

- source-coding theorem
- data compaction
- mutual information
- differential entropy

การแปลงที่ map ความเป็น output ของ source ให้เป็น blocks

min: block หนึ่งประกอบด้วย source symbols หนึ่งถึง  $n$  symbols

ในกรณี block ของ  $n$  symbols นี้ extended source

และใช้ alphabet หนึ่งคือ  $\mathcal{Y}^n$  ใช้แทน min: alphabet ของ  $\mathcal{Y}^n$  แทน alphabet ของ  $\mathcal{Y}$  ที่ใช้สื่อสารก่อนหน้านี้

นั่นคือ

$$H(\mathcal{Y}^n) = nH(\mathcal{Y}) \quad (9.12)$$

Ex  $\mathcal{Y} = \{s_0, s_1, s_2\}$  ที่มี prob. ที่สม่ำเสมอ (uniform)

$$p_0 = \frac{1}{4}, p_1 = \frac{1}{4}, p_2 = \frac{1}{2}$$

ในการหา  $H(\mathcal{Y})$  ให้ใช้

$$H(\mathcal{Y}) = \sum_{k=0}^{K-1} p_k \log_2 \left( \frac{1}{p_k} \right) \quad (9.9)$$

$$\begin{aligned}
&= p_0 \log_2 \left( \frac{1}{p_0} \right) + p_1 \log_2 \left( \frac{1}{p_1} \right) + p_2 \log_2 \left( \frac{1}{p_2} \right) \\
&= \frac{1}{4} \log_2 (4) + \frac{1}{4} \log_2 (4) + \frac{1}{2} \log_2 (2) \\
&= \frac{3}{2} \text{ bits / source symbol}
\end{aligned}$$

Lemma  $\lim_{n \rightarrow 0} n \log n = 0$

where  $n = 2^{-n}$

$$\delta_i ; i = 0, \dots, 8$$

$\delta_i \in \mathcal{S}^2$	$s_0 s_0$	$s_0 s_1$	$s_0 s_2$	$s_1 s_0$	$s_1 s_1$	$s_1 s_2$	$s_2 s_0$	$s_2 s_1$	$s_2 s_2$
$\text{Prob}(\cdot)$	$\frac{1}{16}$	$\frac{1}{16}$	$\frac{1}{8}$	$\frac{1}{16}$	$\frac{1}{16}$	$\frac{1}{8}$	$\frac{1}{8}$	$\frac{1}{8}$	$\frac{1}{4}$

110:

$$\begin{aligned}
H(\mathcal{S}^2) &= \sum_{i=0}^8 p(\delta_i) \log_2 \frac{1}{p(\delta_i)} \\
&= 4 \left( \frac{1}{16} \log_2 (16) \right) + 4 \left( \frac{1}{8} \log_2 (8) \right) + \frac{1}{4} \log_2 (4) \\
&= 3 \text{ bits / block of source symbol}
\end{aligned}$$

$$>>> 4 * (1/16) * \log(16, 2) + (4 * 1/8) * \log(8, 2) + (1/4) * \log(4, 2)$$

3

9;wvvi  $2 H(\mathcal{S}) = H(\mathcal{S}^2)$

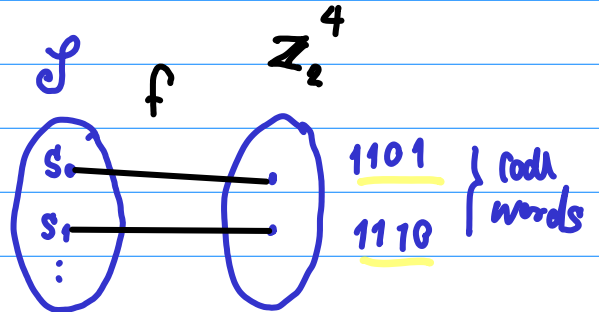
### 9.3 Source-Coding Theorem

Source - Coding ကိုပိုမို **efficient** လုပ်ဆောင်စေရန်  
 နည်းလမ်း: Symbol ၇၄ alphabet ကိုကွဲပြားစွာ ကိုယ်စားပြု  
 လုပ်ဆောင်ရန် ချိတ်ဆက်ပေးရန်

Symbol    ကို ကော်ဒ် ဟုလည်းခေါ်ပြီး    Code    ခေါ်  
             မှတ်    "                      code    ပြန်               } Morse code

11/10/2023 15:10

1 code word in binary form

$$2 \quad f^n \quad 1-1$$


பின் average code - word length,  $\bar{L}$  தீர்வு

$$\vec{L} = \sum_{k=0}^{K-1} p_k \ell_k \quad \text{dim / source symbol} \quad (9.18)$$

15.  $L_k$  number of bits in symbol  $n^k$

ประสิทธิภาพของ source encoder สามารถวัดได้โดยสมการ

$$\eta = \frac{H(\mathcal{X})}{\bar{L}} \quad (9.21)$$

เมื่อให้ source เป็น discrete memoryless source

## 9.4 Data Compaction

โดยทั่วไป physical sources ปรากฏว่ามีความซ้ำซ้อนมากเกิน  
เมื่อที่เรียกว่า **redundant** จึงมักถูกกำจัดทิ้ง! remove

