情報理論 配布資料 #4

担当：井上 純一 (情報科学研究科棟 8-13)
URL：http://chaosweb.complex.eng.hokudai.ac.jp/~j_inoue/

平成 17 年 5 月 16 日

目 次

3.1 KL 情報量と相互情報量 23
3.2 KL 情報量の非負性 24
3.2.1 イェンセンの不等式 24

4 情報源符号化 27
4.1 情報源圧縮の例 27
4.2 復号可能条件 27
4.3 クラフト不等式 28
4.4 クラフト不等式の証明 29
4.5 クラフト不等式を用いた語頭符号の構成 30

演習問題 3 の解答例

(1) 1 回目の試行は仕込みのない完全なコイン投げであるから, $P_A(\mathsf{H}) = P_A(\mathsf{T}) = 1/2$ である. また, 2 回目のコイン投げは 1 回目の試行結果に依存し, 確率 q で 1 回目とは逆向きが出るので, これを条件付き確率で表現すれば

$$P_{B|A}(\mathsf{H}|\mathsf{T}) = q,\ \ P_{B|A}(\mathsf{T}|\mathsf{T}) = 1-q \tag{92}$$

$$P_{B|A}(\mathsf{H}|\mathsf{H}) = 1-q,\ \ P_{B|A}(\mathsf{T}|\mathsf{H}) = q \tag{93}$$

と書ける. これは H を 1 に, T を 0 に対応付ければ, 既に見た 2 元対称通信路であり, 図 10 のようにグラフ表現できることになる.

(2) ベイズの公式から直ちに

$$P_{A|B}(\mathsf{H}|\mathsf{H}) = \frac{P_{B|A}(\mathsf{H}|\mathsf{H})P_A(\mathsf{H})}{P_{B|A}(\mathsf{H}|\mathsf{H})P_A(\mathsf{H}) + P_{B|A}(\mathsf{H}|\mathsf{T})P_A(\mathsf{T})} = \frac{(1-q)\times\frac{1}{2}}{(1-q)\times\frac{1}{2} + q\times\frac{1}{2}} = 1-q \tag{94}$$

$$P_{A|B}(\mathsf{H}|\mathsf{T}) = \frac{P_{B|A}(\mathsf{T}|\mathsf{H})P_A(\mathsf{H})}{P_{B|A}(\mathsf{T}|\mathsf{H})P_A(\mathsf{H}) + P_{B|A}(\mathsf{T}|\mathsf{T})P_A(\mathsf{T})} = \frac{(q\times\frac{1}{2}}{q\times\frac{1}{2} + (1-q)\times\frac{1}{2}} = q \tag{95}$$

$$P_{A|B}(\mathsf{T}|\mathsf{H}) = \frac{P_{B|A}(\mathsf{H}|\mathsf{T})P_A(\mathsf{T})}{P_{B|A}(\mathsf{H}|\mathsf{T})P_A(\mathsf{H}) + P_{B|A}(\mathsf{H}|\mathsf{T})P_A(\mathsf{T})} = \frac{q\times\frac{1}{2}}{(1-q)\times\frac{1}{2} + q\times\frac{1}{2}} = q \tag{96}$$

$$P_{A|B}(\mathsf{T}|\mathsf{T}) = \frac{P_{B|A}(\mathsf{T}|\mathsf{T})P_A(\mathsf{T})}{P_{B|A}(\mathsf{T}|\mathsf{H})P_A(\mathsf{H}) + P_{B|A}(\mathsf{T}|\mathsf{T})P_A(\mathsf{T})} = \frac{(1-q)\times\frac{1}{2}}{q\times\frac{1}{2} + (1-q)\times\frac{1}{2}} = 1-q \tag{97}$$

のように求めることができる.

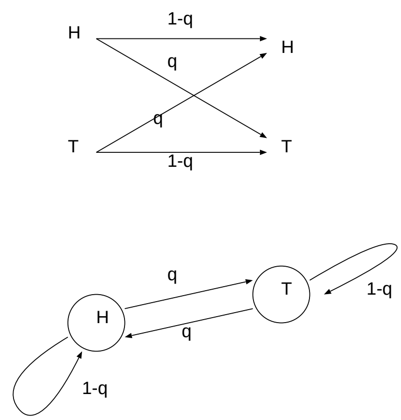

図 10: ここで取り上げたコイン投げのグラフ表現. 上下どちらでもよい.

