# CHAPTER 14

14-1 It suffices to show that [see (14-41)]

$$H(A \cdot B|B_j) = H(A|B_j)$$

Since

$$A_iB_kB_j = \begin{cases} A_iB_j & k = j \\ \{\emptyset\} & k \neq j \end{cases} \qquad \text{and } P(A_iB_j|B_j) = P(A_i|B_j)$$

(14-40) yields

$$H(A \cdot B|B_j) = -\sum_{i,k} P(A_iB_k|B_j) \log P(A_iB_k|B_j)$$

$$= -\sum_i P(A_i|B_j) \log P(A_i|B_j) = H(A|B_j)$$

---

14-2 If $\alpha < \beta$, then $\phi'(\alpha) > \phi'(\beta)$ because

$$\phi'(\alpha) - \phi'(\beta) = \log(\beta/\alpha) > 0. \quad \text{Hence,}$$

$$\int_a^b \phi'(\alpha)d\alpha > \int_{a+c}^{b+c} \phi'(\alpha)d\alpha \qquad c > 0$$

This yields

$$\phi(p_1 + p_2) - \phi(p_1) = \int_{p_1}^{p_1+p_2} \phi'(\alpha)d\alpha < \int_0^{p_2} \phi'(\alpha)d\alpha = \phi(p_2)$$

Similarly

$$\phi(p_1 + \varepsilon) - \phi(p_1) - \phi(p_2) + \phi(p_2 - \varepsilon)$$

$$= \int_{p_1}^{p_1+\varepsilon} \phi'(\alpha)d\alpha - \int_{p_2-\varepsilon}^{p_2} \phi'(\alpha)d\alpha > 0$$

---

14-3 Applying the identity

$$H(A_1 \cdot A_2) = H(A_1) + H(A_2|A_1) \qquad (i)$$

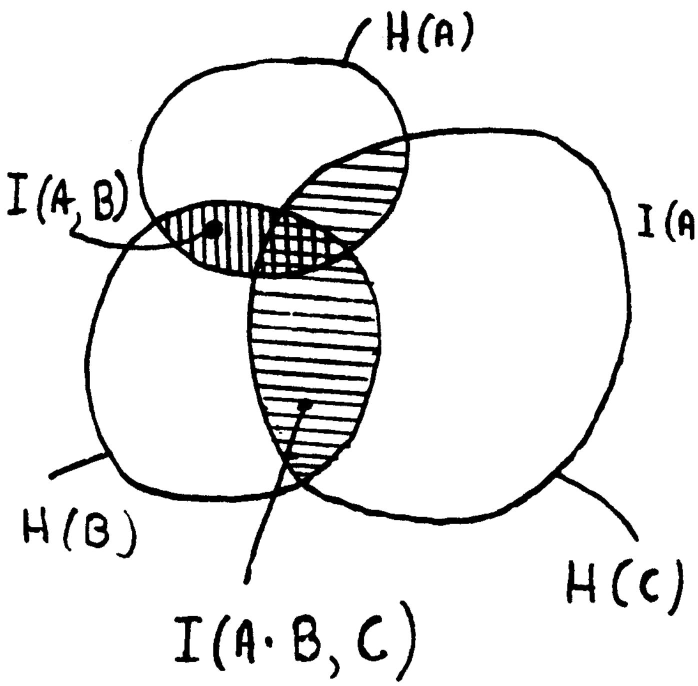
to the partitions $A_1 = A$, $A_2 = B \cdot C$ and $A_1 = A \cdot B$, $A_2 = C$, we obtain the first line. The second line follows from the first [see (i)]. The third line is a consequence of the first two.

