Playing with Signals:
Compressing Source Symbols

Sibi Raj B Pillai
srbpteach@gmail
Subject:EE113-RollNo

January 10, 2022

Information Entropy

1	3	5	7	9
11	13	15	17	19
21	23	25	27	29
31	33	35	37	39
41	43	45	47	49
51	53	55	57	59

Card 1

2	3	6	7	10
11	14	15	18	19
22	23	26	27	30
31	34	35	38	39
42	43	46	47	50
51	54	55	58	59

Card 2

4	5	6	7	12
13	14	15	20	21
22	23	28	29	30
31	36	37	38	39
44	45	46	47	52
53	54	55	60	♣

Card 3

8	9	10	11	12
13	14	15	24	25
26	27	28	29	30
31	40	41	42	43
44	45	46	47	56
57	58	59	60	♦

Card 4

16	17	18	19	20
21	22	23	24	25
26	27	28	29	30
31	48	49	50	51
52	53	54	55	56
57	58	59	60	♥

Card 5

32	33	34	35	36
37	38	39	40	41
42	43	44	45	46
47	48	49	50	51
52	53	54	55	56
57	58	59	60	♠

Card 6

	6	5	4	3	2	1
$X =$						

Information Entropy

1	3	5	7	9
11	13	15	17	19
21	23	25	$\boxed{27}$	29
31	33	35	37	39
41	43	45	47	49
51	53	55	57	59

Card 1

2	3	6	7	10
11	14	15	18	19
22	23	26	27	30
31	34	35	38	39
42	43	46	47	50
51	54	55	58	59

Card 2

4	5	6	7	12
13	14	15	20	21
22	23	28	29	30
31	36	37	38	39
44	45	46	47	52
53	54	55	60	♣

Card 3

8	9	10	11	12
13	14	15	24	25
26	27	28	29	30
31	40	41	42	43
44	45	46	47	56
57	58	59	60	♦

Card 4

16	17	18	19	20
21	22	23	24	25
26	27	28	29	30
31	48	49	50	51
52	53	54	55	56
57	58	59	60	♥

Card 5

32	33	34	35	36
37	38	39	40	41
42	43	44	45	46
47	48	49	50	51
52	53	54	55	56
57	58	59	60	♠

Card 6

	6	5	4	3	2	1
$X =$						

Information Entropy

1	3	5	7	9
11	13	15	17	19
21	23	25	$\boxed{27}$	29
31	33	35	37	39
41	43	45	47	49
51	53	55	57	59

Card 1

2	3	6	7	10
11	14	15	18	19
22	23	26	$\boxed{27}$	30
31	34	35	38	39
42	43	46	47	50
51	54	55	58	59

Card 2

4	5	6	7	12
13	14	15	20	21
22	23	28	29	30
31	36	37	38	39
44	45	46	47	52
53	54	55	60	♣

Card 3

8	9	10	11	12
13	14	15	24	25
26	$\boxed{27}$	28	29	30
31	40	41	42	43
44	45	46	47	56
57	58	59	60	♦

Card 4

16	17	18	19	20
21	22	23	24	25
26	$\boxed{27}$	28	29	30
31	48	49	50	51
52	53	54	55	56
57	58	59	60	♥

Card 5

32	33	34	35	36
37	38	39	40	41
42	43	44	45	46
47	48	49	50	51
52	53	54	55	56
57	58	59	60	♠

Card 6

$X =$

6	5	4	3	2	1

Information Entropy

1	3	5	7	9
11	13	15	17	19
21	23	25	$\boxed{27}$	29
31	33	35	37	39
41	43	45	47	49
51	53	55	57	59

Card 1

2	3	6	7	10
11	14	15	18	19
22	23	26	$\boxed{27}$	30
31	34	35	38	39
42	43	46	47	50
51	54	55	58	59

Card 2

4	5	6	7	12
13	14	15	20	21
22	23	28	29	30
31	36	37	38	39
44	45	46	47	52
53	54	55	60	♣

Card 3

8	9	10	11	12
13	14	15	24	25
26	$\boxed{27}$	28	29	30
31	40	41	42	43
44	45	46	47	56
57	58	59	60	♦

Card 4

16	17	18	19	20
21	22	23	24	25
26	$\boxed{27}$	28	29	30
31	48	49	50	51
52	53	54	55	56
57	58	59	60	♥

