EE 231A Shengze Ye 205418959 HW2

1. Relative Entropy AEP

(a) $\lim_{n\to\infty} -\frac{1}{n}\log q(X_1, X_2 \cdots, X_n)$, where $X_1, X_2 \cdots$ are i.i.d $\sim p(x)$

$\because X_1, X_2 \cdots X_n$ are i.i.d

$$\lim_{n\to\infty} -\frac{1}{n}\log q(X_1, X_2 \cdots X_n) = \lim_{n\to\infty} -\frac{1}{n}\log q(X_1)q(X_2)\cdots q(X_n)$$

$$= \lim_{n\to\infty} -\frac{1}{n}\sum_{i=1}^{n}\log q(X_i)$$

$$= -E_p \log q(X) \quad \text{convergence in probability}$$

$$= -\sum_{i=1}^{n} p(x)\log q(x)$$

$$D(p\|q) = \sum_{i=1}^{n} p(x)\log\frac{p(x)}{q(x)} = \sum_{i=1}^{n} p(x)\log p(x) - \sum_{i=1}^{n} p(x)\log q(x)$$

$$\therefore D(p\|q) - \sum_{i=1}^{n} p(x)\log p(x) = -\sum_{i=1}^{n} p(x)\log q(x)$$

$$\therefore D(p\|q) + H(p) = -\sum_{i=1}^{n} p(x)\log q(x)$$

$\therefore \lim_{n\to\infty} -\frac{1}{n}\log q(X_1, X_2 \cdots X_n)$ can be represented by $D(p\|q) + H(p)$

(b) evaluate the limit of $-\frac{1}{n}\log\frac{q(X_1 \cdots X_n)}{p(X_1 \cdots X_n)}$ when $X_1, X_2 \cdots X_n$ are i.i.d $\sim p(x)$

$$\lim_{n\to\infty} -\frac{1}{n}\log\frac{q(x_1,\cdots x_n)}{p(x_1,\cdots,x_n)} = \lim_{n\to\infty} -\frac{1}{n}\sum_{i=1}^{n}\log\frac{q(x_i)}{p(x_i)}$$

$$= -E_p \log\frac{q(x)}{p(x)} \quad \text{convergence in probability}$$

$$= -\sum_{i=1}^{n} p(x)\log\frac{q(x)}{p(x)} = \sum_{i=1}^{n} p(x)\log\frac{p(x)}{q(x)}$$

$$= D(p\|q)$$

$\therefore$ The limit is $D(p\|q)$

2. Take the limit

$$p(x^n, y^n) = \prod_{i=1}^{n} p(x_i, y_i)$$

$$\lim_{n\to\infty} \frac{1}{n}\log\frac{p(x^n,y^n)}{p(x^n)\,p(y^n)} = \lim_{n\to\infty}\frac{1}{n}\sum_{i=1}^{n}\log\frac{p(x_i,y_i)}{p(x_i)\,p(y_i)}$$

$$= E_{p(x,y)}\log\frac{p(x,y)}{p(x)\,p(y)} \quad \text{convergence in probability}$$

$$= \sum_{i=1}^{n} p(x_i,y_i)\log\frac{p(x_i,y_i)}{p(x_i)\,p(y_i)} = I(X;Y)$$

$\therefore$ The limit is the mutual information $I(X;Y)$

3. The AEP in action

Consider a Bernouli random variable X with $p(1) = 3/4$, $p(0) = 1/4$

(a) Show that $H(X) = 2 - \frac{3}{4}\log 3$

$$H(x) = -\sum_i p(x_i)\log p(x_i)$$

$$= -\frac{3}{4}\log\frac{3}{4} - \frac{1}{4}\log\frac{1}{4} = -\frac{3}{4}[\log 3 - \log 4] + \frac{1}{4}\log 4$$

$$= \frac{3}{2} - \frac{3}{4}\log 3 + \frac{1}{2} = 2 - \frac{3}{4}\log 3$$

(b) show that

$$-\frac{1}{n}\log p(x_1, x_2 \cdots x_n) = 2 - \frac{k}{n}\log 3$$

$x_1, x_2 \cdots x_n$ are i.i.d k is the number of ones

$$-\frac{1}{n}\log p(x_1, x_2, \cdots x_n) = -\frac{1}{n}\log p(x_1)\, p(x_2) \cdots p(x_n)$$