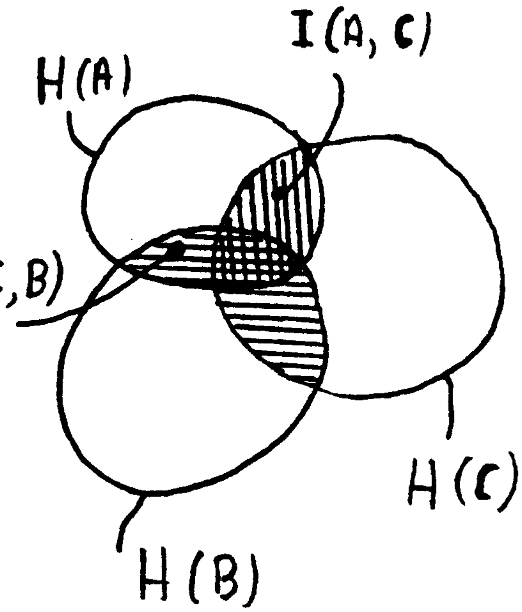
---

14-4 It follows if we apply the identity

$$I(A_1, A_2) = H(A_1) + H(A_2) - H(A_1 \cdot A_2)$$

to the partitions $A_1 = A \cdot B$, $A_2 = C$.




---

14-5 **(a)** From (14-53)

$$I(A,B \cdot C) = H(A) + H(B \cdot C) - H(A \cdot B \cdot C)$$

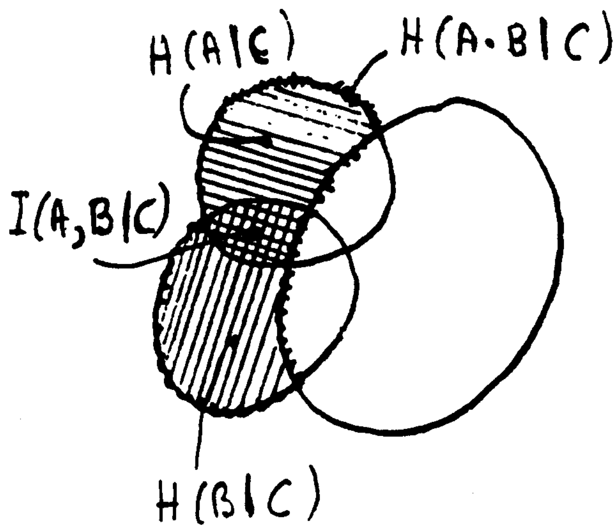
$$I(A,C) = H(A) + H(C) - H(A \cdot C)$$

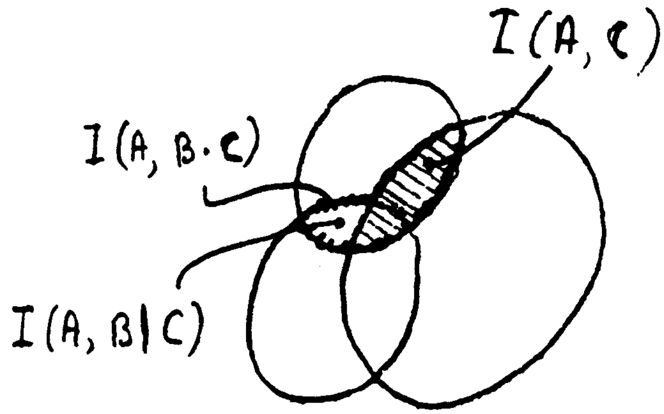
and since (see Prob. 14-4)

$$H(A \cdot B \cdot C) - H(A \cdot C) = H(A \cdot B|C) - H(A|C)$$

we conclude with (14-49) that

$$I(A,B \cdot C) - I(A \cdot C) = H(B|C) + H(A|C) - H(A \cdot B|C)$$




(b) If B • C is observed, then the resulting prediction in the uncertainty of A equals I(A, B • C). But, if B • C is observed, then C is observed, hence, the reduction in the uncertainty of A is at least I(A,C). Hence

$$I(A,B \cdot C) \geq I(A,C)$$

with equality only if $I(A,B|C) = 0$, i.e., if in the subsequence of trials in which C occurred, knowledge of the occurrence of B gives no information about A.