Card 5

32	33	34	35	36
37	38	39	40	41
42	43	44	45	46
47	48	49	50	51
52	53	54	55	56
57	58	59	60	♠

Card 6

$$X = \begin{array}{cccccc} 6 & 5 & 4 & 3 & 2 & 1 \\ \hline \multicolumn{1}{|c|}{0} & \multicolumn{1}{c|}{1} & \multicolumn{1}{c|}{1} & \multicolumn{1}{c|}{0} & \multicolumn{1}{c|}{1} & \multicolumn{1}{c|}{1} \\ \hline \end{array}$$

Information Entropy

Card 1

1	3	5	7	9
11	13	15	17	19
21	23	25	$\boxed{27}$	29
31	33	35	37	39
41	43	45	47	49
51	53	55	57	59

Card 2

2	3	6	7	10
11	14	15	18	19
22	23	26	$\boxed{27}$	30
31	34	35	38	39
42	43	46	47	50
51	54	55	58	59

Card 3

4	5	6	7	12
13	14	15	20	21
22	23	28	29	30
31	36	37	38	39
44	45	46	47	52
53	54	55	60	♣

Card 4

8	9	10	11	12
13	14	15	24	25
26	$\boxed{27}$	28	29	30
31	40	41	42	43
44	45	46	47	56
57	58	59	60	♦

Card 5

16	17	18	19	20
21	22	23	24	25
26	$\boxed{27}$	28	29	30
31	48	49	50	51
52	53	54	55	56
57	58	59	60	♥

Card 6

32	33	34	35	36
37	38	39	40	41
42	43	44	45	46
47	48	49	50	51
52	53	54	55	56
57	58	59	60	♠

$$X = \begin{array}{cccccc} 6 & 5 & 4 & 3 & 2 & 1 \\ \hline \boxed{0} & \boxed{1} & \boxed{1} & \boxed{0} & \boxed{1} & \boxed{1} \end{array}$$

Information: $H(X) \leq \log_2 |\mathcal{X}|$ bits

Compressing Sequences

1 0 0 0 0 0 0 0 0 0 0 0 0 0 0 0 0 0 1 0 0 0 0 1 0 0 0 0 0 0 0 0 0 0 1 0 0
0 0 0 1 0 0 0 0 0 0 0 0 0 0 0 0 0 0 0 0 0 0 1 0 0 0 0 0 0 0 0 0 0 0 0 0 0
0 0 0 0 0 0 0 0 0 0 1 0 0 0 0 0 0 0 0 0 0 0 0 0 0 0 0 0 0 0 0 0 0 0 0 0 0
0 0 0 0 1 0 1 0 0 0 0 0 0 0 1 1 1 0 0 0 0 0 0 1 0 0 0 0 0 0 0 0 0 0 0 0 0
0 0 0 0 1 0 0 0 0 0 0 0 0 0 0 0 0 0 0 0 1 0 0 0 0 0 0 0 0 0 0 0 0 0 0 0 1
0 0 0 0 1 0 0 0 0 0 0 0 0 0 0

Compressing Sequences

1 0 0 0 0 0 0 0 0 0 0 0 0 0 0 0 0 0 1 0 0 0 0 1 0 0 0 0 0 0 0 0 0 0 1 0 0
0 0 0 1 0 0 0 0 0 0 0 0 0 0 0 0 0 0 0 0 0 0 1 0 0 0 0 0 0 0 0 0 0 0 0 0 0
0 0 0 0 0 0 0 0 0 0 1 0 0 0 0 0 0 0 0 0 0 0 0 0 0 0 0 0 0 0 0 0 0 0 0 0 0
0 0 0 0 1 0 1 0 0 0 0 0 0 0 1 1 1 0 0 0 0 0 0 1 0 0 0 0 0 0 0 0 0 0 0 0 0
0 0 0 0 1 0 0 0 0 0 0 0 0 0 0 0 0 0 0 0 1 0 0 0 0 0 0 0 0 0 0 0 0 0 0 0 1
0 0 0 0 1 0 0 0 0 0 0 0 0 0 0

17 4 10 5 18 24 30 1 7 0 0 6 17 15 15 4 10 (**Run Length Coding**)