$$= -\frac{1}{n}\log\left(\frac{3}{4}\right)^k\left(\frac{1}{4}\right)^{n-k} = -\frac{k}{n}\log\frac{3}{4} - \frac{n-k}{n}\log\frac{1}{4}$$

$$= -\frac{k}{n}[\log 3 - \log 4] - \frac{n-k}{n}(-2)$$

$$= -\frac{k}{n}\log 3 + \frac{k}{n}\times 2 + 2 - \frac{k}{n}\times 2$$

$$= 2 - \frac{k}{n}\log 3$$

(c) Compute $Pr(A_\varepsilon^{(n)})$ and $|A_\varepsilon^{(n)}|$ for $n=8$, $\varepsilon=0.2$

we know from Theorem 3.1.2 that

$$H(X) - \varepsilon \le -\frac{1}{n}\log p(x_1, x_2 \cdots x_n) \le H(X) + \varepsilon$$

if $(x_1, x_2 \cdots x_n) \in A_\varepsilon^{(n)}$ (typical set)

$H(X) = 2 - \frac{3}{4}\log 3 = 0.8113$ $\varepsilon = 0.2$, $n=8$

$\therefore 0.8113 - 0.2 \le 2 - \frac{k}{8}\log_2 3 \le 0.8113 + 0.2$

$\Rightarrow \therefore$ $k \le n(2 - 0.6113)/\log_2 3$

$k \ge n(2 - 1.0113)/\log_2 3$

$\therefore$ $k \le 7.0094$ $k \ge 4.9904$

$\therefore$ k could be 5, 6, 7

for $n=8$, the sequence is a typical sequence if there are 5, 6, or 7 ones in the sequence

$$\therefore Pr(A_\varepsilon^{(n)}) = C_8^5 \left(\frac{1}{4}\right)^3 \left(\frac{3}{4}\right)^5 + C_8^6 \left(\frac{1}{4}\right)^2 \left(\frac{3}{4}\right)^6 + C_8^7 \left(\frac{1}{4}\right)\left(\frac{3}{4}\right)^7$$

$$= \frac{1}{65536}[56\times 243 + 28\times 729 + 8\times 2187] = \frac{51516}{65536}$$

$= 0.7861$ $\therefore P(A_\varepsilon^{(n)}) = 0.7861$

$$|A_\varepsilon^{(n)}| = C_8^5 + C_8^6 + C_8^7 = 56 + 28 + 8 = 92$$

(d) Repeat for $n=16$ and $\varepsilon = 0.2$

Same with (c)

$$H(X) - \varepsilon \le 2 - \frac{k}{n}\log_2 3 \le H(X) + \varepsilon$$

$$\therefore \quad 2 - 0.6113 \le 2 - \frac{k}{16}\log_2 3 \le 2 - 1.0113$$

$$\Rightarrow \quad k \le 16(2 - 0.6113)/\log_2 3$$

$$k \ge 16(2 - 1.0113)/\log_2 3$$

$$k \le 14.0188 \qquad k \ge 9.9808$$

k could be $10, 11, 12, 13, 14$

$$\therefore Pr(A_\varepsilon^{(n)}) = C_{16}^{10}\left(\frac{1}{4}\right)^6\left(\frac{3}{4}\right)^{10} + C_{16}^{11}\left(\frac{1}{4}\right)^5\left(\frac{3}{4}\right)^{11}$$

$$+ C_{16}^{12}\left(\frac{1}{4}\right)^4\left(\frac{3}{4}\right)^{12} + C_{16}^{13}\left(\frac{1}{4}\right)^3\left(\frac{3}{4}\right)^{13}$$

$$+ C_{16}^{14}\left(\frac{1}{4}\right)^2\left(\frac{3}{4}\right)^{14}$$

$$= \frac{1}{4294967296}[59049 \times 8008 + 4368 \times 177147$$

$$+ 1820 \times 531441 + 560 \times 1594323 + 120 \times 4782969]$$

$$= \frac{1}{4294967296}[472864392 + 773778096 +$$

$$967222620 + 892820880 + 573956280]$$

$$= 0.8570 \qquad \therefore Pr(A_\varepsilon^{(n)}) = 0.8570$$

$$|A_\varepsilon^{(n)}| = C_{16}^{10} + C_{16}^{11} + C_{16}^{12} + C_{16}^{13} + C_{16}^{14}$$
$$= 8008 + 4368 + 1820 + 560 + 120 = 14876$$
$$|A_\varepsilon^{(n)}| = 14876$$

1e) Did $Pr(A_\varepsilon^{(n)})$ increase with n?