14-6 The partition $H(A^3)$ has eight elements with respective probabilities

$$p^3, p^2q, p^2q, p^2q, pq^2, pq^2, pq^2, q^3$$

Hence

$$H(A^3) = -p^3\log p^3 - 3p^2q\log p^2q - 3pq^2\log pq^2 - q^3\log q^3$$

$$= -3p(p^2+2pq+q^2)\log p - 3q(p^2+2pq+q^2)\log q$$

$$= -3p\log p - 3q\log q = 3H(A)$$

---

14-7 The density of the RV $\underset{\sim}{w} = \underset{\sim}{x} + a$ equals $f_x(w-a)$. Hence,

$$H(\underset{\sim}{x}+a) = -\int_{-\infty}^{\infty} f_x(w-a)\log f_x(w-\alpha)\,dw$$

$$= -\int_{-\infty}^{\infty} f_x(x)\log f_x(x)\,dx = H(\underset{\sim}{x})$$

The joint density of the RVs $\underset{\sim}{x}$ and $\underset{\sim}{z} = \underset{\sim}{x}+\underset{\sim}{y}$ equals $f_{xy}(x,z-x)$. Hence [see (14-90)]

$$H(\underset{\sim}{z}|\underset{\sim}{x}) = -\int_{-\infty}^{\infty}\int_{-\infty}^{\infty} f_{xy}(x,z-x)\log f_{xy}(x,z-x)/f_x(x)\,dx\,dz$$

$$= -\int_{-\infty}^{\infty}\int_{-\infty}^{\infty} f_{xy}(x,y)\log f_{xy}(x,y)/f_x(x)\,dx\,dy = H(\underset{\sim}{y}|\underset{\sim}{x})$$

---

14-8 The RVs $\underset{\sim}{x}$ and $\underset{\sim}{y}$ take the values $x_i$ and $y_j$ respectively then $\underset{\sim}{z} = x_i + y_j$ iff $\underset{\sim}{x} = x_i$ and $\underset{\sim}{y} = y_j$ (assumption). Hence,

$$\{\underset{\sim}{z} = x_i + y_j\} = \{\underset{\sim}{x} = x_i\} \cap \{\underset{\sim}{y} = y_j\}$$

This shows that $A_z = A_x \cdot B_y$. Furthermore, since the RVs $\underset{\sim}{x}$ and $\underset{\sim}{y}$ are independent, the events $\{\underset{\sim}{x} = x_i\}$ and $\{\underset{\sim}{y} = y_j\}$ are also independent. This shows that the partitions $A_x$ and $B_y$ are independent and [see (14-44) and Prob. 14-1]

$$H(A_z|A_x) = H(A_x \cdot A_y|A_x) = H(A_y|A_x) = H(A_y)$$

From this it follows that $H(\underset{\sim}{z}|\underset{\sim}{x}) = H(\underset{\sim}{y})$ because [see (14-88) and (14-41)]

$$H(\underset{\sim}{z}|\underset{\sim}{x}) = H(A_z|A_x)$$

---

14-9 As we see from (14-80)

$H(\underset{\sim}{x}) = \ln a$ where we assume that $a = N\delta$. The RV $\underset{\sim}{y}$ takes the values $0, \delta, \ldots, (N-1)\delta$ with probability $1/N$. The conditional density of $\underset{\sim}{x}$ assuming $\underset{\sim}{y} = k\delta$ is uniform in the interval $(k\delta, k\delta + \delta)$. Hence,

$$H(\underset{\sim}{x}|\underset{\sim}{y} = k\delta) = -\int_{k\delta}^{k\delta+\delta} f(x|\underset{\sim}{y} = k\delta) \ln f(x|\underset{\sim}{y} = k\delta)\,dx = \ln \delta$$

And as in (14-41)

$$H(\underset{\sim}{x}|\underset{\sim}{y}) = \sum_{k=0}^{N} H(\underset{\sim}{x}|\underset{\sim}{y} = k\delta) P\{\underset{\sim}{y} = k\delta\} = \ln \delta$$

Finally [see (14-95)]

$$I(\underset{\sim}{x},\underset{\sim}{y}) = H(\underset{\sim}{x}) - H(\underset{\sim}{x}|\underset{\sim}{y}) = \ln a - \ln \delta$$

---

14-10 If $y_i = g(x_i)$, $y_j = g(x_j)$ and $y_i = y_j$ then $x_i = x_j$. Hence,

$$p_{ij} = \begin{cases} p_i & i = j \\ 0 & i \neq j \end{cases} \qquad p_i = P\{\underset{\sim}{x} = x_i\}$$

and

$$H(\underset{\sim}{x},\underset{\sim}{y}) = -\sum_{i,j} p_{ij} \log p_{ij} = -\sum_i p_i \log p_i = H(\underset{\sim}{x})$$

