通信路容量

通信路容量

$$A = \{a_1, \cdots, a_r\} \quad \xrightarrow{X} \boxed{\text{通信路}} \xrightarrow{Y} \quad B = \{b_1, \cdots, b_s\}$$

$$T = \{p_{ij}\}$$

通信路行列（r行，s列）

r行

$$p_{ij} = P(b_j | a_i)$$

s列

- （通信路Tによる）XとYの相互情報量：

$$I(X;Y) = H(Y) - H(Y|X) = -\sum_j q_j \log_2 q_j + \sum_{i,j} p_i p_{ij} \log_2 p_{ij}$$

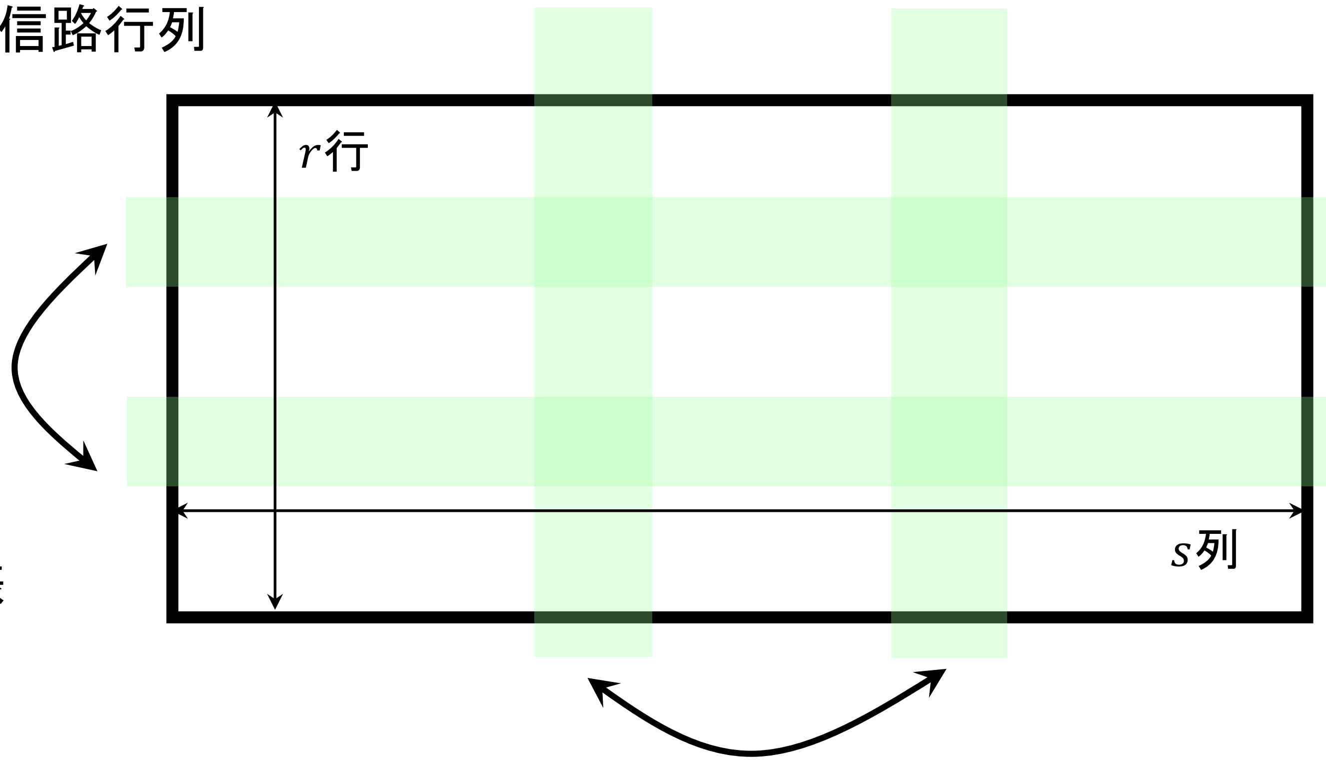
- 通信路Tの通信路容量：

$$C = \max_{\boldsymbol{p}} I(X;Y)$$

ここで，$\boldsymbol{p} = (p_1, \cdots, p_r)$は入力の確率分布

記憶のない一様通信路の通信路容量

通信路行列



どの行をとっても集合として同じ

入力について一様

どの列をとっても集合として同じ

出力について一様

記憶のない一様通信路の通信路容量

入力について一様な通信路の通信路容量

$$
\begin{aligned}
C &= \max_{\boldsymbol{p}} I(X;Y) \\
&= \max_{\boldsymbol{p}} \left(-\sum_{j=1}^{s} q_j \log_2 q_j + \sum_{i,j} p_i p_{ij} \log_2 p_{ij} \right) \\
&= \max_{\boldsymbol{p}} \left(-\sum_{j=1}^{s} q_j \log_2 q_j + \boxed{\sum_{i=1}^{r} \left(p_i \cdot \left(\sum_{j=1}^{s} p_{ij} \log_2 p_{ij} \right) \right)} \right) \\
&= \max_{\boldsymbol{p}} \left(H(Y) + \boxed{\sum_{i=1}^{r} \left(p_i \cdot \left(\sum_{j=1}^{s} p_{1j} \log_2 p_{1j} \right) \right)} \right) \\
&= \max_{\boldsymbol{p}} H(Y) + \sum_{j=1}^{s} p_{1j} \log_2 p_{1j}
\end{aligned}
$$

$\{p_{\alpha 1}, \cdots, p_{\alpha s}\} = \{p_{\beta 1}, \cdots, p_{\beta s}\}$だから

$$
\begin{aligned}
&\sum_{j=1}^{s} p_{\alpha j} \log_2 p_{\alpha j} \\