(3) まず, $H(A)$ は定義に従って簡単に

$$H(A) = -\sum_{i=\mathsf{H},\mathsf{T}} P_A(i)\log P_A(i) = \log 2 = 1(\text{ビット}) \tag{98}$$

と計算することができる. $H(A|B)$ を計算するために, $P_B(\mathsf{H}), P_B(\mathsf{T})$ を求めると

$$\begin{aligned} P_B(\mathsf{H}) &= \sum_{i=\mathsf{H},\mathsf{T}} P_{B|A}(j|i)P_A(i) \\ &= P_{B|A}(\mathsf{H}|\mathsf{H})P_A(\mathsf{H}) + P_{B|A}(\mathsf{H}|\mathsf{T})P_A(\mathsf{T}) \\ &= q\times\frac{1}{2} + (1-q)\times\frac{1}{2} = \frac{1}{2} \end{aligned} \tag{99}$$

と計算できる.　明らかに, $P_B(\mathsf{T}) = 1 - P_B(\mathsf{H}) = 1/2$ であるから, 求める条件付きエントロピーは

$$\begin{aligned} H(A|B) &= -P_{A|B}(\mathsf{H}|\mathsf{H})P_B(\mathsf{H})\log P_{A|B}(\mathsf{H}|\mathsf{H}) - P_{A|B}(\mathsf{H}|\mathsf{T})P_B(\mathsf{T})\log P_{A|B}(\mathsf{H}|\mathsf{T}) \\ &= -P_{A|B}(\mathsf{T}|\mathsf{H})P_B(\mathsf{H})\log P_{A|B}(\mathsf{T}|\mathsf{H}) - P_{A|B}(\mathsf{T}|\mathsf{T})P_B(\mathsf{T})\log P_{A|B}(\mathsf{T}|\mathsf{T}) \\ &= -q\log q - (1-q)\log(1-q) \end{aligned} \tag{100}$$

のように求めることができる. 従って, 相互情報量は

$$\begin{aligned} I(A;B) &= H(A) - H(A|B) \\ &= 1 + q\log q + (1-q)\log(1-q) \end{aligned} \tag{101}$$

のように q の関数として書き下すことができる. 図 11 にこれを図示する.
従って, B, つまり, この「通信路」の出力値である 2 回目のコインの向きを知ることによって増加する 1 回目の試行におけるコイン投げの向きは「通信路」の反転確率 q が $q = 1/2$ のときに最小で, $q = 0$ または $q = 1$ になる. つまり, 2 階目の試行の結果が完全に 1 回目の試行によって確定するときに最大となる.

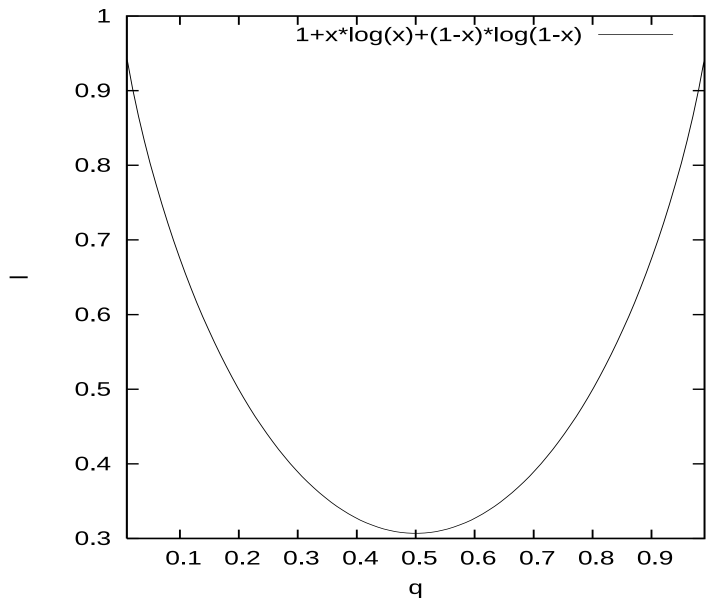

図 11: ここで求めた相互情報量の q 依存性.

