Elias Coding

Mário A. T. Figueiredo,
Departamento de Engenharia Electrotécnica e de Computadores,
Instituto Superior Técnico, Lisboa, Portugal

October 2009

This short lecture note describes and analyzes two techniques proposed by Elias to build instantaneous binary codes for arbitrary integers.

1 The Elias Gamma Code

The *Elias gamma code* is the simplest of the Elias codes and is defined as follows. To code a natural number $x \in \mathbb{N} = \{1, 2, 3, ...\}$, write its binary representation preceded by $\lfloor \log_2 x \rfloor$ zeros. Notice that $\lfloor \log_2(x) \rfloor + 1$ is the number of digits required to write x in binary.

For example, for $x = 13$, the binary representation is 1101, and $\lfloor \log_2 13 \rfloor = 3$ thus the Elias gamma code word is $C_\gamma(13) = 0001101$.

It is very easy to verify that this constitutes an instantaneous code, simply by studying the decoding procedure. The decoder starts by counting the number, say n, of zeros in the beginning of code word; if n is zero, then the decoded number is 1; if n is not zero, then the decoder reads the following $n + 1$ binary digits and decodes the corresponding binary number. In Table 1, the Elias gamma codes for several integers are listed.

2 The Elias Delta Code

The *Elias delta code* is somewhat more sophisticated and uses the Elias gamma code as a building block. We begin by presenting a modified version, which we will denote as $\widetilde{C}_\delta$. To code a natural number $x \in \mathbb{N} = \{1, 2, 3, ...\}$, the code word $\widetilde{C}_\delta(x)$ is obtained as follows: write its binary representation preceded by $C_\gamma(\lfloor \log_2(x) \rfloor + 1)$. Recall that $\lfloor \log_2(x) \rfloor + 1$ is the number of digits required to write x in binary.

For example, for $x = 13$, the binary representation is 1101; computing $\lfloor \log_2 13 \rfloor = 3$, we have $C_\gamma(4) = 00100$; finally, $\widetilde{C}_\delta(13) = 001001101$.

The final version of the Elias delta code, denoted C_δ, is obtained by observing that we don't need the first "1" in the binary representation of x, since any binary representation starts with a "1". Thus, in the example above, we have $C_\delta(13) = 00100101$. As another example, consider

Table 1: A few examples of Elias gamma code words for integers.

x	$C_\gamma(x)$
1	1
2	010
3	011
4	00100
5	00101
6	00110
7	00111
8	0001000
9	0001001
10	0001010
$\vdots$	$\vdots$
19	000010011
$\vdots$	$\vdots$
147	000000010010011

$x = 7$: the binary representation is 111; next, we have $\lfloor \log_2 7 \rfloor + 1 = 3$ and $C_\gamma(3) = 011$; finally, $C_\delta(7) = 01111$.

Again, to verify that C_δ is decodable instantaneously, it suffices to study the decoding procedure: first, we decode the Elias gamma code that resides in the first bits of the code word (we verified above that this can be made without any ambiguity); the result of this decoding provides the decoder with knowledge of the number of digits, say b, in the binary representation of the coded number; finally, the decoder reads the following $b - 1$ digits, inserts a "1" at the beginning, and decodes the resulting binary representation. As an example, we describe the decoding of the code word 001010001: first, to decode the Elias gamma code at the beginning, we count the number of zeros, which is 2; this means we need to read the next 3 digits, 101; decoding 101, gives 5, meaning that the coded number has six digits, the first of which is a 1 (it always is), that is, 10001; finally, decoding this binary representation gives 17.

Finally, Table 2 shows some examples of Elias delta code words.

Table 2: A few examples of Elias delta code words for integers.

x	$C_\delta(x)$
1	1
2	0100
3	0101
4	01100
5	01101
6	01110
7	01111
8	00100000
9	00100001
10	00100010
$\vdots$	$\vdots$
19	001010011
$\vdots$	$\vdots$
147	00010000010011

3 Comparison of the Two Codes

Consider the use of the two codes above described to encode integers in the set $\{1, 2, ..., N\}$, where $N >> 1$.

Let $l_\gamma(x)$ and $l_\delta(x)$ denote the number of bits of $C_\gamma(x)$ and $C_\delta(x)$, respectively. Then, its clear that

$$l_\gamma(x) = 2\lfloor \log_2 x \rfloor + 1$$

and

$$\begin{aligned} l_\delta(x) &= l_\gamma(\lfloor \log_2 x \rfloor + 1) + \lfloor \log_2 x \rfloor \\ &= 2\lfloor \log_2(\lfloor \log_2 x \rfloor + 1) \rfloor + 1 + \lfloor \log_2 x \rfloor. \end{aligned} \quad (1)$$

Consider that X is a random variable with uniform distribution on $\{1, 2, ..., N\}$, thus with entropy $H(X) = \log_2 N$ bits/symbol. The expected length of the Elias gamma code is

$$E[l_\gamma(X)] = \frac{1}{N} \sum_{x=1}^{N} (2\lfloor \log_2 x \rfloor + 1).$$

Observing that $\lfloor a \rfloor > a - 1$, for any a, we can write

$$E[l_\gamma(X)] > \frac{1}{N} \sum_{x=1}^{N} (2 \log_2 x - 1) = \frac{2}{N} \log \left(\prod_{x=1}^{N} x \right) - 1 = \frac{2}{N} \log_2(N!) - 1. \quad (2)$$

Finally, using the fact that

$$\lim_{t\to\infty} \frac{\log(t!)}{t\ \log(t)} = 1, \tag{3}$$

we can conclude that

$$\lim_{N\to\infty} \frac{E\left[l_\gamma(X)\right]}{\log N} \geq 2$$

showing that for very large N, the expected length of the Elias gamma code approaches twice the entropy, thus being clearly non-optimal.

Considering now the Elias delta code, and proceeding as above,

$$\begin{aligned} E\left[l_\delta(X)\right] &\leq \frac{1}{N}\sum_{x=1}^{N}\left(2\ \log_2(\log_2 x+1)+1+\log_2 x\right) \\ &= \frac{1}{N}\log\left(\prod_{x=1}^{N} x\right)+\frac{2}{N}\sum_{x=1}^{N}\log_2(\log_2 x+1)+1 \\ &\leq \frac{1}{N}\log\left(N!\right)+2\ \log_2(\log_2 N+1)+1, \end{aligned} \tag{4}$$

because $\log_2(\log_2 x+1) \leq \log_2(\log_2 N+1)$, for any $x \leq N$. Finally, using (3) and the fact that

$$\lim_{t\to\infty} \frac{\log(\log t+1)}{\log(t)} = 0, \tag{5}$$

we can conclude that

$$\lim_{N\to\infty} \frac{E\left[l_\delta(X)\right]}{\log N} = 1$$

showing that for very large N, the expected length of the Elias delta code approaches the entropy, thus being asymptotically optimal.

References

[1] P. Elias, "Universal codeword sets and representations of the integers." *IEEE Transactions on Information Theory*, vol. 21, no. 2, pp. 194-203, March 1975.