&\quad = \sum_{j=1}^{s} p_{\beta j} \log_2 p_{\beta j} \\
&\quad = \sum_{j=1}^{s} p_{1j} \log_2 p_{1j}
\end{aligned}
$$

$$
\begin{aligned}
&\sum_{i=1}^{r} \left(p_i \cdot \left(\sum_{j=1}^{s} p_{1j} \log_2 p_{1j} \right) \right) \\
&\quad = \left(\sum_{j=1}^{s} p_{1j} \log_2 p_{1j} \right) \cdot \left(\sum_{i=1r} p_i \right) \\
&\quad = \sum_{j=1}^{s} p_{1j} \log_2 p_{1j}
\end{aligned}
$$

記憶のない一様通信路の通信路容量

さらに出力について一様であれば

$$C = \max_{\boldsymbol{p}} H(Y) + \sum_{j=1}^{s} p_{1j} \log_2 p_{1j}$$
$$= \log_2 s + \sum_{j=1}^{s} p_{1j} \log_2 p_{1j}$$

なぜならば,

(1)　一般に,

$$\max_{\boldsymbol{p}} H(Y) = -(q_1 \log_2 q_1 + \cdots + q_s \log_2 q_s) \le \log_2 s$$

かつ，等号が成立するのは，$q_1 = \cdots = q_s = \frac{1}{s}$のときである.

(2) 実際に，統合が成立するときがあるか？　→　答えは，YES

$p_1 = \cdots = p_r = \frac{1}{r}$のとき，$q_1 = \cdots = q_s = \frac{1}{s}$となる.

(3) なぜならば，任意のj, kについて,

$$q_j = \sum_{i=1}^{r} p(a_i) p(b_j|a_i) = \sum_{i=1}^{r} p_i p_{ij} = \sum_{i=1}^{r} \frac{1}{r} p_{ij} = \frac{1}{r} \sum_{i=1}^{r} p_{ij}$$

$$q_k = \sum_{i=1}^{r} p(a_i) p(b_k|a_i) = \sum_{i=1}^{r} p_i p_{ik} = \sum_{i=1}^{r} \frac{1}{r} p_{ik} = \frac{1}{r} \sum_{i=1}^{r} p_{ik}$$

となるが，$\{p_{1j}, \cdots, p_{rj}\} = \{p_{1k}, \cdots, p_{rk}\}$であるので，$\sum_{i=1}^{r} p_{ij} = \sum_{i=1}^{r} p_{ik}$,

すなわち$q_j = q_k = \frac{1}{s}$が導かれる.