3.1 KL 情報量と相互情報量

2 つの確率分布の間の距離として次の **KL 情報量 (Kulback-Leibler)** を定義する.

$$D(P||Q) = \sum_{x \in X} P(x) \log \left\{ \frac{P(x)}{Q(x)} \right\} = \sum_{x \in X} P(x) \log \{\log P(x) - \log Q(x)\} \tag{102}$$

これは別名を**相対エントロピー**と言い, 前回学んだ相互情報量との間には次に示すような関係がある. まず, 同時分布 $P_{XY}(x,y)$ の確率変数 Y, 及び, X に関する**周辺分布**を

$$P_X(x) = \sum_{y \in Y} P_{XY}(x,y), \quad P_Y(y) = \sum_{x \in X} P_{XY}(x,y) \tag{103}$$

で定義すれば, X と Y の間の相互情報量 $I(X;Y)$ は

$$\begin{aligned} I(X;Y) &= H(X) + H(Y) - H(X,Y) \\ &= -\sum_{x \in X} P_X(x) \log P_X(x) - \sum_{y \in Y} P_Y(y) \log P_Y(y) + \sum_{x,y \in X,Y} P_{XY}(x,y) \log P_{XY}(x,y) \\ &= -\sum_{x,y \in X,Y} P_{XY}(x,y) \log P_X(x) - \sum_{x,y \in X,Y} P_{XY}(x,y) \log P_Y(y) \\ &+ \sum_{x,y \in X,Y} P_{XY}(x,y) \log P_{XY}(x,y) \\ &= \sum_{x,y \in X,Y} P_{XY}(x,y) \log \left\{ \frac{P_{XY}(x,y)}{P_X(x) P_Y(y)} \right\} \\ &= D(P_{XY}||P_X P_Y) \end{aligned} \tag{104}$$

と書ける. 従って, **相互情報量は同時分布 $P_{X,Y}(x,y)$ と周辺分布の積 : $P_X(x)P_Y(y)$ の間の距離を KL 情報量で計ったもの**である.

3.2　KL 情報量の非負性

ここでは KL 情報量の持つ性質として, その非負性について見ておこう. そのための準備として次に示すような**イェンセンの不等式 (Jensen)** を学ぶことにする.

3.2.1　イェンセンの不等式

上に凸な関数の性質として, 例えば, 図 12 のような関数 $f(x)$ を考えると, x 軸上の内分点 $\lambda x_1 + (1-\lambda)x_2$ での関数値 $f(\lambda x_1 + (1-\lambda)x_2)$ は必ず関数値の内分点 $\lambda f(x_1) + (1-\lambda)f(x_2)$ よりも大きいことから, 次の不等式が成り立つ.

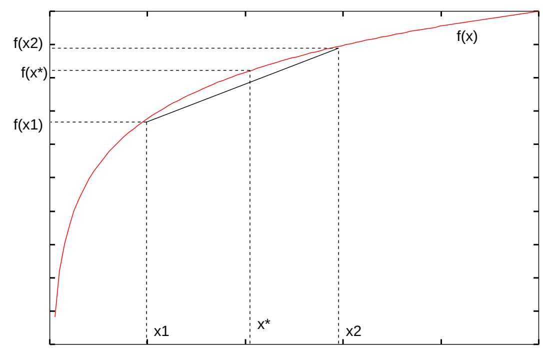

図 12: 上に凸な関数 $f(x)$ とその性質. x_1 と x_2 の内分点 $x_* = \lambda x_1 + (1-\lambda)x_2$ での関数値 $f(x_*)$ は関数値の内分点 $\lambda f(x_1)L + (1-\lambda)f(x_2)$ よりも必ず大きい.

$$f(\lambda x_1 + (1-\lambda)x_2) \quad \geq \quad \lambda f(x_1) + (1-\lambda)f(x_2) \tag{105}$$

ここで, $0 \leq \lambda \leq 1$ を確率変数 x が x_1 をとる確率, つまり, x_1 の出現する確率, $1-\lambda$ を確率変数 x が $x = x_2$ の値をとる確率と解釈するならば, $\lambda x_1 + (1-\lambda)x_2$ は確率変数 x の期待値を表すことに注意する. つまり, 式で書けば

$$E(x) \quad = \quad \lambda x_1 + (1-\lambda()x_2 \tag{106}$$

と書ける. なお, これらの確率と期待値の定義に従えば

$$E(f(x)) \quad = \quad \lambda f(x_1) + (1-\lambda)f(x_2) \tag{107}$$

であるから, (105) 式は

$$f(E(x)) \quad \geq \quad E(f(x)) \tag{108}$$

を表していることになる (これらの不等式は期待値をより一般的に $E(x) = \sum_{x\in X} p(x)x, E(f(x)) = \sum_{x\in X} p(x)f(x)$ として成立する). この不等式を**イェンセン (Jensen) の不等式**と呼んでいる. この不等式を用いることにより, KL 不等式の非負性を示すことができる.

