[Cover] Elements of Information Theory-Wiley-Interscience

chapter 2

- 2-1 Coin flips
  - (a)
    $\sum_{n=0}^{\infty} r^n = \frac{1}{1-r}$ $\sum_{n=0}^{\infty} nr^n = \frac{r}{(1-r)^2}$
    $P(X = n) = pq^{n-1}$
    $H(X) = -\sum_{n=1} pq^{n-1}log(pq^{n-1}) = -\sum_{n=0} (pq^n logp + npq^n logq) =$
    $-\frac{plogp+qlogq}{p} = \frac{H(p)}{p}$
    then let p=1/2
  - (b)
    X=1
    - X=2
      - X=3

        ...

    so $\sum_{n=0} n(\frac{1}{2^n}) = 2$
    and $E \geq H(X), (X,Y) \sim (1,1), (2,01, (3,001)...$ Huffman code
- 2-2 Entropy of functions
  for $y = g(x)$,$p(y) = \sum_{x:y=g(x)} p(x)$
  $\sum_{x:y=g(x)} p(x)logp(x) \leq \sum_{x:y=g(x)} p(x)logp(y) = p(y)logp(y)$
  so $H(X) = -\sum_x p(x)logp(x)$
  $= -\sum_y \sum_{x:y=g(x)} p(x)logp(x) \geq -\sum_y p(y)logp(y)$
  $= H(Y)$
  iff Y-X is one-to-one
  - (a)true
  - (b)false so $H(X) \geq H(Y)$
- 2-3 Minimum entropy
  $H(p_1, ..., p_n) = H(p) = -\sum_i p_i logp_i, \mathbf{p} = (p_1, ..., p_n)$
  for $-p_i logp_i \geq 0$ iff $p_i = 0$ or 1
  so those $p_i = 1$ for some i and $p_j = 0, j \neq i$, as $(1,0,0...), (0,1,0,...), (...,0,0,1)$
  the minimum value is 0
- 2-4 Entropy of functions of a random variable
  - (a) $H(X, g(X)) = H(X) + H(g(X)|X)$ chain rule

- (b) $H(g(X)|X) = 0$ since for any X, g(X) is fixed and $H(g(X)|X) = \sum_x p(x)H(g(X)|X = x) = \sum_x 0 = 0$
- (c) $H(X, g(X)) = H(g(X)) + H(X|g(X))$ chain rule
- (d) $H(X|g(X)) \geq 0$ iff X and g(X) is one-to-one, so $H(X, g(X)) \geq H(g(X))$

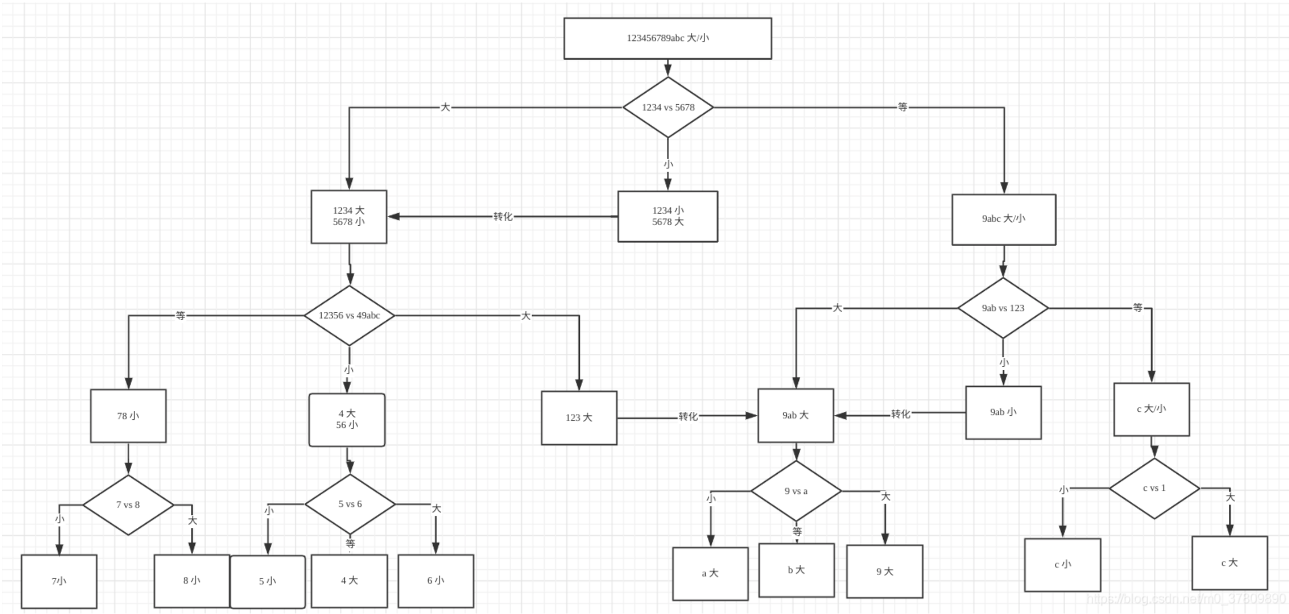
- 2-5 Zero conditional entropy
  assume x_0 and y_1, y_2 i.e. $p(x_0, y1) > 0, p(x_0, y_2) > 0$ and $p(x_0) \geq p(x_0, y_1) + p(x_0, y_2) > 0$ and those are not 0 and 1