記憶のない一様通信路の通信路容量

- 通信路が入力について一様な場合

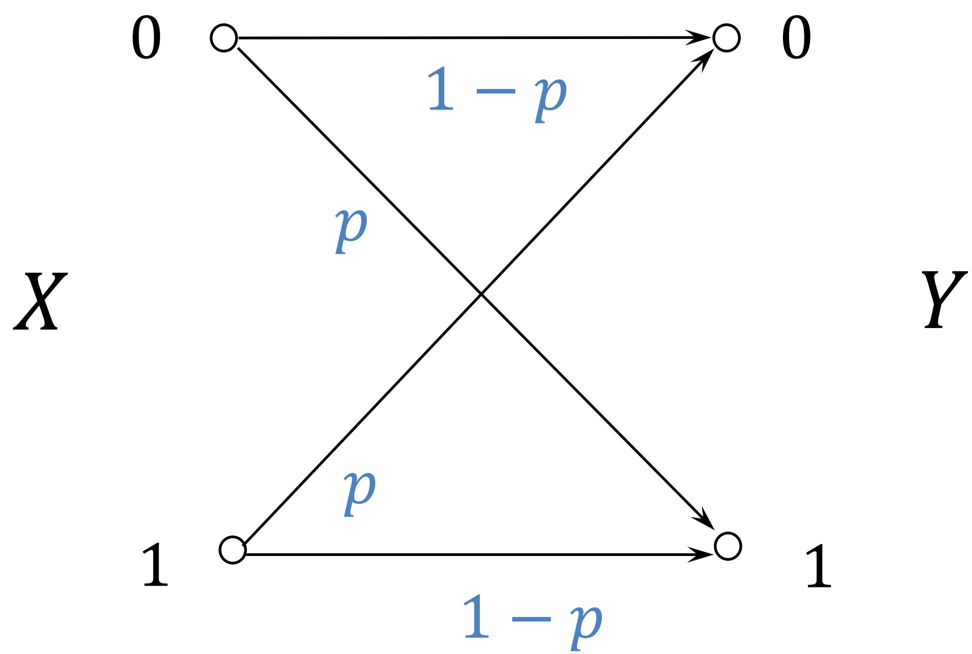
$$\max_{\boldsymbol{p}} H(Y) + \sum_{j=1}^{s} p_{1j} \log_2 p_{1j}$$