Compressing Sequences

1 0 0 0 0 0 0 0 0 0 0 0 0 0 0 0 0 0 1 0 0 0 0 1 0 0 0 0 0 0 0 0 0 0 1 0 0
0 0 0 1 0 0 0 0 0 0 0 0 0 0 0 0 0 0 0 0 0 0 1 0 0 0 0 0 0 0 0 0 0 0 0 0 0
0 0 0 0 0 0 0 0 0 0 1 0 0 0 0 0 0 0 0 0 0 0 0 0 0 0 0 0 0 0 0 0 0 0 0 0 0
0 0 0 0 1 0 1 0 0 0 0 0 0 0 1 1 1 0 0 0 0 0 0 1 0 0 0 0 0 0 0 0 0 0 0 0 0
0 0 0 0 1 0 0 0 0 0 0 0 0 0 0 0 0 0 0 0 1 0 0 0 0 0 0 0 0 0 0 0 0 0 0 0 1
0 0 0 0 1 0 0 0 0 0 0 0 0 0 0

17 4 10 5 18 24 30 1 7 0 0 6 17 15 15 4 10 **(Run Length Coding)**

1 0 1 0 0 0 1 1 0 0 0 0 0 0 1 0 0 1 0 0 1 1 0 0 0 0 0 1 0 1 0 1 0 0 1 0 1
1 0 0 0 1 1 0 0 0 0 1 1 1 1 0 0 0 0 0 0 1 0 0 0 0 0 1 0 0 0 0 0 0 0 0 0 0
1 0 1 1 1 0 0 0 0 0 0 0 1 0 1 1 0 0 0 0 0 0 0 0 0 1 0 0 1 0 1 0 0 0 1 0 1
0 0 0 0 1 1 0 0 0 1 0 0 0 0 1 0 0 0 0 1 0 1 1 1 1 0 0 0 1 1 1 1 0 0 0 0 0
1 0 0 0 1 1 0 0 0 1 0 0 0 0 0 1 0 0 0 0 0 0 0 0 0 0 1 0 1 0 0 0 0 0 0 1 0
0 0 1 1 0 0 0 0 0 1 0 0 0 0 0

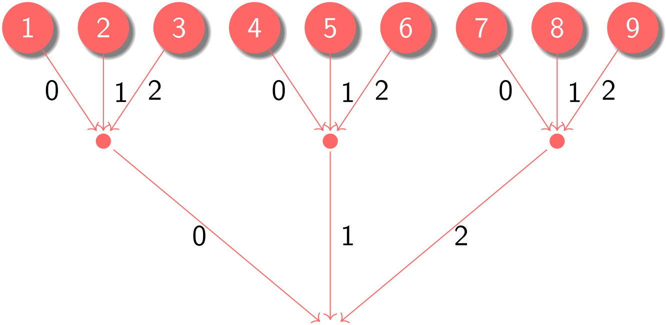
1 3 0 6 2 2 0 5 1 1 2 1 0 3 0 4 0 0 0 6 5 10 1 0 0 7 1 0 9 2 1 3 1 4
0 3 4 4 1 0 0 0 3 0 0 0 5 3 0 3 5 10 1 6 3 0 5 (Run Length Coding)

9 Ball Game

Suppose there are 9 balls, look alike, but one of them is heavier than the rest (GOLD!). With two weighings (measurements) on a common balance, can you identify the odd one.

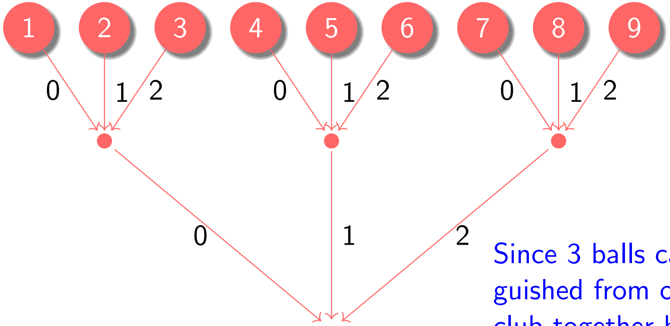
9 Ball Game

Suppose there are 9 balls, look alike, but one of them is heavier than the rest (GOLD!). With two weighings (measurements) on a common balance, can you identify the odd one.



9 Ball Game

Suppose there are 9 balls, look alike, but one of them is heavier than the rest (GOLD!). With two weighings (measurements) on a common balance, can you identify the odd one.