Yes. for $\varepsilon = 0.2$ $n = 8$ $Pr(A_\varepsilon^{(n)}) = 0.7861$

for $\varepsilon = 0.2$ $n = 16$ $Pr(A_\varepsilon^{(n)}) = 0.8570$

(f) Confirm that the inequality between $|A_\varepsilon^{(n)}|$ and $2^{n(H+\varepsilon)}$ is satisfied

$$1 = \sum_{x \in X^n} p(x) \geq \sum_{x \in A_\varepsilon^{(n)}} p(x)$$
$$\geq \sum_{x \in A_\varepsilon^{(n)}} 2^{-n(H(x)+\varepsilon)} = 2^{-n(H(x)+\varepsilon)} |A_\varepsilon^{(n)}|$$
$$\therefore |A_\varepsilon^{(n)}| \leq 2^{n(H(x)+\varepsilon)}$$

for $n = 8$ $|A_\varepsilon^{(n)}| = 92 < 2^8 < 2^{8(1.0113)}$

for $n = 16$ $|A_\varepsilon^{(n)}| = 14876 < 2^{16} < 2^{16(1.0113)}$

$\therefore$ it is confirmed that for both $n = 8$ and $n = 16$, it satisfies

$$|A_\varepsilon^{(n)}| \leq 2^{n(H(x)+\varepsilon)}$$

4. The AEP and source coding

$p(1) = 0.005$ and $p(0) = 0.995$. The digits are taken 100 at a time and a binary codeword is provided

(a) Assuming that all codewords are the same length. find the minimum length required to provide codewords for all sequences with three or fewer ones.

The number of sequence that required is

$C_{100}^{0} + C_{100}^{1} + C_{100}^{2} + C_{100}^{3}$

$= 1 + 100 + 4950 + 161700 = 166751$

$\lceil \log_2 166751 \rceil = 18$ bits

So we need at least 18 bits

(b) How does it compare to "brute-force" representation?

It reduces a lot of bits than the "brute-force". But the "brute-force" codes every possible sequence, the method in (a) only codes sequences with very large possibility, for rare cases, it does not provide a code