まず, KL 情報量の定義から

$$D(P||Q) \quad = \quad \sum_{x\in A} P(x)\log\frac{P(x)}{Q(x)} = -\sum_{x\in A} P(x)\log\frac{Q(x)}{P(x)} \tag{109}$$

であるが, イェンセンの不等式で $f(x)\mapsto\log x, x\mapsto Q(x)/P(x)$ とおくと

$$\begin{aligned}\sum_{x\in A}P(x)\log\left\{\frac{Q(x)}{P(x)}\right\} &= E\left(f\left(\frac{Q(x)}{P(x)}\right)\right)\\ &\leq f\left(E\left(\frac{Q(x)}{P(x)}\right)\right)=\log\left\{\sum_{x\in A}P(x)\cdot\frac{Q(x)}{P(x)}\right\}=0 \qquad (110)\end{aligned}$$

が成り立つ. 従って, 任意の確率分布 $P(x), Q(x)$ の間の距離である KL 情報量は $D(P||Q)\geq 0$ であり, 絶対に負にはならない. また, このことから直ちに P_{XY} と $P_X P_Y$ にとの間の KL 情報量も非負であるから,

$$I(X;Y) = D(P_{XY}||P_X P_Y)\geq 0 \qquad (111)$$

となり, 相互情報量もまた決して負にはならない量であることがわかる.

例題 5

情報源 X の**エントロピーレート**：

$$\mathcal{H}(X) = \lim_{n\to\infty}\frac{1}{n}H(X_1, X_2, \cdots, X_n) \qquad (112)$$

に関して以下の問いに答えよ.

(1) 記号 $0, 1$ をそれぞれ $p, 1-p$ で生成する**定常的無記憶情報源**のエントロピーレートを求めよ.

(2) 定常分布 $p(x)$ が

$$p(x) = p\,\delta_{x,1}+(1-p)\,\delta_{x,0} \qquad (113)$$

状態遷移確率 $p(x^{'}|x)$ が

$$p(x^{'}|x) = q+(1-2q)\,\delta_{x,x'} \qquad (114)$$

で与えられる**定常的単純マルコフ情報源**を考えよう ($\delta_{a,b}$ はクロネッカ・デルタである). このとき

- この状態遷移確率 $p(x^{'}|x)$ を表す状態遷移図.
- この情報源のエントロピーレート.

のそれぞれを求めよ.

(解答例)

(1) 定常的無記憶情報源の場合, 結合確率は

$$p(x_1, x_2, \cdots, x_n) = \prod_{i=1}^{n}p(x_i) \qquad (115)$$

と書け, この中の確率 $p(x_i)$ のそれぞれが

$$p(x_i) = p\,\delta_{x_i,1}+(1-p)\,\delta_{x_i,0} \qquad (116)$$

で与えられるので, この系のエントロピーは

$$
\begin{aligned}
H(x_1, x_2, \cdots, x_n) &= -\sum_{x_1}\sum_{x_2}\cdots\sum_{x_n}\prod_{i=1}^{n} p(x_i)\log\prod_{i=1}^{n} p(x_i) \\
&= -\sum_{x_1}\sum_{x_2}\cdots\sum_{x_n} p(x_1)p(x_2)\cdots p(x_n)\{\log p(x_1)+\cdots+\log p(x_n)\} \\
&= -n\sum_{x_1=0,1} p(x_1)\log p(x_1)\sum_{x_2}\sum_{x_3}\cdots\sum_{x_n} p(x_2)\cdots p(x_n) \\
&= -n\sum_{x_1=0,1}\{p\,\delta_{x,1}+(1-p)\,\delta_{x,0}\}\log\{p\,\delta_{x,1}+(1-p)\,\delta_{x,0}\} \\
&= -n\{(1-p)\log(1-p)+p\log p\} \qquad (117)
\end{aligned}
$$

であるから, エントロピーレート $\mathcal{H}$ は

$$
\mathcal{H} = \lim_{n\to\infty}\frac{1}{n}H(x_1, x_2, \cdots, x_n) = -p\log p-(1-p)\log(1-p) \qquad (118)
$$

となる.

(2) まず, この定常的マルコフ情報源の状態遷移をグラフで表すと図 13 のようになる. この場合のエント

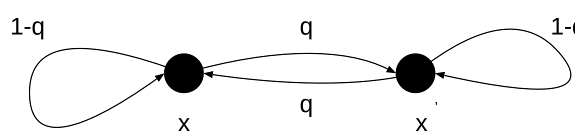

図 13: 遷移確率 $p(x^{'}|x) = q+(1-2q)\delta_{x^{'},x}$ のグラフ表現.