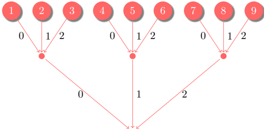


Since 3 balls can be distinguished from one test, we club together blocks of 3

Balls and Sources

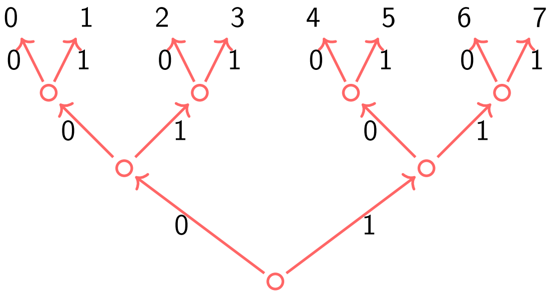
- Suppose we repeatedly perform the first experiment, using a statistical machine that shuffles the golden ball
- The random variable representing the output of the machine is called a *source*.
- Every time a source symbol $S_i \in \mathcal{X}$ occurs, we will convey its branch labels.

Question: For a given source and a label-alphabet,
what is the **optimal** tree?



Binary Number System

- A binary tree representation for numbers.



- If the source is fair, this indeed is the **optimal** tree.
- This also gives the simple principle

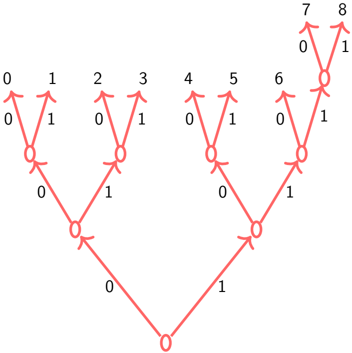
$$\mathit{Info}(X) \leq \log_D |\mathcal{X}| + 1 \quad \text{DiTs}$$

Gimme another Drop

- Suppose $\mathcal{X} = \{0, 1, \cdots, 8\}$ with $p_0 \geq p_1 \geq \cdots \geq p_8$.

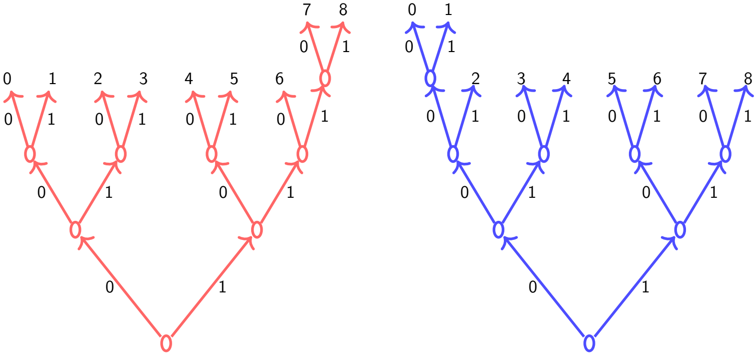
Gimme another Drop

- Suppose $\mathcal{X} = \{0, 1, \cdots, 8\}$ with $p_0 \geq p_1 \geq \cdots \geq p_8$.



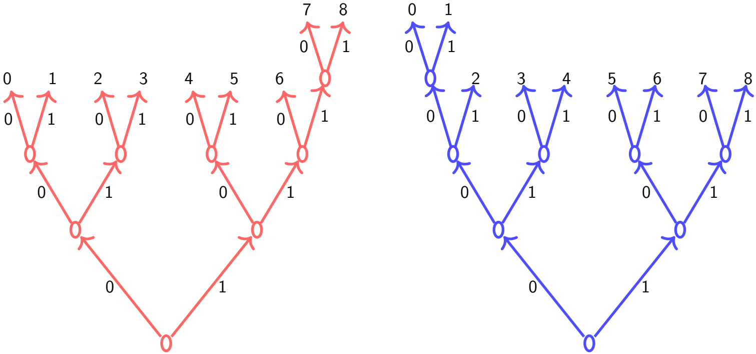
Gimme another Drop

- Suppose $\mathcal{X} = \{0, 1, \cdots, 8\}$ with $p_0 \geq p_1 \geq \cdots \geq p_8$.



Gimme another Drop

- Suppose $\mathcal{X} = \{0, 1, \cdots, 8\}$ with $p_0 \geq p_1 \geq \cdots \geq p_8$.