(c) The probability for observing is

$$C_{100}^{0}(0.995)^{100} + C_{100}^{1}(0.995)^{99}(0.005)$$

$$+ C_{100}^{2}(0.995)^{98}(0.005)^2 + C_{100}^{3}(0.995)^{97}(0.005)^3$$

$$= 0.60577 + 0.30441 + 0.7572 + 0.01243 = 0.99833$$

$1 - 0.99833 = 0.00167$

$\therefore$ The probability that the sequence has no codeword is 0.00167

5. Time's arrow Let $\{X_i\}_{i=-\infty}^{+\infty}$ be a stationary stochastic process. Prove that

$$H(X_0 \mid X_{-1}, X_{-2}, \cdots X_{-n}) = H(X_0 \mid X_1, X_2 \cdots X_n)$$

proof $H(X_0 \mid X_{-1}, X_{-2} \cdots X_{-n})$

$$= H(X_0, X_{-1}, X_{-2} \cdots X_{-n}) - H(X_{-1}, X_{-2} \cdots, X_{-n})$$

$$= H(X_n, X_{n-1}, \cdots X_0) - H(X_n, X_{n-1}, \cdots X_1)$$

$$= H(X_0 \mid X_n, X_{n-1}, \cdots X_1)$$

$$= H(X_0 \mid X_1, X_2 \cdots X_n) \quad \square$$

6. Average entropy per element vs conditioning entropy. Show that

$H(X_1, X_2 \cdots X_n)/n \geq H(X_n \mid X_{n-1}, \cdots X_1)$

For a stationary stochastic process $X_1, X_2 \cdots X_n$

$H(X_n \mid X_{n-1}, X_{n-2}, \cdots X_1) \leq H(X_i \mid X_{i-1}, \cdots X_1)$

for $1 \leq i \leq n$

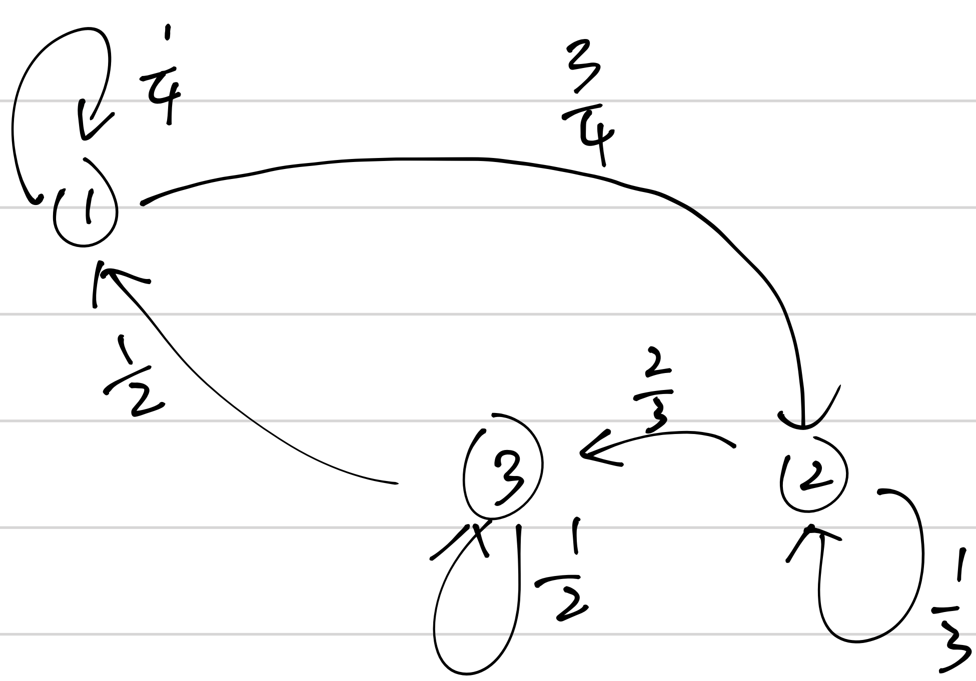
$\frac{1}{n} H(X_1, X_2 \cdots X_n) = \frac{1}{n} \sum_{i=1}^{n} H(X_i \mid X_{i-1}, \cdots X_1)$

$\geq \frac{1}{n} \sum_{i=1}^{n} H(X_n \mid X_{n-1}, \cdots X_1) = \frac{1}{n} \times n \times H(X_n \mid X_{n-1} \cdots X_1)$

$= H(X_n \mid X_{n-1}, \cdots X_1)$

$\therefore \frac{1}{n} H(X_1, X_2 \cdots X_n) \geq H(X_n \mid X_{n-1}, \cdots X_1)$

7. Adam's Seat Selection



From the graph above we can draw the transfer matrix

$$P = \begin{bmatrix} \frac{1}{4} & \frac{3}{4} & 0 \\ 0 & \frac{1}{3} & \frac{2}{3} \\ \frac{1}{2} & 0 & \frac{1}{2} \end{bmatrix} \qquad \mu P = \mu \qquad \mu = [\mu_1, \mu_2, \mu_3]$$

$$\therefore \begin{cases} \frac{1}{4}\mu_1 + \frac{1}{2}\mu_3 = \mu_1 \\ \frac{3}{4}\mu_1 + \frac{1}{3}\mu_2 = \mu_2 \\ \frac{2}{3}\mu_2 + \frac{1}{2}\mu_3 = \mu_3 \end{cases} \Rightarrow \begin{cases} \mu_1 + 2\mu_3 = 4\mu_1 \\ 9\mu_1 + 4\mu_2 = 12\mu_2 \\ 4\mu_2 + 3\mu_3 = 6\mu_3 \end{cases}$$

and $\mu_1 + \mu_2 + \mu_3 = 1$

$$\begin{cases} 2\mu_3 = 3\mu_1 \\ 9\mu_1 = 8\mu_2 \\ 4\mu_2 = 3\mu_3 \end{cases} \qquad \mu_1 + \frac{9}{8}\mu_1 + \frac{3}{2}\mu_1 = 1$$

$$\therefore \mu_1 = \frac{8}{29} \quad \mu_2 = \frac{9}{29} \quad \mu_3 = \frac{12}{29}$$