- さらに，通信路が出力について一様な場合

$$C = \log_2 s + \sum_{j=1}^{s} p_{1j} \log_2 p_{1j}$$

さまざまな通信路の通信路容量

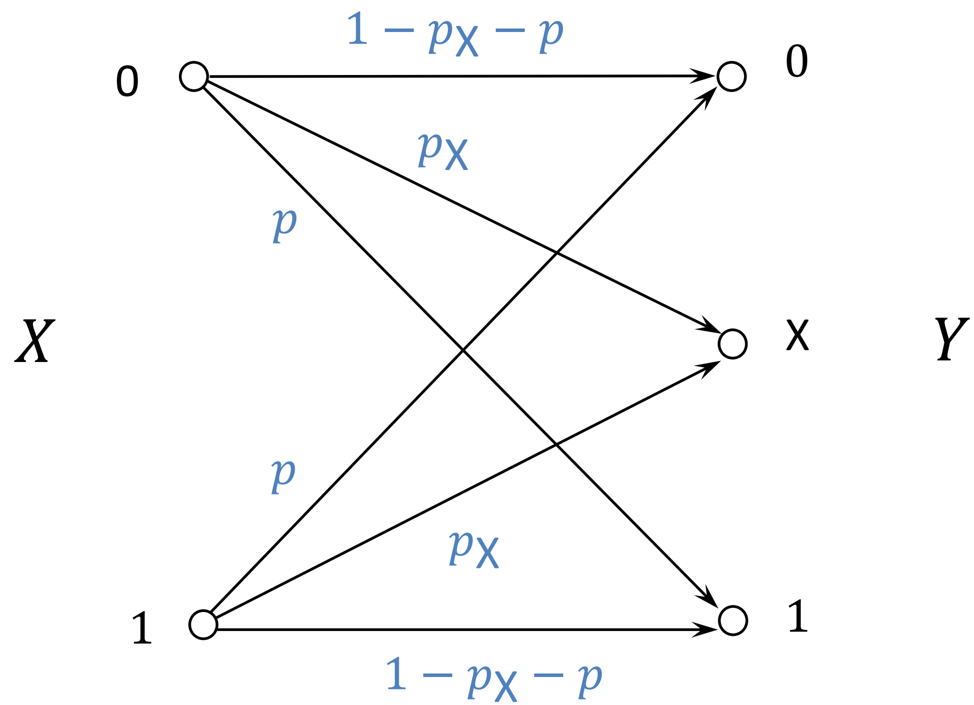
2元対称通信路



$$C = \log_2 2 + \sum_{j=1}^{2} p_{1j} \log_2 p_{1j} = 1 + p \log_2 p + (1-p) \log_2 (1-p) = 1 - \mathcal{H}(p)$$

さまざまな通信路の通信路容量

2元対称消失通信路



$$T=\begin{bmatrix} 1-p_\mathrm{X}-p & p_\mathrm{X} & p \\ p & p_\mathrm{X} & 1-p_\mathrm{X}-p \end{bmatrix}$$

さまざまな通信路の通信路容量

□ 通信路:$T=\begin{bmatrix} 1-p_{\mathsf{X}}-p & p_{\mathsf{X}} & p \\ p & p_{\mathsf{X}} & 1-p_{\mathsf{X}}-p \end{bmatrix}$の通信容量

(1) 入力について一様であるので,

$$\begin{aligned} C &= \max_{\boldsymbol{p}} H(Y) + \textstyle\sum_{j=1}^{s} p_{1j} \log_2 p_{1j} \\ &= \max_{\boldsymbol{p}} H(Y) + (1-p_{\mathsf{X}}-p)\log_2(1-p_{\mathsf{X}}-p) + p_{\mathsf{X}}\log_2 p_{\mathsf{X}} + p\log_2 p \end{aligned}$$

(2) つまり,

$$H(Y) = -\left(q_0 \log_2 q_0 + q_{\mathsf{X}} \log_2 q_{\mathsf{X}} + q_1 \log_2 q_1\right)$$

の最大値がわかればよい.

さまざまな通信路の通信路容量

(3) 変数の動く範囲：入力で0の起きる確率をx，1の起きる確率をyとすると，

$$x + y = 1, \qquad 0 \le \{x, y\} \le 1$$

(4) q_0, q_{X}, q_1をx, yで表す．

$$\begin{cases} q_0 &=& (1 - p_{\mathsf{X}} - p)x + py \\ q_{\mathsf{X}} &=& p_{\mathsf{X}}x + p_{\mathsf{X}}y = p_{\mathsf{X}} \\ q_1 &=& (1 - p_{\mathsf{X}} - p)y + px \end{cases}$$

(5) ラグランジュの未定乗数法を適用する：

$$J = -\Big((1 - p_{\mathsf{X}} - p)x + py\Big)\log_2\Big((1 - p_{\mathsf{X}} - p)x + py\Big) - q_{\mathsf{X}}\log_2 q_{\mathsf{X}}$$
$$-\Big((1 - p_{\mathsf{X}} - p)y + px\Big)\log_2\Big((1 - p_{\mathsf{X}} - p)y + px\Big) + \lambda(x + y - 1)$$

と置いて

$$\frac{\partial J}{\partial x} = \frac{\partial J}{\partial y} = \frac{\partial J}{\partial \lambda} = 0$$

となるx, y, λを求める．

さまざまな通信路の通信路容量

(6) 式を変形して，

$$\frac{\partial J}{\partial x} = -(1-p_\chi-p)\log_2\left((1-p_\chi-p)x+py\right)-\frac{1-p_\chi-p}{\ln 2}$$
$$-p\log_2\left((1-p_\chi-p)y+px\right)-\frac{p}{\ln 2}+\lambda$$

$$\frac{\partial J}{\partial y} = -p\log_2\left((1-p_\chi-p)x+py\right)-\frac{p}{\ln 2}$$
$$-(1-p_\chi-p)\log_2\left((1-p_\chi-p)y+px\right)-\frac{1-p_\chi-p}{\ln 2}+\lambda$$

$$\frac{\partial J}{\partial z} = x+y-1$$

から$x=y$が得られる.

