtein)
	Dictionary type	Byte pair encoding · DEFLATE · Lempel–Ziv (LZ77 / LZ78 (LZ1 / LZ2) · LZJB · LZMA · LZO · LZRW · LZS · LZSS · LZW · LZWL · LZX · LZ4 · Statistical)
	Other types	BWT · CTW · Delta · DMC · MTF · PAQ · PPM · RLE
Audio	Concepts	Bit rate (average (ABR) · constant (CBR) · variable (VBR)) · Companding · Convolution · Dynamic range · Latency · Nyquist–Shannon theorem · Sampling · Sound quality · Speech coding · Sub-band coding
	Codec parts	A-law · μ-law · ACELP · ADPCM · CELP · DPCM · Fourier transform · LPC (LAR · LSP) · MDCT · Psychoacoustic model · WLPC
Image	Concepts	Chroma subsampling · Coding tree unit · Color space · Compression artifact · Image resolution · Macroblock · Pixel · PSNR · Quantization · Standard test image
	Methods	Chain code · DCT · EZW · Fractal · KLT · LP · RLE · SPIHT · Wavelet
Video	Concepts	Bit rate (average (ABR) · constant (CBR) · variable (VBR)) · Display resolution · Frame · Frame rate · Frame types · Interlace · Video characteristics · Video quality
	Codec parts	Lapped transform · DCT · Deblocking filter · Motion compensation
Theory	Entropy · Kolmogorov complexity · Lossy · Quantization · Rate–distortion · Redundancy · Timeline of information theory	

Compression formats · Compression software (codecs)

Categories: Non-standard positional numeral systems | Lossless compression algorithms | Fibonacci numbers