จึงลดจำนวนบิตที่ส่งต่อ

เป็นเหตุให้เกิดกระบวนการ Data Compaction

**prefix code** เป็นหนึ่งใน code ที่ไม่มี code word ใด

เป็น prefix

ม.ว. ของ non-prefix code  $s_0 \rightarrow 0$   
 $s_1 \rightarrow 00$

ที่ใดที่รับจ. ไม่สามารถหา Symbol ได้

prefix Code  $s_0 \rightarrow 0$   
 $s_1 \rightarrow 10$

ที่ใดที่รับจ. สามารถหา Symbol ได้

၁ prefix code သည် စုံလင်သော ကိုယ်စားပြုကုဒ် ဖြစ်သည်

သို့သော် decode လုပ်ရန် symbol ကို အသုံးပြုရသည် (1-1)

prefix code ကို discrete memoryless source ဝန်ဆောင်မှုပေးရန်

အသုံးပြုသော Kraft - McMillan ကိုယ်စားပြုကုဒ်

$$\sum_{k=0}^{K-1} 2^{-l_k} \leq 1 \quad (9.22)$$

prefix code ကို discrete memoryless source သည်

entropy ပုံစံ  $H(\mathcal{X})$  ကို အသုံးပြုရန်

$$H(\mathcal{X}) \leq \bar{L} < H(\mathcal{X}) + 1 \quad (9.23)$$

သို့သော် extended code ကို prefix code ကို အသုံးပြုရန်

$$H(\mathcal{X}^n) \leq \bar{L}_n < H(\mathcal{X}^n) + 1 \quad (9.27)$$

၇၇ (9.17) ကို

$$n H(\mathcal{X}) \leq \bar{L}_n < n H(\mathcal{X}) + 1$$

4. n မှာ မြင့်မား

$$H(\mathcal{P}) \leq \frac{\bar{L}_n}{n} < H(\mathcal{P}) + \frac{1}{n} \quad (9.28)$$

10.:

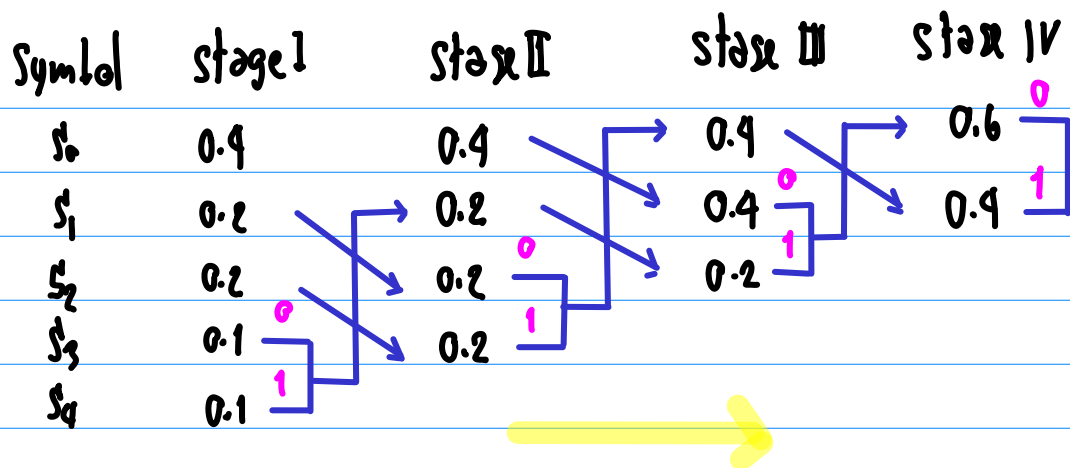
$$\lim_{n \rightarrow \infty} \frac{1}{n} \bar{L}_n = H(\mathcal{P}) \quad (9.29)$$

စာ (9.29) ကို အသုံးပြု၍ ကိုယ်၏ block size,  $n$  ကို ကိန်း  $\frac{\bar{L}_n}{n}$  ကိုယ်၏ (ခန့်မှန်း)  $H(\mathcal{P})$  (ကိုယ်၏ ကုဒ်များ၏ ပျမ်းမျှအလျင်) ဝင်ရောက် )

Huffman Coding  $\bar{L} \rightarrow H(\mathcal{P})$

အလုပ်သမား (algorithm)

1. ပေးထားသော source symbols ကို prob မှာ  $\rightarrow$  prob ပို့
2. Combined two source symbols (ကိုယ်၏ prob ကို သုံးစွဲ)  
ကိုယ်၏ new source symbols ကို ကိုယ်၏ ပြုပြင်မှုမှ ကိုယ်၏  
သော symbol ကိုယ်၏ prob မှာ ကိုယ်၏ ကိုယ်၏ 2 source symbols
3. Assign binary 0 ကို 1



EX

Symbol $\in \mathcal{S}$	Prob.	Code word
$s_0$	0.4	00
$s_1$	0.2	10
$s_2$	0.2	11
$s_3$	0.1	010
$s_4$	0.1	011

$$I_A \quad \bar{L} = \sum_{k=0}^4 p_k l_k$$

$$= 0.4(1) + 0.2(2) + 0.2(2) + 0.1(3) + 0.1(3)$$

$$= 2.2$$

110:

$$H(\mathcal{S}) = \sum_{k=0}^4 p_k \log_2 \left( \frac{1}{p_k} \right)$$

$$= 0.4 \log_2 \left( \frac{1}{0.4} \right) + 0.2 \log_2 \left( \frac{1}{0.2} \right)$$

$$+ 0.2 \log_2 \left( \frac{1}{0.2} \right) + 0.1 \log_2 \left( \frac{1}{0.1} \right) + 0.1 \log_2 \left( \frac{1}{0.1} \right)$$

$$= 2.12193 \quad \text{Lits}$$

$$\gamma_n \quad \eta = \frac{H(\mathcal{L})}{\bar{L}} = \frac{2.12193}{2.2} \approx 0.9693$$