It's Been 1,000,000 Years Since Huffman

Alistair Moffat

The University of Melbourne, Australia
`ammoffat@unimelb.edu.au`

David Huffman's algorithm for computing minimum-redundancy prefix-free codes [1] has legendary status in the computing disciplines. Its elegant blend of simplicity and applicability has made it a favorite example in algorithms courses, and if all of the class assignments are included, it is perhaps one of the most implemented techniques in computer science. Huffman's seminal paper now has over 5,000 citations.

The origins of Huffman coding are captured by Gary Stix [3], who recounts a tale that Huffman told to a number of people. While enrolled as a graduate student at MIT in 1951 in a class taught by coding pioneer Robert Fano, Huffman and his fellow students were told that they would be exempted from the final exam if they solved a coding challenge as part of a term paper. Not realizing that the task was an open problem that Fano had been unable to solve himself, Huffman elected to work on the term paper. After months of unsuccessful struggle, and with the final exam just days away, Huffman threw his attempts in the bin, and started to prepare for the exam. A flash of insight the next morning had him realize that the attempt he had thrown in the trash was in fact a bottom-up strategy that would lead to a solution to the problem. Huffman coding was born at that moment. Following publication of his paper in *Proceedings of the Institute of Radio Engineers* (the predecessor of *Proceedings of the IEEE*) in 1952, the new technique quickly replaced the previous Shannon-Fano coding as the method of choice for data compression applications.

David Huffman died in October 1999, at the age of 74, shortly after being awarded the 1999 IEEE Richard W. Hamming Medal "for design procedures of minimum redundancy (Huffman) codes and asynchronous sequential circuits, and contributions to analysis of visual imagery". Huffman codes continue to be relevant, and are embedded in a wide range of critically important communications and storage codecs. With 2015 marking the 64 th anniversary of their development – 1,000,000 years in binary – it is timely to review Huffman and related codes, and the many mechanisms that have been developed for computing and deploying them [2].

[1] D. A. Huffman. A method for the construction of minimum-redundancy codes. *Proc. Inst. Radio Engineers*, 40(9):1098–1101, Sept. 1952.

[2] A. Moffat. Huffman coding. In M.-Y. Kao, editor, *Encyclopedia of Algorithms*. Springer, Oct. 2014. doi: 10.1007/978-3-642-27848-8_633-1.

[3] G. Stix. Encoding the 'neatness' of ones and zeroes: A profile of information theorist David A. Huffman. *Scientific American*, 265(3):54,58, Sept. 1991. Reproduced online at `http://www.huffmancoding.com/my-uncle/scientific-american`, accessed 17 Aug., 2014.