The stationary probability is $\mu = [\frac{8}{29}, \frac{9}{29}, \frac{12}{29}]$

$$H(\{X_i\}) = H(X_2|X_1)$$

$$= \mu_1 H(X_2|X_1 = \mu_1) + \mu_2 H(X_2|X_1 = \mu_2) + \mu_3 H(X_2|X_1 = \mu_3)$$

$$= \frac{8}{29} H(\frac{3}{4}, \frac{1}{4}) + \frac{9}{29} H(\frac{2}{3}, \frac{1}{3}) + \frac{12}{29} H(\frac{1}{2}, \frac{1}{2})$$

$= 0.2759 \times 0.8113 + 0.3103 \times 0.9183 + 0.4138 \times 1$

$= 0.9226$ bit

$\therefore$ The entropy rate is 0.9226 bit

8. Random Walk on chessboard. Find the entropy rate of the Markov chain associated with a random walk of a king on the 3x3 chessboard

1	2	3
4	5	6
7	8	9

A king is likely to move to any adjacent square at each transition

$\therefore H(X_2 | X_1 \in \text{interior}) = \log 8$

$H(X_2 | X_1 \in \text{corner}) = \log 3$

$H(X_2 | X_1 \in \text{edge}) = \log 5$

The stationary distribution is given by $\mu_i = w_i / 2w$, where w_i is number of edges from node i.

We have 4 corners with 3 edges

4 sides with 5 edges

1 interior with 8 edges

$2W = 4\times 3 + 4\times 5 + 1\times 8 = 40$

$\therefore \mu_{interior} = \frac{8}{40}$ $\mu_{corner} = \frac{3}{40}$ $\mu_{side} = \frac{5}{40}$

$\therefore$ The entropy rate is

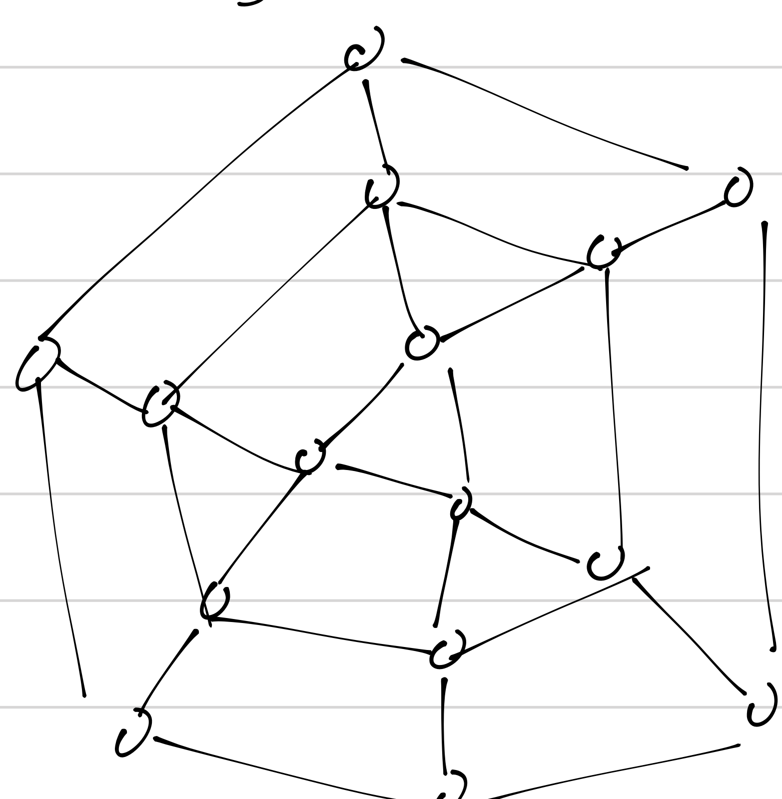
$$H(\{X_i\}) = \sum_{i=1}^{9} \mu_i H(X_2 | X_1 = i)$$

$$= \frac{8}{40}\log 8 + \frac{3}{40}\times 4 \log 3 + \frac{5}{40}\times 4 \log 5$$

$$= 2.24 \text{ bits}$$

$\therefore$ The entropy rate is 2.24 bits

9. Random walk of a Spider



Compute the entropy rate for the random walk of a spider on the web shown above

We can divide these nodes into 2 part

The inner parts :