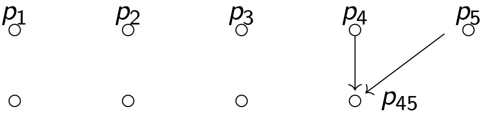
- '*Shorter codes to more frequent symbols*' seems to be the key to compression, we seek the best code tree for a given probability distribution on the symbols.

Binary Huffman Example

Let $p_1 = 0.47, p_2 = 0.18, p_3 = 0.15, p_4 = 0.1, p_5 = 0.1$ and $p_{ij} \triangleq p_i + p_j$.

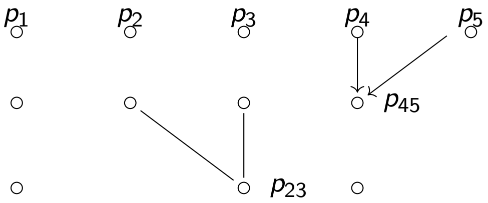
Binary Huffman Example

Let $p_1 = 0.47, p_2 = 0.18, p_3 = 0.15, p_4 = 0.1, p_5 = 0.1$ and $p_{ij} \triangleq p_i + p_j$.



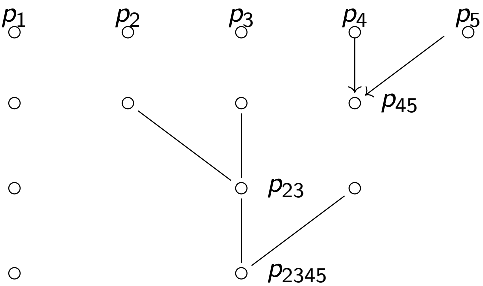
Binary Huffman Example

Let $p_1 = 0.47, p_2 = 0.18, p_3 = 0.15, p_4 = 0.1, p_5 = 0.1$ and $p_{ij} \triangleq p_i + p_j$.



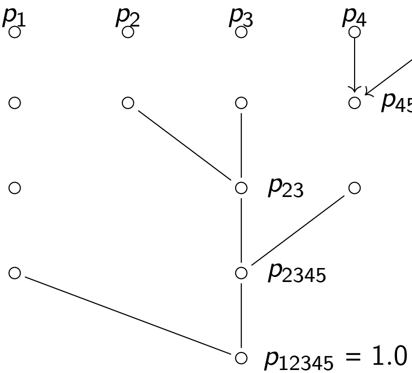
Binary Huffman Example

Let $p_1 = 0.47, p_2 = 0.18, p_3 = 0.15, p_4 = 0.1, p_5 = 0.1$ and $p_{ij} \triangleq p_i + p_j$.



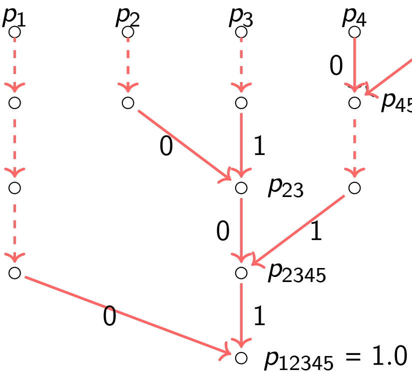
Binary Huffman Example

Let $p_1 = 0.47, p_2 = 0.18, p_3 = 0.15, p_4 = 0.1, p_5 = 0.1$ and $p_{ij} \triangleq p_i + p_j$.



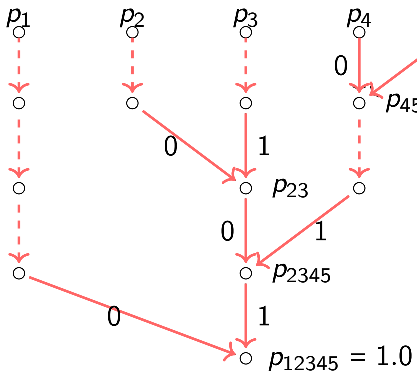
Binary Huffman Example

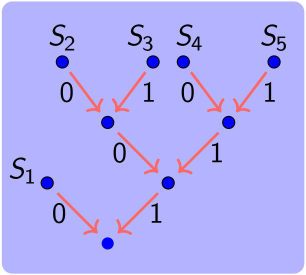
Let $p_1 = 0.47, p_2 = 0.18, p_3 = 0.15, p_4 = 0.1, p_5 = 0.1$ and $p_{ij} \triangleq p_i + p_j$.



