Information Theory

Mohamed Hamada

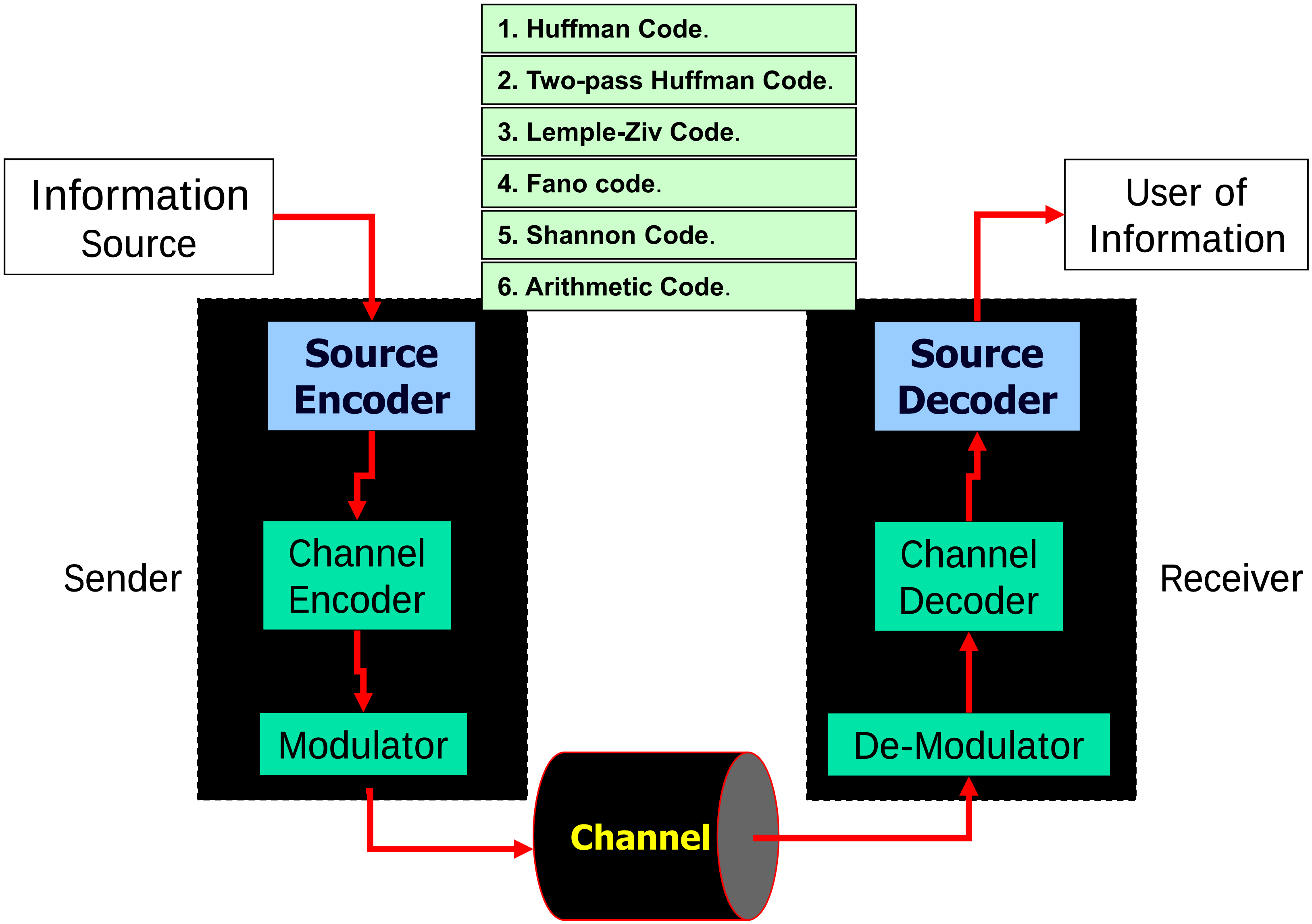
Software Engineering Lab
The University of Aizu

Email: hamada@u-aizu.ac.jp
URL: http://www.u-aizu.ac.jp/~hamada

Today's Topics

- Source Coding Techniques
- Fano Code
- Coding Examples
- Shannon Code
- Code Comparison

Source Coding Techniques



Source Coding Techniques

1. Huffman Code.

2. Two-pass Huffman Code.

3. Lemple-Ziv Code.

4. Fano code.

5. Shannon Code.

6. Arithmetic Code.

Source Coding Techniques

1. **Huffman Code.**

2. **Two-path Huffman Code.**

3. **Lemple-Ziv Code.**

4. **Fano Code.**

5. **Shannon Code.**

6. **Arithmetic Code.**

4. Fano Code.

(Also called Shannon-Fano code)

The Fano code is a technique for constructing a prefix code

It is suboptimal in the sense that it does not achieve the lowest possible expected code word length like Huffman coding.