ロピーは結合確率が

$$
p(x_1, x_2, \cdots, x_n) = p(x_n|x_{n-1})p(x_{n-1}|x_{n-2})\cdots p(x_3|x_2)p(x_2|x_1)p(x_1) \qquad (119)
$$

と書けることに注意すれば, この系のエントロピーは

$$
\begin{aligned}
H(x_1, x_2, \cdots, x_n) &= -\sum_{x_1}\cdots\sum_{x_n} p(x_1, \cdots, x_n)\{\log p(x_1)+\log p(x_2|x_1)+\cdots+\log p(x_n|x_{n-1})\} \\
&= -\sum_{x_1} p(x_1)\log p(x_1) - \sum_{x_1}\sum_{x_2} p(x_1, x_2)\log p(x_2|x_1) - \cdots \\
&- \sum_{x_n}\sum_{x_{n-1}} p(x_n, x_{n-1})\log p(x_n|x_{n-1}) \\
&= -\sum_{i=1}^{n}\left\{\sum_{x_{i-1}} p(x_{i-1})\sum_{x_i} p(x_i|x_{i-1})\log p(x_i|x_{i-1})\right\} \\
&= -n\sum_{x} p(x)\sum_{x^{'}} p(x^{'}|x)\log p(x^{'}|x) \qquad (120)
\end{aligned}
$$

で与えられる. 従って, エントロピーレート $\mathcal{H}$ は

$$
\mathcal{H} = \lim_{n\to\infty}\frac{1}{n}H(x_1, \cdots, x_n) = -\sum_{x} p(x)\sum_{x^{'}} p(x^{'}|x)\log p(x^{'}|x) \qquad (121)
$$

となる．ここに問題文に与えられた $p(x), p(x^{'}|x)$ を代入し，$x, x^{'}$ に関する和を実行すれば

$$
\begin{aligned}
\mathcal{H} &= -\sum_{x=0,1}\{p\,\delta_{x,0}+(1-p)\,\delta_{x,1}\}\sum_{x'=0,1}\{q+(1-2q)\,\delta_{x,x'}\}\log\{q+(1-2q)\,\delta_{x,x'}\} \\
&= -p\sum_{x'=0,1}\{q+(1-2q)\,\delta_{0,x'}\}\log\{q+(1-2q)\,\delta_{0,x'}\} \\
&- (1-p)\sum_{x'=0,1}\{q+(1-2q)\,\delta_{1,x'}\}\log\{q+(1-2q)\,\delta_{1,x'}\} \\
&= -p\{(1-q)\log(1-q)+q\log q\}-(1-p)\{q\log q+(1-q)\log(1-q)\} \\
&= -q\log q-(1-q)\log(1-q) \qquad (122)
\end{aligned}
$$

となる．

4 情報源符号化

情報源符号化：情報源が出力する系列を，その系列が持っている情報を失うことなく，短い系列に変換する (情報源の圧縮) こと．

4.1 情報源圧縮の例

例えば，ababbabaabbb という系列を 2 つずつに区切って，ab | ab | ba | ba | ab | bb とし，表のような変換規則に従って，0 と 1 の並びからなる「符号語」に変換する．

アルファベット	**符号語**	**符号長**
aa	110	3
ab	0	1
ba	10	2
bb	111	3

とすると，ababbabaabbb → 0010100111 となる．この変換で 12 ビットの情報が 10 ビットに減少したことに注意しょう．この例の場合，表の対応関係を逆に用いることにより，もとのアルファベット列 ababbabaabb を復元できるので，この場合の圧縮を**可逆圧縮**と呼ぶ．逆にこのような復元ができないものを**非可逆圧縮**と呼ぶ．ここで示した例を一般化しておこう．

$\mathcal{A}=\{a_1, a_2, \cdots, a_M\}$：情報源を作る記号の集合．上の例で言うと $\mathcal{A}=\{a, b\}$.
$\mathcal{B}=\{b_1, b_2, \cdots, b_K\}$：通信路で使われる記号の集合．例で言うと $\mathcal{B}=\{0, 1\}$.