  $$H(Y|X) = -\sum_x p(x) \sum_y p(y|x) logp(y|x)$$

  so
  $$\begin{aligned} &\geq p(x_0)(-p(y_1|x_0)logp(y_1|x_0) - p(y_2|x_0)logp(y_2|x_0)) \\ &> 0 \end{aligned}$$

  contracted!
  moreover we can find $H(Y|X) = H(X, Y) - H(X)$
  that is when conditional entropy is equal to zero, X is defined and then Y isn't random so Y is the funciton of X
- 2-6 Conditional mutual information and unconditional mutual information
  - (a) $I(X; Y|Z) < I(X; Y)$
    if $X \rightarrow Y \rightarrow Z$ i.e.
    if $p(x, y|z) = p(x|z)p(y|z)$ then $I(X; Y) \geq I(X; Y|Z)$ equality holds iff $I(X; Z) = 0$
    let $X = Y = Z$ so $I(X; Y) = H(X) - H(X|Y) = H(X) = 1$
    and $I(X; Y|Z) = H(X|Z) - H(X|Y, Z) = 0$
  - (b) $I(X; Y|Z) > I(X; Y)$
    let $Z = X + Y$ and X, Y be independent and binary random variables such as 0 or 1
    then $I(X; Y|Z) = H(X|Z) = \frac{1}{2}, I(X; Y) = 0$
- 2-7 Coin weighing
  - (a) we have 2n+1 possible sotuations i.e.
    - one of the n coins is heavier n states
    - one of the n coins is lighter n states
    - They are all of equal weight 1 states
      basing on the rule, with k weighings, there are 3^k possible outcomes so $2n + 1 \leq 3^k$ or $k \geq \frac{log_2(2n+1)}{log_2 3}$
      which means each weighing gives at most $log_2 3$ and the maximum entropy of $log_2(2n + 1)$ bits so one would require at least $\frac{log_2(2n+1)}{log_2 3}$ weighings
  - (b)
    solution

the reason: the question information entropy is $12 \times 2 = 24$ or $2n + 1 = 25$
with 3 weighing, we can get the information $3^3 = 27$
The information entropy of each weighing should be reduced as much as possible
also: we can
$$\begin{aligned} &min\ k \\ &s.t.\ klog_2 3 \geq log_2(2n) \end{aligned}$$



more: we use $\{-12, -11, ..., 0, .., 11, 12\}$ and the state $-1, 0, 1$ to express theses number such as $8 = (-1, 0, 1) = -1 \times 3^0 + 0 \times 3^1 + 1 \times 3^2$

- on left pan when $n_i = -1$
- aside when $n_i = 0$
- on right pan when $n_i = 1$

Then the three weighings give the ternary expansion of the index of the odd coin. If the expansion is the same as the expansion in the matrix, it indicates that the coin is heavier. If the expansion is of the opposite sign, the coin is lighter
for example $(0, 1, 1)$ indicates 12 so the coin #12 is heavy
$(1, 0, -1)$ indicates -8 so the coin #8 is light $(0, 0, 0)$ indicates no odd coin
and we can try by coding called Reduce and Conquer Method.
or we can try by Hamming code method

- 2-8 Drawing with and without replacement
  r red w white b black balls
  - With replacement

$$X_i = \begin{cases} red & with\ prob. \frac{r}{r+w+b} \\ white & with\ prob. \frac{w}{r+w+b} \\ black & with\ prob. \frac{b}{r+w+b} \end{cases}$$

so $H(X_i|X_{i-1}, ..., X_1) = H(X_i) = -\frac{r}{r+w+b} log \frac{r}{r+w+b} - \frac{w}{r+w+b} log \frac{w}{r+w+b} - \frac{b}{r+w+b} log \frac{b}{r+w+b}$

- Without replacement
  the **conditional probalility** of the i-th ball being red is $\frac{r}{r+w+b}$
  so the unconditional entropy $H(X_i)$ is the same as with replacement
  but the conditional entropy $H(X_i|X_{i-1}, ..., X_1)$ is less than the unconditional entropy