9 nodes. for each node, the spider has 4 choices to move to an adjacent node

$H(X_2|X_1 = \text{inner}) = \log 4$

The outer parts

6 nodes. For each node, the spider has 3 choices to move to an adjacent node

$H(X_2|X_1 = \text{outer}) = \log_2 3$

The stationary probability is given by $u_i = w_i / w$

we have :

9 inner nodes with 4 edges

6 outer nodes with 3 edges

$2w = 9 \times 4 + 6 \times 3 = 36 + 18 = 54$

$\mu_{inner} = \frac{4}{54}$ $\quad \mu_{outer} = \frac{3}{54}$

The entropy rate $H(\{X_i\}) =$

$\sum_{i=1}^{15} \mu_i H(X_2 | X_1 = i)$

$$= \frac{4}{54} \times 9 \log 4 + \frac{3}{54} \times 6 \log 3$$

$$= \frac{2}{3} \times 2 + \frac{1}{3} \log_2 3 = 1.8617 \text{ bit}$$

$\therefore$ The entropy rate is 1.8617 bits

10 why entropy involves a logarithm

(a) use the Grouping axiom to show that

$$H_m(p_1, \ldots p_m) = H_{m-k+1}(p_1+p_2+\cdots+p_k, p_{k+1}, \cdots, p_m)$$
$$+ (p_1+p_2+\cdots+p_k) H_k(p_1/(p_1+\cdots+p_k), \ldots, p_k(p_1+\cdots+p_k))$$

The grouping axiom is:

$$H_m(p_1, p_2, \ldots p_m) = H_{m-1}(p_1+p_2, p_3 \cdots, \ldots p_m) +$$
$$(p_1+p_2) H_2(p_1/(p_1+p_2), p_2/(p_1+p_2))$$

$$\therefore H_m(p_1, p_2 \cdots p_m) = H_{m-2}(p_1+p_2+p_3, p_4 \cdots p_m) +$$
$$(p_1+p_2+p_3) H_2\left(\frac{p_1+p_2}{p_1+p_2+p_3}, \frac{p_3}{p_1+p_2+p_3}\right)$$
$$+ (p_1+p_2) H_2\left(\frac{p_1}{p_1+p_2}, \frac{p_2}{p_1+p_2}\right)$$

In order to make it simple, we use

$S_k = p_1+p_2+ \cdots + p_k$

$$\therefore H_m(p_1, p_2 \cdots p_m) = H_{m-1}(S_2, p_3, p_4 \cdots, p_m) + S_2 H_2\left(\frac{S_1}{S_2}, \frac{p_2}{S_2}\right)$$

$$= H_{m-2}(S_3, p_4, \cdots p_m) + S_3 H_2\left(\frac{S_2}{S_3}, \frac{p_3}{S_3}\right) + S_2 H_2\left(\frac{S_1}{S_2}, \frac{p_2}{S_2}\right)$$

$$\vdots$$

$$= H_{m-(k-1)}(S_k, p_{k+1}, \cdots p_m) + \Big[S_k H_2\left(\frac{S_{k-1}}{S_k}, \frac{p_k}{S_k}\right)$$

$$+ S_{k-1} H_2\left(\frac{S_{k-2}}{S_{k-1}}, \frac{p_{k-1}}{S_{k-1}}\right) + S_{k-2} H_2\left(\frac{S_{k-3}}{S_{k-2}}, \frac{p_{k-2}}{S_{k-2}}\right)$$

$$+ \cdots + S_2 H_2\left(\frac{S_1}{S_2}, \frac{p_2}{S_2}\right) \Big]$$

we notice that in the "[ ]"

$$\Rightarrow S_2 H_2\left(\frac{S_1}{S_2}, \frac{p_2}{S_2}\right) + S_3 H_2\left(\frac{S_2}{S_3}, \frac{p_3}{S_3}\right) + \cdots + S_k H_2\left(\frac{S_{k-1}}{S_k}, \frac{p_k}{S_k}\right)$$

$$= S_3 \left[\boxed{H_2\left(\frac{S_2}{S_3}, \frac{p_3}{S_3}\right) + \frac{S_2}{S_3} H_2\left(\frac{S_1}{S_2}, \frac{p_2}{S_2}\right)} \right] + \cdots + S_k H_2\left(\frac{S_{k-1}}{S_k}, \frac{p_k}{S_k}\right)$$