---

14-11 From Prob. 10-10 it follows with $g(x) = x$ that $H(\underset{\sim}{x},\underset{\sim}{x}) = H(\underset{\sim}{x})$. And since [see (14-103)] $H(\underset{\sim}{x},\underset{\sim}{x}) = H(\underset{\sim}{x}|\underset{\sim}{x}) + H(\underset{\sim}{x})$ we conclude that $H(\underset{\sim}{x}|\underset{\sim}{x}) = 0$. From Prob. 14-3 it follows that

$$H(\underset{\sim}{y},\underset{\sim}{x}|\underset{\sim}{x}) = H(A_y \cdot A_x|A_x) = H(A_x \cdot A_x) + H(A_y|A_x \cdot A_x)$$

$$= H(A_y|A_x) = H(\underset{\sim}{y}|\underset{\sim}{x})$$

because $A_x \cdot A_x = A_x$ and $H(A_x \cdot A_x) = H(\underset{\sim}{x},\underset{\sim}{x}) = 0$.

---

14-12 $E\{\underset{\sim}{x}_n\} = 0$ $E\{\underset{\sim}{x}_n^2\} = 5$ $E\{\underset{\sim}{y}_n\} = 0$

$$E\{\underset{\sim}{y}_n^2\} = \sum_{k=0}^{\infty} 2^{-2k} E\{\underset{\sim}{x}_{n-k}^2\} = \frac{20}{3} \qquad E\{\underset{\sim}{x}_n\underset{\sim}{y}_n\} = E\{\underset{\sim}{x}_n^2\} = 5$$

(a) From (14-95), (14-84), and (15-86) with $\mu_{11} = 5$, $\mu_{22} = 20/3$, and $\mu_{12} = 5$

$$H(\underset{\sim}{x}) = \ln\sqrt{10\pi e} \qquad H(\underset{\sim}{y}) = \ln\sqrt{40\pi e/3} \qquad H(\underset{\sim}{x},\underset{\sim}{y}) = \ln 10\pi e/\sqrt{3}$$

$$I(\underset{\sim}{x},\underset{\sim}{y}) = \ln 2$$

(b) The process $\underset{\sim}{y}(t)$ is the output of the system

$$L(z) = \frac{1}{1-0.5\,z^{-1}} \qquad \ell_o = 1$$

with input $\underset{\sim}{x}_n$. Since $\bar{H}(\underset{\sim}{x}) = H(\underset{\sim}{x})$ and [see (12A-1)

$$\frac{1}{2\pi}\int_{-\pi}^{\pi} \ln|L(e^{j\phi})|\,d\phi = \ln\ell_o = 0$$

(14-133) yields $\bar{H}(\underset{\sim}{y}) = \bar{H}(\underset{\sim}{x}) = H(\underset{\sim}{x}) = \ln\sqrt{10\pi e}$.

---

14-13
$$\bar{H}(\underset{\sim}{x}) = H(\underset{\sim}{x}) = -\frac{1}{2}\int_4^6 \ln\frac{1}{2}\,dx = \ln 2$$

And as in Prob. 14-12 with $\ell_0 = 5$,

$$\bar{H}(\underset{\sim}{y}) = \bar{H}(x) + \ln 5 = \ln 10$$

---

14-14 Given that f(x) = 0 for |x|>1 and $E\{\underset{\sim}{x}\} = 0.3$, find f(x). With g(x) = x, (14-143) yields $f(x) = Ae^{-\lambda x}$ where

$$A\int_{-1}^{1} e^{-\lambda x}\,dx = \frac{A}{\lambda}(e^{\lambda} - e^{-\lambda}) = 1$$

$$A\int_{-1}^{1} xe^{-\lambda x}\,dx = \frac{A}{\lambda^2}(e^{\lambda} - e^{-\lambda}) - \frac{A}{\lambda}(e^{\lambda} - e^{-\lambda}) = 0.31$$

Solving, we obtain $A \simeq 0.425$, $\lambda \simeq -1$

---

14-15 $f(x) = Ae^{-\lambda x}$ for 1<x<5 and 0 otherwise,

$$A\int_{1}^{5} e^{-\lambda x}\,dx = 0.31 \qquad A\int_{1}^{5} xe^{-\lambda x}dx = 3\,\frac{37}{60}$$