符号化：$\mathcal{A}$ からなる系列を $\mathcal{B}$ からなる系列に変換すること．
復号化：$\mathcal{B}$ からなる系列から，元の $\mathcal{A}$ からなる系列を再生すること．
$x_i \in \mathcal{B}\ \ (i=1,\cdots,k)$ から**符号語 (code words)** $x=x_1x_2\cdots x_k$ を作った場合，k を**符号長**と呼ぶ．

4.2 復号可能条件

情報源系列から符号語への写像 ϕ を次のように定義する．

$$\phi : \mathcal{A} \quad\to\quad \mathcal{B} \qquad (123)$$

$\chi = \{0, 1\}$ とすると, χ^+ を次で定義する.

$$\chi^+ = \{0, 1, 00, 01, 10, 11, 000, \cdots\} \tag{124}$$

つまり, χ^+ は χ の記号を 1 つ以上並べた系列の集合である. また, χ^n を

$$\chi^n = \{x_1 x_2 \cdots x_n : x_i \in \chi, i = 1, \cdots, n\} \tag{125}$$

で定義する. また, $\#S$ を有限集合 S の要素数とすると, 復号できるための条件は

復号できるための条件
$A \subset \mathcal{A}^n, B \subset \mathcal{B}^+$ とし, 符号化は写像：$\psi : A \to B$ の元で行われるものとすると, $\forall_{x,x'} \in A$ に対し,

$$x \neq x' \quad \Rightarrow \quad \psi(x) \neq \psi(x') \tag{126}$$

のとき, 符号 ψ は**正則**であると言う. また,

$$x \neq x' \quad \Leftrightarrow \quad \psi(x) \neq \psi(x') \tag{127}$$

が成り立つとき, この符号 ψ は**一意復号可能**である. という.

語頭条件：どの符号語も他の符号語の先頭部分にはなっていないという条件.

語頭条件を満たしている正則な符号は一意復号可能である.

4.3 クラフト不等式

クラフト不等式：
$\phi : \mathcal{A} \to \mathcal{B}^+$ は一意復号可能であるとし, $l(x)$ を x に対する符号語 $\phi(x)$ の長さとすると, 次の不等式が成り立つ.

$$\sum_{x \in \mathcal{A}} K^{-l(x)} \leq 1 \tag{128}$$

ここに, K は $\mathcal{B}$ の要素数である ($K = \#\mathcal{B}$).

(適用例 1)：
$\psi_1 : \mathcal{A} = \{$ 赤, 青, 黄 $\} \to \mathcal{B} = \{0, 1\}$ とすると, $K = 2$ であり, 変換 ψ_1 は具体的に $\psi_1(\text{赤}) = 0, \psi_1(\text{青}) = 1, \psi_1(\text{黄}) = 10$ で与えられ, それぞれの符号長は $l(\text{赤}) = 1, l(\text{青}) = 1, l(\text{黄}) = 2$ である. このとき, クラフト不等式 (128) の左辺は

$$\sum_{x \in \mathcal{A}} K^{-l(x)} = 2^{-1} + 2^{-1} + 2^{-2} = \frac{1}{2} + \frac{1}{2} + \frac{1}{4} = \frac{5}{4} > 1 \tag{129}$$

となり, クラフトの不等式を満たさない. 従って, 符号 ψ_1 は一意復号不可能である.
(適用例 2)：

$\psi_2 : \mathcal{A} = \{aa, ab, ba, bb\}, \mathcal{B} = \{0, 1\}$ とすると, $K = 2$ である. 具体的な変換規則は $\psi_2(aa) = 00, \psi_2(ab) = 10, \psi_2(ba) = 11, \psi_2(bb) = 110$ であり, それぞれの符号長は $l(aa) = l(ab) = l(ba) = 2, l(bb) = 3$ である. このとき, クラフト不等式 (128) の左辺は

$$\sum_{x \in \mathcal{A}} K^{-l(x)} \;=\; 2^{-2} + 2^{-2} + 2^{-2} + 2^{-3} = \frac{3}{4} + \frac{1}{8} \leq 1 \tag{130}$$

となり, クラフトの不等式を満たす. 従って ψ_2 は一意復号可能である.