Binary Huffman Example

Let $p_1 = 0.47, p_2 = 0.18, p_3 = 0.15, p_4 = 0.1, p_5 = 0.1$ and $p_{ij} \triangleq p_i + p_j$.





Collapse (delete) the dashed lines to get the highlighted tree.

Ternary Huffman Code

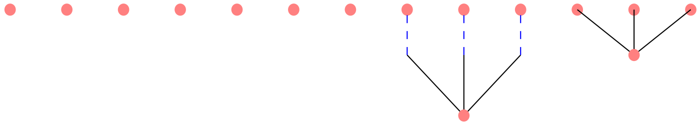
Source	S_1	S_2	S_3	S_4	S_5	S_6	S_7	S_8	S_9	S_{10}	S_{11}	S_{12}	S_{13}
Pr(S)	$\frac{1}{9}$	$\frac{1}{9}$	$\frac{1}{9}$	$\frac{1}{9}$	$\frac{1}{9}$	$\frac{1}{9}$	$\frac{1}{9}$	$\frac{1}{27}$	$\frac{1}{27}$	$\frac{1}{27}$	$\frac{1}{27}$	$\frac{1}{27}$	$\frac{1}{27}$

Ternary Huffman Code

Source	S_1	S_2	S_3	S_4	S_5	S_6	S_7	S_8	S_9	S_{10}	S_{11}	S_{12}	S_{13}
Pr(S)	$\frac{1}{9}$	$\frac{1}{9}$	$\frac{1}{9}$	$\frac{1}{9}$	$\frac{1}{9}$	$\frac{1}{9}$	$\frac{1}{9}$	$\frac{1}{27}$	$\frac{1}{27}$	$\frac{1}{27}$	$\frac{1}{27}$	$\frac{1}{27}$	$\frac{1}{27}$

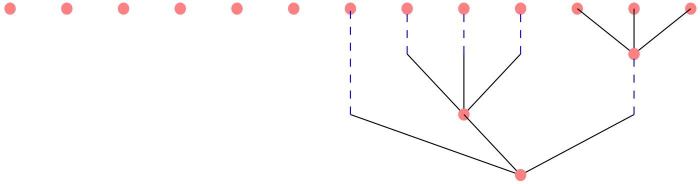
Ternary Huffman Code

Source	S_1	S_2	S_3	S_4	S_5	S_6	S_7	S_8	S_9	S_{10}	S_{11}	S_{12}	S_{13}
Pr(S)	$\frac{1}{9}$	$\frac{1}{9}$	$\frac{1}{9}$	$\frac{1}{9}$	$\frac{1}{9}$	$\frac{1}{9}$	$\frac{1}{9}$	$\frac{1}{27}$	$\frac{1}{27}$	$\frac{1}{27}$	$\frac{1}{27}$	$\frac{1}{27}$	$\frac{1}{27}$



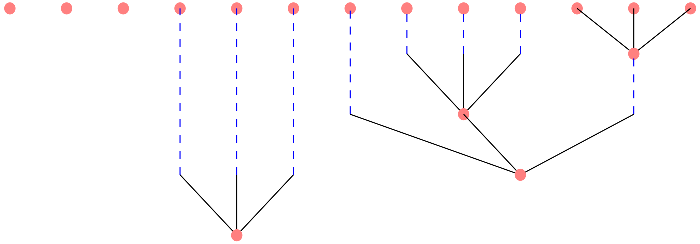
Ternary Huffman Code

Source	S_1	S_2	S_3	S_4	S_5	S_6	S_7	S_8	S_9	S_{10}	S_{11}	S_{12}	S_{13}
Pr(S)	$\frac{1}{9}$	$\frac{1}{9}$	$\frac{1}{9}$	$\frac{1}{9}$	$\frac{1}{9}$	$\frac{1}{9}$	$\frac{1}{9}$	$\frac{1}{27}$	$\frac{1}{27}$	$\frac{1}{27}$	$\frac{1}{27}$	$\frac{1}{27}$	$\frac{1}{27}$