However unlike Huffman coding, it does guarantee that all code word lengths are within one bit of their theoretical ideal $log(1/P(x))$.

Fano coding is used in the IMPLODE compression method, which is part of the ZIP file format.

4. Fano Code.

The Fano code is performed as follows:

1. arrange the information source symbols in order of decreasing probability
2. divide the symbols into two equally probable groups, as possible as you can
3. each group receives one of the binary symbols (i.e. 0 or 1) as the first symbol
4. repeat steps 2 and 3 per group as many times as this is possible.
5. stop when no more groups to divide

4. Fano Code.

Example 1:

1. arrange the information source symbols in order of decreasing probability

Symbol	Probability	Fano Code
A	1/4	
B	1/4	
C	1/8	
D	1/8	
E	1/16	
F	1/16	
G	1/32	
H	1/32	
I	1/32	
J	1/32	

4. Fano Code.

Example 1:

2. divide the symbols into two equally probable groups, as possible as you can

Symbol	Probability	Fano Code
A	1/4	
B	1/4	
C	1/8	
D	1/8	
E	1/16	
F	1/16	
G	1/32	
H	1/32	
I	1/32	
J	1/32	

4. Fano Code.

Example 1:

3. each group receives one of the binary symbols (i.e. 0 or 1) as the first symbol

Symbol	Probability	Fano Code
A	1/4	0
B	1/4	0
C	1/8	1
D	1/8	1
E	1/16	1
F	1/16	1
G	1/32	1
H	1/32	1
I	1/32	1
J	1/32	1

4. Fano Code.

Example 1:

4. repeat steps 2 and 3 per group as many times as this is possible.

Symbol	Probability	Fano Code
A	1/4	0
B	1/4	0
C	1/8	1
D	1/8	1
E	1/16	1
F	1/16	1
G	1/32	1
H	1/32	1
I	1/32	1
J	1/32	1

4. Fano Code.

Example 1:

2. divide the symbols into two equally probable groups, as possible as you can

Symbol	Probability	Fano Code
A	1/4	0
B	1/4	0
C	1/8	1
D	1/8	1
E	1/16	1
F	1/16	1
G	1/32	1
H	1/32	1
I	1/32	1
J	1/32	1

4. Fano Code.

Example 1:

3. each group receives one of the binary symbols (i.e. 0 or 1) as the first symbol

Symbol	Probability	Fano Code
A	1/4	0 0
B	1/4	0 1
C	1/8	1
D	1/8	1
E	1/16	1
F	1/16	1
G	1/32	1
H	1/32	1
I	1/32	1
J	1/32	1

4. Fano Code.

Example 1:

Symbol	Probability	Fano Code
A	1/4	0 0
B	1/4	0 1
C	1/8	1
D	1/8	1
E	1/16	1
F	1/16	1
G	1/32	1
H	1/32	1
I	1/32	1
J	1/32	1

4. Fano Code.

Example 1:

2. divide the symbols into two equally probable groups, as possible as you can

Symbol	Probability	Fano Code
A	1/4	0 0
B	1/4	0 1
C	1/8	1
D	1/8	1
E	1/16	1
F	1/16	1
G	1/32	1
H	1/32	1
I	1/32	1
J	1/32	1

4. Fano Code.

Example 1:

3. each group receives one of the binary symbols (i.e. 0 or 1) as the first symbol

Symbol	Probability	Fano Code
A	1/4	0 0
B	1/4	0 1
C	1/8	1 0
D	1/8	1 0
E	1/16	1 1
F	1/16	1 1
G	1/32	1 1
H	1/32	1 1
I	1/32	1 1
J	1/32	1 1

4. Fano Code.

Example 1:

4. repeat steps 2 and 3 per group as many times as this is possible.