4.4　クラフト不等式の証明

まず, J を任意の正の整数として, 次の関係式が成り立つことに注意する.

$$\left(\sum_{x \in \mathcal{A}} K^{-l(x)}\right)^J \;=\; \sum_{x_1 \in \mathcal{A}} \sum_{x_2 \in \mathcal{A}} \cdots \sum_{x_J \in \mathcal{A}} K^{-(l(x_1) + l(x_2) + \cdots + l(x_J))} \tag{131}$$

そこで, $A_J(n)$ を全体の長さが n になるような J 個の符号語の並べ方の総数とする. つまり,

$$A_J(n) \;=\; \#\left\{x_1 x_2 \cdots x_J \in \mathcal{A}^J : l(x_1) + l(x_2) + \cdots l(x_J) = n\right\} \tag{132}$$

とする. ここで, $\exists_{x_1 \cdots x_J} \in \mathcal{A}^J$ に対し,

$$\sum_{n=1}^{J_{l_{\max}}} \delta\left(l(x_1) + l(x_2) + \cdots + l(x_J), n\right) \;=\; 1 \tag{133}$$

に注意して ($J_{l_{\max}}$ は J 個の符号語を並べた符号長の最大値), (131) 式の右辺を書き直すと

$$\sum_{n=1}^{J_{l_{\max}}} \sum_{x_1 \in \mathcal{A}} \cdots \sum_{x_J \in \mathcal{A}} \delta\left(l(x_1) + l(x_2) + \cdots + l(x_J), n\right) K^{-n} \;=\; \sum_{n=1}^{J_{l_{\max}}} A_J(n) K^{-n} \tag{134}$$

となる. 従って

$$\left(\sum_{x \in \mathcal{A}} K^{l(x)}\right)^J \;=\; \sum_{n=1}^{J_{l_{\max}}} A_J(n) K^{-n} \tag{135}$$

が成り立つ. ところで, $\mathcal{B}$ の記号を n 個並べてできる語の総数が K^n であるから,

$$A_J(n) \;>\; K^n \tag{136}$$

であると仮定すると, $\exists_{a,a'} \in \mathcal{A}$ に対して, $\phi(a) = \phi(a^{'})$ となってしまい, ϕ が一意復号可能であることに反する. よって

$$A_J(n) \;\leq\; K^n \tag{137}$$

であることが必要. 従って, $A_J(n) K^{-n} \leq 1$ より, (131) 式は

$$\left(\sum_{x \in \mathcal{A}} K^{-l(x)}\right)^J \;=\; \sum_{n=1}^{J_{l_{\max}}} A_J(n) K^{-n} \leq \sum_{n=1}^{J_{l_{\max}}} 1 = J_{l_{\max}} \tag{138}$$

すなわち, $J \to \infty$ の極限で

$$\sum_{x \in \mathcal{A}} K^{-l(x)} \;\leq\; (J_{l_{\max}})^{1/J} = 1 \tag{139}$$

となり, クラフト不等式が成立する.

4.5　クラフト不等式を用いた語頭符号の構成

$\mathcal{B}=\{0,1\}$ $(K=2)$, $l_1=1, l_2=2, l_3=l_4=3$ であるとする．このとき，クラフト不等式は

$$2^{-1}+2^{-2}+2^{-3}+2^{-3} \;=\; \frac{1}{2}+\frac{1}{4}+\frac{1}{8}\times 2=1\le 1 \tag{140}$$

となり，クラフト不等式を満たす．$x_1=0\,(l_1=1)$ と決めると，x_2 は 10 か 01 のいずれかであるので，ここでは $x_2=10\,(l_2=2)$ とする．ここで，クラフト不等式から $2\le 2^3-2-2^2=2$ であるから，x_3, x_4 としては $\underline{0}01, \underline{0}01, \underline{0}10, \underline{0}11$ の 4 つと，$\underline{10}0, \underline{10}1$ の 2 つは使えない．これを**符号の木**で表すと図 14 のようになる．従って，x_3, x_4 は $x_3=111, x_4=110$ となればよい．

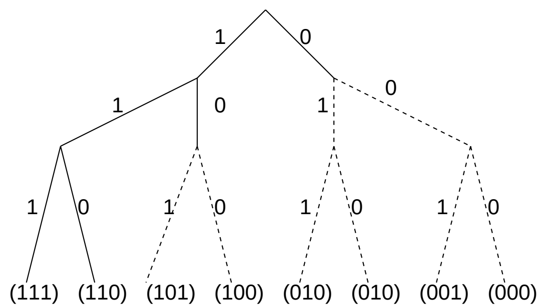

図 14: 符号の木．頭語条件を考えると，図の破線の部分の枝は該当しない (頭語条件を満たさない)．従って，x_3, x_4 として可能な符号は $x_3=111$ と $x_4=110$ である．