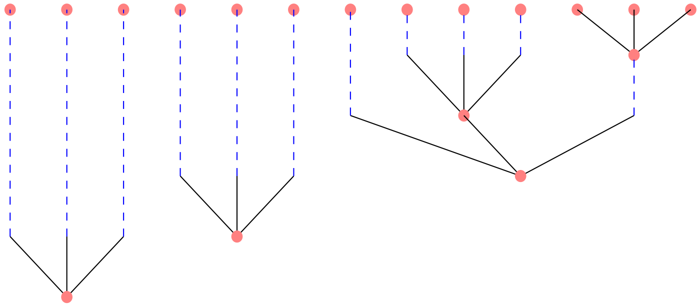
Ternary Huffman Code

Source	S_1	S_2	S_3	S_4	S_5	S_6	S_7	S_8	S_9	S_{10}	S_{11}	S_{12}	S_{13}
Pr(S)	$\frac{1}{9}$	$\frac{1}{9}$	$\frac{1}{9}$	$\frac{1}{9}$	$\frac{1}{9}$	$\frac{1}{9}$	$\frac{1}{9}$	$\frac{1}{27}$	$\frac{1}{27}$	$\frac{1}{27}$	$\frac{1}{27}$	$\frac{1}{27}$	$\frac{1}{27}$



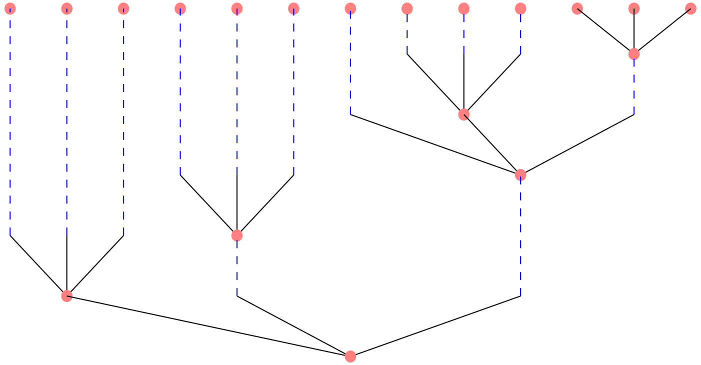
Ternary Huffman Code

Source	S_1	S_2	S_3	S_4	S_5	S_6	S_7	S_8	S_9	S_{10}	S_{11}	S_{12}	S_{13}
Pr(S)	$\frac{1}{9}$	$\frac{1}{9}$	$\frac{1}{9}$	$\frac{1}{9}$	$\frac{1}{9}$	$\frac{1}{9}$	$\frac{1}{9}$	$\frac{1}{27}$	$\frac{1}{27}$	$\frac{1}{27}$	$\frac{1}{27}$	$\frac{1}{27}$	$\frac{1}{27}$



Ternary Huffman Code

Source	S_1	S_2	S_3	S_4	S_5	S_6	S_7	S_8	S_9	S_{10}	S_{11}	S_{12}	S_{13}
Pr(S)	$\frac{1}{9}$	$\frac{1}{9}$	$\frac{1}{9}$	$\frac{1}{9}$	$\frac{1}{9}$	$\frac{1}{9}$	$\frac{1}{9}$	$\frac{1}{27}$	$\frac{1}{27}$	$\frac{1}{27}$	$\frac{1}{27}$	$\frac{1}{27}$	$\frac{1}{27}$



Huffman Coding

- We will describe **Huffman Coding** when $\mathcal{A} = \{0, 1\}$ (binary).
- Let all labels be empty, and let $m = |\mathcal{X}|$.

1. Rearrange sources such that $p_1 \geq p_2 \geq \cdots \geq p_m$.
2. Append labels 0 and 1 respectively to the last two sources.
3. Merge the last two sources to form a new source X'_{m-1}, having probability $p_{m-1} + p_m$.
4. Put $m \leftarrow m - 1$ and go to step 1, using the new source set.

- Terminate by assiging 0 and 1 to the two remaining sources.

Lossy Source Coding (JPEG/MPEG)

- Taking a block of data and apply a sparsifying Transform.
- Throw away the not so relevant values (based on demand).
- Store the remaining small set of values losslessly.
- JPEG uses Discrete Cosine Transfrom (DCT) and zig-zag run length coding to compress by ≈ 30 for similar visual quality.