Symbol	Probability	Fano Code
A	1/4	0 0
B	1/4	0 1
C	1/8	1 0
D	1/8	1 0
E	1/16	1 1
F	1/16	1 1
G	1/32	1 1
H	1/32	1 1
I	1/32	1 1
J	1/32	1 1

4. Fano Code.

Example 1:

2. divide the symbols into two equally probable groups, as possible as you can

Symbol	Probability	Fano Code
A	1/4	0 0
B	1/4	0 1
C	1/8	1 0
D	1/8	1 0
E	1/16	1 1
F	1/16	1 1
G	1/32	1 1
H	1/32	1 1
I	1/32	1 1
J	1/32	1 1

4. Fano Code.

Example 1:

3. each group receives one of the binary symbols (i.e. 0 or 1) as the first symbol

Symbol	Probability	Fano Code
A	1/4	0 0
B	1/4	0 1
C	1/8	1 0 0
D	1/8	1 0 1
E	1/16	1 1
F	1/16	1 1
G	1/32	1 1
H	1/32	1 1
I	1/32	1 1
J	1/32	1 1

4. Fano Code.

Example 1:

Symbol	Probability	Fano Code
A	1/4	0 0
B	1/4	0 1
C	1/8	1 0 0
D	1/8	1 0 1
E	1/16	1 1
F	1/16	1 1
G	1/32	1 1
H	1/32	1 1
I	1/32	1 1
J	1/32	1 1

4. Fano Code.

Example 1:

2. divide the symbols into two equally probable groups, as possible as you can

Symbol	Probability	Fano Code
A	1/4	0 0
B	1/4	0 1
C	1/8	1 0 0
D	1/8	1 0 1
E	1/16	1 1
F	1/16	1 1
G	1/32	1 1
H	1/32	1 1
I	1/32	1 1
J	1/32	1 1

4. Fano Code.

Example 1:

3. each group receives one of the binary symbols (i.e. 0 or 1) as the first symbol

Symbol	Probability	Fano Code
A	1/4	0 0
B	1/4	0 1
C	1/8	1 0 0
D	1/8	1 0 1
E	1/16	1 1 0
F	1/16	1 1 0
G	1/32	1 1 1
H	1/32	1 1 1
I	1/32	1 1 1
J	1/32	1 1 1

4. Fano Code.

Example 1:

2. divide the symbols into two equally probable groups, as possible as you can

Symbol	Probability	Fano Code
A	1/4	0 0
B	1/4	0 1
C	1/8	1 0 0
D	1/8	1 0 1
E	1/16	1 1 0
F	1/16	1 1 0
G	1/32	1 1 1
H	1/32	1 1 1
I	1/32	1 1 1
J	1/32	1 1 1

4. Fano Code.

Example 1:

3. each group receives one of the binary symbols (i.e. 0 or 1) as the first symbol

Symbol	Probability	Fano Code
A	1/4	0 0
B	1/4	0 1
C	1/8	1 0 0
D	1/8	1 0 1
E	1/16	1 1 0 0
F	1/16	1 1 0 1
G	1/32	1 1 1
H	1/32	1 1 1
I	1/32	1 1 1
J	1/32	1 1 1

4. Fano Code.

Example 1:

2. divide the symbols into two equally probable groups, as possible as you can

Symbol	Probability	Fano Code
A	1/4	0 0
B	1/4	0 1
C	1/8	1 0 0
D	1/8	1 0 1
E	1/16	1 1 0 0
F	1/16	1 1 0 1
G	1/32	1 1 1
H	1/32	1 1 1
I	1/32	1 1 1
J	1/32	1 1 1

4. Fano Code.

Example 1:

3. each group receives one of the binary symbols (i.e. 0 or 1) as the first symbol

Symbol	Probability	Fano Code
A	1/4	0 0
B	1/4	0 1
C	1/8	1 0 0
D	1/8	1 0 1
E	1/16	1 1 0 0
F	1/16	1 1 0 1
G	1/32	1 1 1 0
H	1/32	1 1 1 0
I	1/32	1 1 1 1
J	1/32	1 1 1 1

4. Fano Code.

Example 1:

2. divide the symbols into two equally probable groups, as possible as you can

Symbol	Probability	Fano Code
A	1/4	0 0
B	1/4	0 1
C	1/8	1 0 0
D	1/8	1 0 1
E	1/16	1 1 0 0
F	1/16	1 1 0 1
G	1/32	1 1 1 0
H	1/32	1 1 1 0
I	1/32	1 1 1 1
J	1/32	1 1 1 1

4. Fano Code.

Example 1:

3. each group receives one of the binary symbols (i.e. 0 or 1) as the first symbol