今回学んだ事項を確認しておくために次の**例題 6** を見ておこう．

例題 6
$K=3, \mathcal{B}=\{0,1,2\}, l_1=1, l_2=2, l_3=l_4=3$ の符号語に関して以下の問いに答えよ．

(1) この符号は一意復号可能かどうかを判定せよ (そのように判定した理由も明記すること)．
(2) (1) で判定した結果，一意復号可能である場合，語頭条件を満たすような符号語 x_1, x_2, x_3, x_4 を一つ挙げよ．
(3) 教科書 p. 35 の図 3.4 にならって，この符号語の符号の木を描け．

(解答例)

(1) クラフト不等式が満たされているかどうかを調べればよい．$K=3, l_1=1, l_2=2, l_3=l_4=3$ であるから，クラフト不等式は

$$3^{-1}+3^{-2}+3^{-3}+3^{-3} \;=\; \frac{1}{3}+\frac{1}{9}+\frac{1}{27}+\frac{1}{27}=\frac{14}{27}<1 \tag{141}$$

となり満たされている．よってこの符号は一意復号可能である．

(2) 例えば $\underline{x_1=0}$ と選ぶと，語頭条件から $l_2=2$ である x_2 として 01, 00, 02 の 3 つは使えない．従って，$3^2=9$ 通りの可能性の中で，x_2 として用いることのできるのは 10, 11, 12, 20, 21, 22 の 6 通りである．ここでは $\underline{x_2=20}$ と選ぶことにしよう．符号長が 3 である x_3, x_4 に対しては語頭条件のために $3^3=27$

通りの可能性の中で 0 を先頭に持つ 000, 001, 002, 010, 011, 012, 020, 021, 022, 及び, 20 を先頭に持つ 200, 201, 202 の計 12 通りの符号は用いることができない．これ以外の符号長 3 の符号として例えば $\underline{x_3 = 100}$, $\underline{x_4 = 101}$ のように選ぶことができる．よって上記を表にまとめれば

x_1	0
x_2	20
x_3	110
x_4	101

となる．

(3) 符号の木を描くと図 15 のようになる．

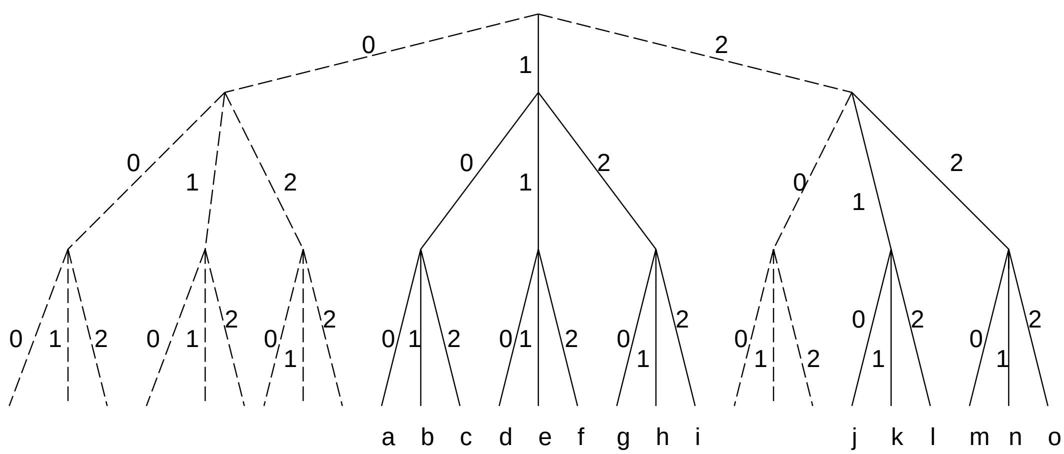

図 15: この問題での符号語に対する符号の木．$x_1 = 0, x_2 = 20$ のように符号を選んだ場合, x_3, x_4 として採用できる長さ 3 の符号としては図の a〜o の中のいずれかとなる．(破線でたどられる符号を用いることはできない)

演習問題 4

演習問題 3 において, $P_{A,B}$ と $P_A \cdot P_B$ との間の KL 情報量：

$$D(P_{A,B}||P_A \cdot P_B) \quad = \quad \sum_{x=\{\mathrm{H},\mathrm{T}\}} \sum_{y=\{\mathrm{H},\mathrm{T}\}} P_{A,B}(x,y) \log \left\{ \frac{P_{A,B}(x,y)}{P_A(x) P_B(y)} \right\} \tag{142}$$

を求め, $I(A;B) = D(P_{A,B}||P_A \cdot P_B)$ が成立するか否かを確かめよ．