Hence, $A \simeq 1.06$, $\lambda \simeq 0.5$

---

14-16 From (14-151) with $x_k=k$, $g_1(x_k) = g_1(k) = k$, $k=1, \ldots, 6$

$$g_2(x_k)=\begin{cases} 0 & k=1,3,5 \\ 1 & k=2,4,6 \end{cases} \qquad p_k=\begin{cases} Ae^{-\lambda_1 k} & k=1,3,5 \\ Ae^{-\lambda_1 x-\lambda_2} & k=2,4,6 \end{cases}$$

Since $p_1 + p_3 + p_5 = 0.5$ and $E\{\underset{\sim}{x}\} = 4.44$, we conclude with $z = e^{-\lambda_2}$ and $w = e^{-\lambda_2}$ that

$$A(z+z^3+z^5) = Aw(z^2+z^4z^6)$$

$$A(\underset{\sim}{z}+3z^3+5z^5) + Aw(2z^2+4z^4+6z^6) = 4.44$$

This yields $A \simeq 0.0437$, $\underset{\sim}{z} = 1/w \simeq 1.468$

---

14-17 (a) The transformation $\underset{\sim}{y} = 3\underset{\sim}{x}$ is one-to-one, hence, $H(\underset{\sim}{y}) = H(\underset{\sim}{x})$

(b) From (14-113) with $g(x) = 3x$: $H(\underset{\sim}{y}) = H(\underset{\sim}{x}) + \ell n\ 3$

---

14-18 (a) For fair dice, $P\{7\} = \frac{1}{6}$, $P(11) = \frac{1}{18}$, $P\{\text{neither 7 nor 11}\} = \frac{14}{18}$

$$H(A) = -\left(\frac{1}{6}\ \ell n \frac{1}{6} + \frac{1}{18} \ell n \frac{1}{18} + \frac{14}{18}\ \ell n \frac{14}{18}\right) = 0.655$$

(b) From (14-10) with n=100 and N=3:

$$n_T \simeq e^{nH(A)} \simeq 2.79 \times 10^{28} \qquad n_a \simeq N^n \simeq 5.16 \times 10^{47}$$

---

14-19 The process $x_n$ is WSS with entropy rate $\overline{H}(x)$. Show that, if

$$w_n = \sum_{k=0}^{n} x_{n-k}\, \ell_k$$

then

$$\lim_{n\to\infty} \frac{1}{n+1} H(w_o,\ldots,w_n) = \overline{H}(x) + \ln|\ell_o| \qquad \text{(i)}$$

<u>Proof.</u> The RVs $w_o,\ldots,w_n$ are linear transformations of the RVs $x_o,\ldots,x_n$ and the transformation matrix equals

$$\begin{bmatrix} \ell_0 & 0 & \ldots & 0 \\ \ell_1 & \ell_0 & \ldots & 0 \\ \hline \ell_n & \ell_{n-1} & \ldots & 0 \end{bmatrix}$$

Since the determinant of this transformation equals $|\ell_o|^{n+1}$, (14-115) yields

$$H(w_o,\ldots,w_n) = H(x_o,\ldots x_n) + (n+1)\ln|\ell_o|$$

Dividing by (n+1) we obtain (i) as $n\to\infty$.

---

14-20 As in Example 14-19, $f(p) = A\, e^{-\lambda p}$. To find $\lambda$, we use the $\lambda$-$\eta$ curve of Fig. 14-16. This yields

$$\lambda \simeq -1.23 \qquad f(p) \simeq 0.51\, e^{1.23p}$$

14-21 As in Example 14-22, $p_k = A\ e^{-\lambda k}$. To find $\lambda$, we use the w-η curve of Fig. 14-17. This yields (see also Jaynes)

$$w \simeq 1.449 \qquad \lambda \simeq -0.371$$

| $p_1$ | $p_2$ | $p_3$ | $p_4$ | $p_5$ | $p_6$ |
|---|---|---|---|---|---|
| 0.054 | 0.079 | 0.114 | 0.165 | 0.240 | 0.348 |

---

14-22 The unknown density is normal as in (14-157) where

$$\Delta = \begin{vmatrix} 4 & 1 & 1 \\ 1 & 4 & m_{23} \\ 1 & m_{23} & 4 \end{vmatrix} = -4m_{23}^2 + 2m_{23} + 56$$

The moment $m_{23} = E\{\underset{\sim}{x}_2 \underset{\sim}{x}_3\}$ must be such as to maximize $\Delta$. This yields $m_{23} = 0.25$.