Symbol	Probability	Fano Code
A	1/4	0 0
B	1/4	0 1
C	1/8	1 0 0
D	1/8	1 0 1
E	1/16	1 1 0 0
F	1/16	1 1 0 1
G	1/32	1 1 1 0 0
H	1/32	1 1 1 0 1
I	1/32	1 1 1 1
J	1/32	1 1 1 1

4. Fano Code.

Example 1:

2. divide the symbols into two equally probable groups, as possible as you can

Symbol	Probability	Fano Code
A	1/4	0 0
B	1/4	0 1
C	1/8	1 0 0
D	1/8	1 0 1
E	1/16	1 1 0 0
F	1/16	1 1 0 1
G	1/32	1 1 1 0 0
H	1/32	1 1 1 0 1
I	1/32	1 1 1 1
J	1/32	1 1 1 1

4. Fano Code.

Example 1:

3. each group receives one of the binary symbols (i.e. 0 or 1) as the first symbol

Symbol	Probability	Fano Code
A	1/4	0 0
B	1/4	0 1
C	1/8	1 0 0
D	1/8	1 0 1
E	1/16	1 1 0 0
F	1/16	1 1 0 1
G	1/32	1 1 1 0 0
H	1/32	1 1 1 0 1
I	1/32	1 1 1 1 0
J	1/32	1 1 1 1 1

4. Fano Code.

Example 1:

5. stop when no more groups to divide

Symbol	Probability	Fano Code
A	1/4	0 0
B	1/4	0 1
C	1/8	1 0 0
D	1/8	1 0 1
E	1/16	1 1 0 0
F	1/16	1 1 0 1
G	1/32	1 1 1 0 0
H	1/32	1 1 1 0 1
I	1/32	1 1 1 1 0
J	1/32	1 1 1 1 1

4. Fano Code.

Note that: **If it was not possible to divide precisely the probabilities into equally probable groups, we should try to make the division as good as possible, as we can see from the following example.**

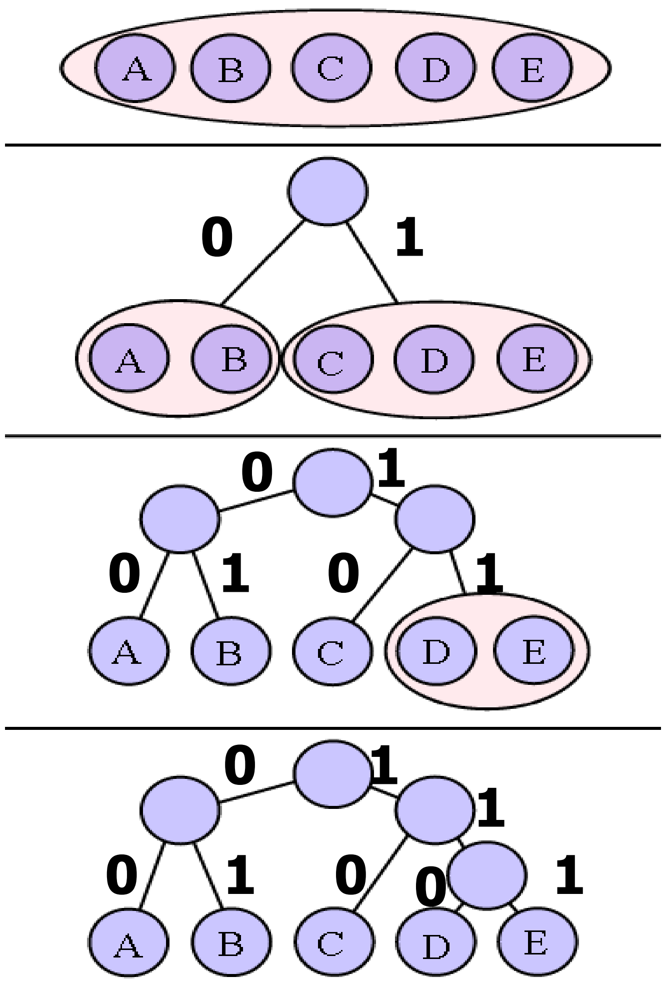
Example 2:

Symbol	Probability	Fano Code
T	**1/3**	**0 0**
U	**1/3**	**0 1**
V	**1/9**	**1 0**
W	**1/9**	**1 1 0**
X	**1/9**	**1 1 1**

4. Fano Code.

Example 3: **Tree form:**

Symbol	Probability	Fano Code
A	0.39	00
B	0.18	01
C	0.15	10
D	0.15	110
E	0.13	111



Source Coding Techniques

1. Huffman Code.

2. Two-path Huffman Code.

3. Lemple-Ziv Code.

4. Fano Code.

5. Shannon Code.

6. Arithmetic Code.

5. Shannon Code.

The Shannon code is the coding method used to prove Shannon's noiseless coding theorem

It constructs a uniquely decodable code

It is suboptimal in the sense that it does not achieve the lowest possible expected code word length like Huffman coding.