(7) さらに，$q_0=q_1=\frac{1-p_\chi}{2}$が得られる.

(8) 以上より，

$$\max_{x+y=1,0\le\{x,y\}\le 1} H(Y) = -\left(p_\chi\log_2 p_\chi+(1-p_\chi)\log_2\frac{1-p_\chi}{2}\right)$$

さまざまな通信路の通信路容量

(8) 通信路容量については，

$$C = -\left(p_X \log_2 p_X + (1 - p_X) \log_2 \frac{1 - p_X}{2} \right)$$

$$+ \left(1 - p_X - p\right) \log_2 \left(1 - p_X - p\right) + p_X \log_2 p_X + p \log_2 p$$

$$= (1 - p_X) \left(1 - \log_2 \left(1 - p_X\right)\right) + \left(1 - p_X - p\right) \log_2 \left(1 - p_X - p\right) + p \log_2 p$$

$$= 1 - p_X + p \log_2 \frac{p}{1 - p_X} + \left(1 - p_X - p\right) \log_2 \frac{1 - p_X - p}{1 - p_X}$$

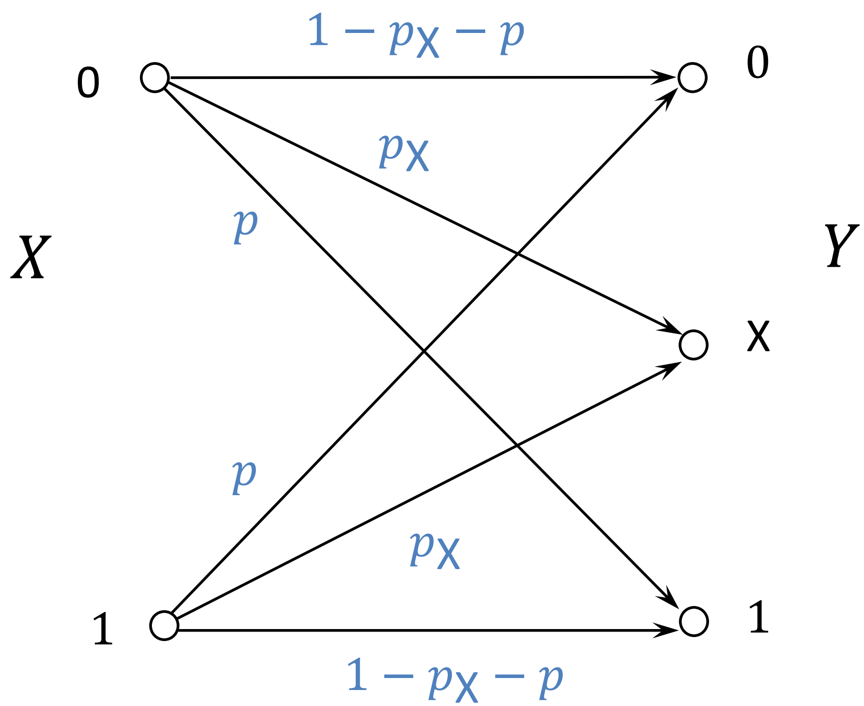
$$= \left(1 - p_X\right) \left(1 + \frac{p}{1 - p_X} \log_2 \frac{p}{1 - p_X} + \left(1 - \frac{p}{1 - p_X}\right) \log_2 \left(1 - \frac{p}{1 - p_X}\right)\right)$$

$$= \left(1 - p_X\right) \left(1 - \mathcal{H}\left(\frac{p}{1 - p_X}\right)\right)$$