---

14-23

| $p_i$ | 0.3 | 0.2 | 0.15 | 0.15 | 0.1 | 0.06 | 0.04 | |
|---|---|---|---|---|---|---|---|---|
| | $\frac{1}{4} \le p_i < \frac{1}{2}$ | $\frac{1}{8} \le p_i < \frac{1}{4}$ | | | $\frac{1}{16} \le p_i < \frac{1}{8}$ | $\frac{1}{32} \le p_i < \frac{1}{6}$ | | $\sum_{i=1}^{7} \frac{1}{2^{n_i}}$ |
| Shannon $n_i$ | 2 | 3 | 3 | 3 | 4 | 5 | 5 | 0.75 |
| | 2 | 3 | 3 | 3 | 4 | 4 | 4 | 0.8125 |
| | 2 | 3 | 3 | 3 | 3 | 4 | 4 | 0.875 |
| L = 2.7 | 2 | 3 | 3 | 3 | 3 | 3 | 4 | 0.9375 |
| | 2 | 3 | 3 | 3 | 3 | 3 | 3 | 1 |
| $x_i$ | 00 | 010 | 011 | 100 | 101 | 110 | 111 | |

Fano

$L = 2.6$

| $p_i$ | 0.3 | 0.2 | 0.15 | 0.15 | 0.1 | 0.06 | 0.04 |
|---|---|---|---|---|---|---|---|
| | $A_0$ 0.5 | | $A_1$ 0.5 | | | | |
| | $A_{00}$ 0.3 | $A_{01}$ 0.2 | $A_{10}$ 0.3 | | $A_{11}$ 0.2 | | |
| | | | $A_{100}$ 0.15 | $A_{101}$ 0.15 | $A_{110}$ 0.1 | $A_{111}$ 0.1 | |
| | | | | | | $A_{1110}$ | $A_{1111}$ |
| $x_i$ | 00 | 01 | 100 | 101 | 110 | 1110 | 1111 |

Huffman

$L = 2.6$

| | | | | | | | |
|---|---|---|---|---|---|---|---|
| | 1 | 2 | 3 | 4 | 5 | 6 | 7 |
| | 1 | 2 | 3 | 4 | 5 | 6 / 0 | 7 / 1 |
| | 1 | 2 | 5 / 0 | 6 / 10 | 7 / 11 | 3 | 4 |
| | 1 | 3 / 0 | 4 / 1 | 2 | 5 / 0 | 6 / 10 | 7 / 11 |
| | 2 / 0 | 5 / 10 | 6 / 110 | 7 / 111 | 1 | 3 / 0 | 4 / 1 |
| | 1 / 0 | 3 / 10 | 4 / 11 | 2 / 0 | 5 / 10 | 6 / 110 | 7 / 111 |
| | 1 / 00 | 3 / 010 | 4 / 011 | 2 / 10 | 5 / 110 | 6 / 1110 | 7 / 1111 |
| $x_i$ | 00 | 10 | 010 | 011 | 110 | 1110 | 1111 |

14-24 If $\underline{x}_n = 0$, then $\bar{x}_n = 000$ and $\underline{y}_n = 1$ iff $\bar{y}_n$ consists of one 0 or no zeros. The probability of one and only one zero equals $3\beta^2(1-\beta)$ [see (3-13)]; the probability of no zeros equals $\beta^3$. Hence,

$$P\{\underline{y}_n = 1 | \underline{x}_n = 0\} = 3\beta^2(1-\beta) + \beta^3$$

Thus, the redundantly coded channel of Example 14-29 is symmetrical as in (14-191) with probability of error $\beta_1 = \beta^2$.

---

14-25 If the received information is always wrong, then

$$P\{\underline{y}_n = 1 | \underline{x}_n = 0\} = \beta = 1, \text{ hence } C = 1 - r(\beta) = 1$$

---