5. Shannon Code.

The Shannon code is performed as follows:

1. calculate a series of cumulative probabilities $q_k = \sum_{i=1}^{k} p_i$ **, k=1,2,…,n**

2. calculate the code length for each symbol using $\log(1 \backslash p_i) \le l_i < \log(1 \backslash p_i) + 1$

3. write q_k **in the form** $c_1 2^{-1} + c_2 2^{-2} + \ldots + c_{li} 2^{-li}$ **where each** c_i **is either 0 or 1**

5. Shannon Code.

Example 4:

1. calculate a series of cumulative probabilities

Symbol	Probability	q_k	Length l_i	Shannon Code
A	1/4 +	0		
B	1/4 +	1/4		
C	1/8 +	1/2		
D	1/8 +	5/8		
E	1/16 +	3/4		
F	1/16 +	13/16		
G	1/32 +	7/8		
H	1/32 +	29/32		
I	1/32 +	15/16		
J	1/32	31/32		

5. Shannon Code.

Example 4:

2. calculate the code length for each symbol using

$\log(1\backslash p_i) \le l_i < \log(1\backslash p_i) + 1$

Symbol	Probability	q_k	Length l_i	Shannon Code
A	1/4	0		
B	1/4	1/4		
C	1/8	1/2		
D	1/8	5/8		
E	1/16	3/4		
F	1/16	13/16		
G	1/32	7/8		
H	1/32	29/32		
I	1/32	15/16		
J	1/32	31/32		

5. Shannon Code.

Example 4:

2. calculate the code length for each symbol using

$\log(1\backslash p_i) \leq l_i < \log(1\backslash p_i) + 1$

$\text{Log}(1/(1/4)) \leq l_1 < \text{Log}(1/(1/4)) + 1$

$2 \leq l_1 < 2 + 1$

Symbol	Probability	q_k	l	
A	1/4	0		
B	1/4	1/4		
C	1/8	1/2		
D	1/8	5/8		
E	1/16	3/4		
F	1/16	13/16		
G	1/32	7/8		
H	1/32	29/32		
I	1/32	15/16		
J	1/32	31/32		

5. Shannon Code.

Example 4:

2. calculate the code length for each symbol using

$\log(1\backslash p_i) \leq l_i < \log(1\backslash p_i) + 1$

$\text{Log}(1/(1/4)) \leq l_1 < \text{Log}(1/(1/4)) + 1$

$2 \leq l_1 < 2 + 1$

$l_1 = 2$

Symbol	Probability	q_k	Length	
A	1/4	0	2	
B	1/4	1/4		
C	1/8	1/2		
D	1/8	5/8		
E	1/16	3/4		
F	1/16	13/16		
G	1/32	7/8		
H	1/32	29/32		
I	1/32	15/16		
J	1/32	31/32		

5. Shannon Code.

Example 4:

2. calculate the code length for each symbol using

$\log(1\backslash p_i) \le l_i < \log(1\backslash p_i) + 1$

Symbol	Probability	q_k	Length l_i	Shannon Code
A	1/4	0	2	
B	1/4	1/4	2	
C	1/8	1/2	3	
D	1/8	5/8	3	
E	1/16	3/4	4	
F	1/16	13/16	4	
G	1/32	7/8	5	
H	1/32	29/32	5	
I	1/32	15/16	5	
J	1/32	31/32	5	

5. Shannon Code.

Example 3:

3. write q_k in the form $c_1 2^{-1} + c_2 2^{-2} + \dots + c_{l_i} 2^{-l_i}$ where each c_i is either 0 or 1

Symbol	Probability	q_k	Length l_i	Shannon Code
A	1/4	0	2	
B	1/4	1/4	2	
C	1/8	1/2	3	
D	1/8	5/8	3	
E	1/16	3/4	4	
F	1/16	13/16	4	
G	1/32	7/8	5	
H	1/32	29/32	5	
I	1/32	15/16	5	
J	1/32	31/32	5	

5. Shannon Code.

Example 4:

3. write q_k in the form $c_1 2^{-1} + c_2 2^{-2} + \ldots + c_{l_i} 2^{-l_i}$ | where each c_i is either 0 or 1

$$q_k \le c_1 2^{-1} + c_2 2^{-2}$$

Symbol	Probability	q_k	Length	
A	1/4	0	2	
B	1/4	1/4	2	
C	1/8	1/2	3	
D	1/8	5/8	3	
E	1/16	3/4	4	
F	1/16	13/16	4	
G	1/32	7/8	5	
H	1/32	29/32	5	
I	1/32	15/16	5	
J	1/32	31/32	5	

5. Shannon Code.

Example 4:

3. write q_k in the form $c_1 2^{-1} + c_2 2^{-2} + \dots + c_{l_i} 2^{-l_i}$ | where each c_i is either 0 or 1

Symbol	Probability	q_k	Length	
A	1/4	0	2	
B	1/4	1/4	2	
C	1/8	1/2	3	
D	1/8	5/8	3	
E	1/16	3/4	4	
F	1/16	13/16	4	
G	1/32	7/8	5	
H	1/32	29/32	5	
I	1/32	15/16	5	
J	1/32	31/32	5	

$$q_k \le c_1 2^{-1} + c_2 2^{-2}$$
$$0 \le c_1 2^{-1} + c_2 2^{-2}$$

5. Shannon Code.

Example 4:

3. write q_k in the form $c_1 2^{-1} + c_2 2^{-2} + \ldots + c_{l_i} 2^{-l_i}$ | where each c_i is either 0 or 1

$$q_k \le c_1 2^{-1} + c_2 2^{-2}$$

$$0 \le c_1 2^{-1} + c_2 2^{-2}$$

$$c_1 = 0, \quad c_2 = 0$$

Symbol	Probability	q_k	Length	
A	1/4	0	2	
B	1/4	1/4	2	
C	1/8	1/2	3	
D	1/8	5/8	3	
E	1/16	3/4	4	
F	1/16	13/16	4	
G	1/32	7/8	5	
H	1/32	29/32	5	
I	1/32	15/16	5	
J	1/32	31/32	5	

5. Shannon Code.

Example 4:

3. write q_k in the form $c_1 2^{-1} + c_2 2^{-2} + \ldots + c_{l_i} 2^{-l_i}$ | where each c_i is either 0 or 1

Symbol	Probability	q_k	Length l_i	Shannon Code
A	1/4	0	2	00
B	1/4	1/4	2	
C	1/8	1/2	3	
D	1/8	5/8	3	
E	1/16	3/4	4	
F	1/16	13/16	4	
G	1/32	7/8	5	
H	1/32	29/32	5	
I	1/32	15/16	5	
J	1/32	31/32	5	

$$q_k \leq c_1 2^{-1} + c_2 2^{-2}$$
$$0 \leq c_1 2^{-1} + c_2 2^{-2}$$

$$c_1 = 0, \quad c_2 = 0$$

5. Shannon Code.

Example 4:

3. write q_k in the form $c_1 2^{-1} + c_2 2^{-2} + \ldots + c_{l_i} 2^{-l_i}$	where each c_i is either 0 or 1

Symbol	Probability	q_k	Length l_i	Shannon Code
A	1/4	0	2	00
B	1/4	1/4	2	01
C	1/8	1/2	3	100
D	1/8	5/8	3	101
E	1/16	3/4	4	1100
F	1/16	13/16	4	1101
G	1/32	7/8	5	11100
H	1/32	29/32	5	11101
I	1/32	15/16	5	11110
J	1/32	31/32	5	11111

5. Shannon Code.

Example 4:

Symbol	Probability	q_k	Length l_i	Shannon Code
A	1/4	0	2	00
B	1/4	1/4	2	01
C	1/8	1/2	3	100
D	1/8	5/8	3	101
E	1/16	3/4	4	1100
F	1/16	13/16	4	1101
G	1/32	7/8	5	11100
H	1/32	29/32	5	11101
I	1/32	15/16	5	11110
J	1/32	31/32	5	11111

5. Shannon Code.

Note that:

from examples 1 and 4 one may conclude that Fano coding and Shannon coding produce the same code, however this is not true in general as we can see from the following example.

Example

Symbol	Probability	q_k	Length l_i	Shannon Code	Fano code
W	0.4	0	2	00	0
X	0.3	0.4	2	01	10
Y	0.2	0.7	3	101	110
Z	0.1	0.9	4	1110	111