さまざまな通信路の通信路容量

まとめ：　2元対称消失通信路

$$T = \begin{bmatrix} 1-p_{\mathsf{X}}-p & p_{\mathsf{X}} & p \\ p & p_{\mathsf{X}} & 1-p_{\mathsf{X}}-p \end{bmatrix}$$

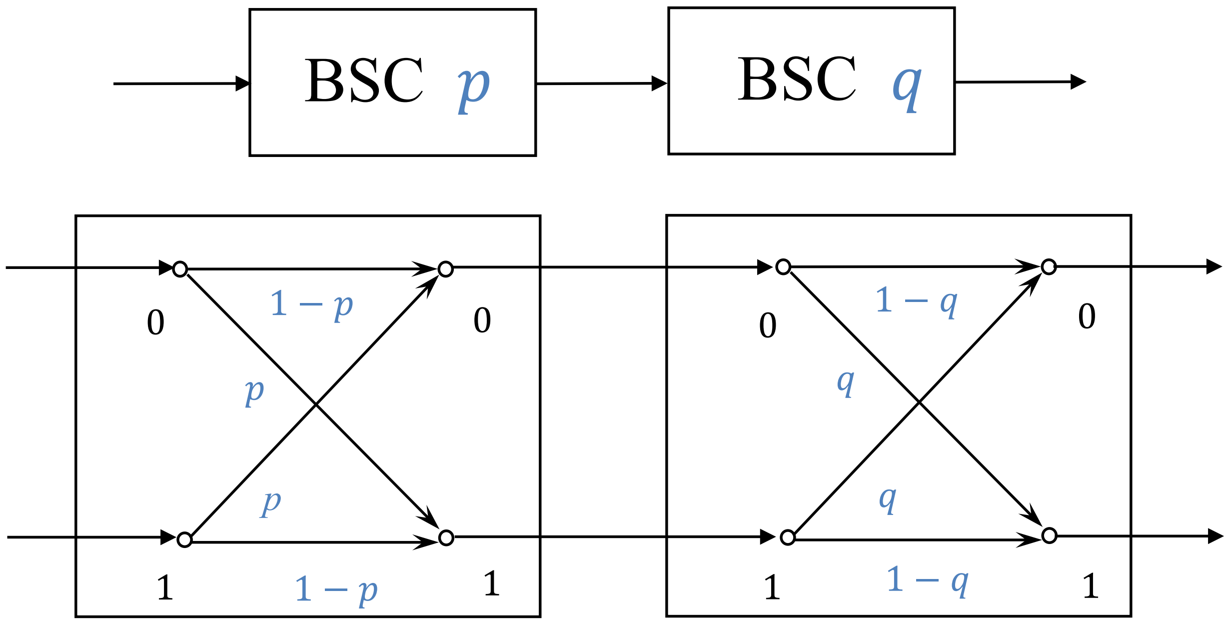


$$C = (1-p_{\mathsf{X}})\left(1-\mathcal{H}\left(\frac{p}{1-p_{\mathsf{X}}}\right)\right)$$

$$= (1-p_{\mathsf{X}})\big(1-\log_2(1-p_{\mathsf{X}})\big) + \big(1-p_{\mathsf{X}}-p\big)\log_2\big(1-p_{\mathsf{X}}-p\big) + p\log_2 p$$