$$H\left(\frac{p_1}{S_3}, \frac{p_2}{S_3}, \frac{p_3}{S_3}\right) = H_2\left(\frac{S_2}{S_3}, \frac{p_3}{S_3}\right) + \frac{S_2}{S_3} H_2\left(\frac{S_1}{S_2}, \frac{p_2}{S_2}\right)$$

$$= \boxed{S_3 H\left(\frac{p_1}{S_3}, \frac{p_2}{S_3}, \frac{p_3}{S_3}\right) + S_4 H_2\left(\frac{S_3}{S_4}, \frac{p_4}{S_4}\right)} + \cdots$$

$$+ S_k H_2\left(\frac{S_{k-1}}{S_k}, \frac{p_k}{S_k}\right)$$

$$H\left(\frac{p_1}{S_4}, \frac{p_2}{S_4}, \frac{p_3}{S_4}, \frac{p_4}{S_4}\right) = H_2\left(\frac{S_3}{S_4}, \frac{p_4}{S_4}\right) + \frac{S_3}{S_4} H\left(\frac{p_1}{S_3}, \frac{p_2}{S_3}, \frac{p_3}{S_3}\right)$$

$= S_4 H_4\left(\frac{p_1}{S_4}, \frac{p_2}{S_4}, \frac{p_3}{S_4}, \frac{p_4}{S_4}\right) + \cdots + S_k H_2\left(\frac{S_{k-1}}{S_k}, \frac{p_k}{S_k}\right)$

$\vdots$

$= S_k H_k\left(\frac{p_1}{S_k}, \frac{p_2}{S_k}, \cdots, \frac{p_k}{S_k}\right)$

$\therefore H_m(p_1, p_2 \cdots, p_m) = H_{m-(k-1)}(S_k, p_{k+1}, \cdots p_m)$

$+ S_k H_k\left(\frac{p_1}{S_k}, \frac{p_2}{S_k}, \cdots, \frac{p_k}{S_k}\right)$

$(S_k = p_1 + p_2 + \cdots + p_k)$ $\square$

(b) Let $f(m)$ be the application of H_m to a uniform distribution on m symbols.

$f(m) = H_m\left(\frac{1}{m}, \frac{1}{m}, \cdots \frac{1}{m}\right)$

Show that $f(mn) = f(m) + f(n)$

$f(mn) = H_{mn}\left(\frac{1}{mn}, \frac{1}{mn}, \cdots \frac{1}{mn}\right)$

$= H_{mn-n+1}\left(\frac{1}{mn}, \frac{1}{mn}, \cdots \frac{1}{mn}\right) + \frac{1}{m} H_n\left(\frac{1}{n}, \frac{1}{n} \cdots \frac{1}{n}\right)$

$= H_{mn-2n+2}\left(\frac{1}{m}, \frac{1}{m}, \frac{1}{mn}, \cdots \frac{1}{mn}\right) + \frac{2}{m} H_n\left(\frac{1}{n}, \frac{1}{n}, \cdots \frac{1}{n}\right)$

$= H_{mn-3n+3}\left(\frac{1}{m}, \frac{1}{m}, \frac{1}{m}, \frac{1}{mn}, \cdots \frac{1}{mn}\right) + \frac{3}{m} H_n\left(\frac{1}{n}, \frac{1}{n}, \cdots \frac{1}{n}\right)$

$\vdots$

$= H_{mn-mn+m}\left(\frac{1}{m}, \frac{1}{m}, \cdots \frac{1}{m}\right) + \frac{m}{m} H_n\left(\frac{1}{n}, \frac{1}{n}, \cdots \frac{1}{n}\right)$

$= H_m\left(\frac{1}{m}, \frac{1}{m} \cdots \frac{1}{m}\right) + H_n\left(\frac{1}{n}, \frac{1}{n} \cdots \frac{1}{n}\right)$

$= f(m) + f(n)$

$\therefore f(mn) = f(m) + f(n)$

(c) $H_2\left(\frac{1}{2}, \frac{1}{2}\right) = 1$

$H_b(x) = \log_b a \; H_a(x)$

$\therefore H_e(x) = \ln 2 \, H_2\left(\frac{1}{2}, \frac{1}{2}\right) = 1$

$\therefore$ The normalization should be

$H_2\left(\frac{1}{2}, \frac{1}{2}\right) = \frac{1}{\ln 2}$