さまざまな通信路の通信路容量

2元対称通信路の直列接続



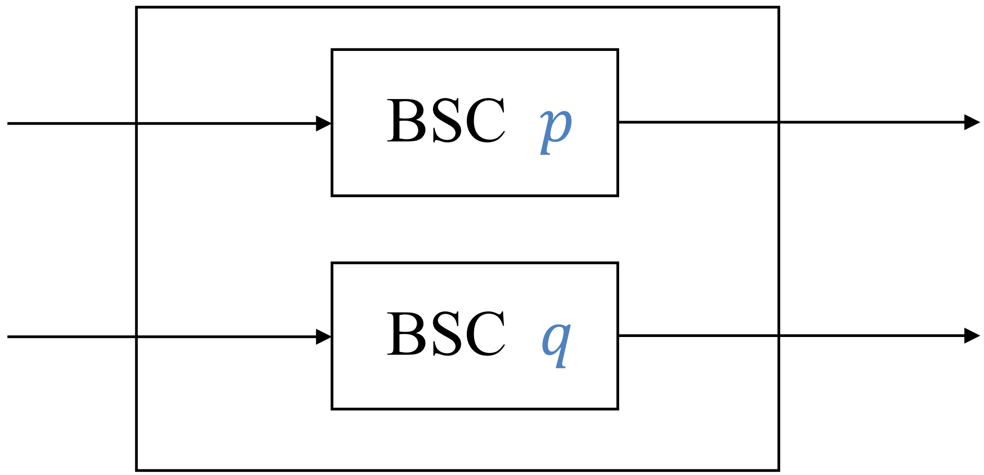
$$\begin{bmatrix} 1-p & p \\ p & 1-p \end{bmatrix} \cdot \begin{bmatrix} 1-q & q \\ q & 1-q \end{bmatrix} = \begin{bmatrix} 1-p-q+2pq & p+q-2pq \\ p+q-2pq & 1-p-q+2pq \end{bmatrix}$$ ⇒ BSC

誤り率pのBSCの通信路容量は$1-\mathcal{H}(p)$であるので

$$C = 1-\mathcal{H}(p+q-2pq) = 1-\mathcal{H}\big(1-(p+q-2pq)\big) = 1-\mathcal{H}(1-p-q+2pq)$$

さまざまな通信路の通信路容量

2元対称通信路の並列接続



$$\begin{bmatrix} (1-p)(1-q) & (1-p)q & p(1-q) & pq \\ (1-p)q & (q-p)(1-q) & pq & p(1-q) \\ p(1-q) & pq & (1-p)(1-q) & (1-p)q \\ pq & p(1-q) & (1-p)q & (1-p)(1-q) \end{bmatrix}$$ ⇒　2重に一様

さまざまな通信路の通信路容量

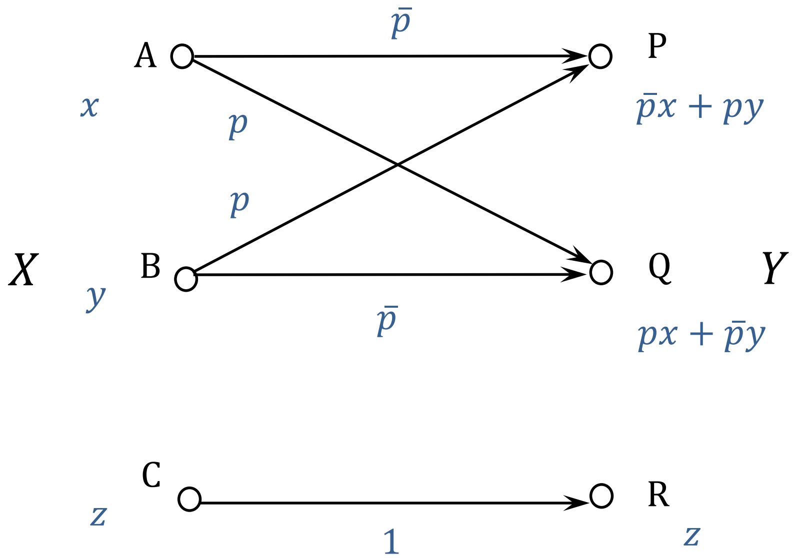
2元対称通信路の並列接続（続き）

$$
\begin{aligned}
C &= \log_2 4 + \sum_{j=1}^{4} p_{1j} \log_2 p_{1j} \\
&= 2 + \bar{p}\bar{q}\log_2(\bar{p}\bar{q}) + \bar{p}q\log_2(\bar{p}q) + p\bar{q}\log_2(p\bar{q}) + pq\log_2(pq) \\
&= 2 + \bar{p}\bar{q}(\log_2\bar{p} + \log_2\bar{q}) + \bar{p}q(\log_2\bar{p} + \log_2 q) \\
&\qquad + p\bar{q}(\log_2 p + \log_2\bar{q}) + pq(\log_2 p + \log_2 q) \\
&= 2 + p\log_2 p + \bar{p}\log_2\bar{p} + q\log_2 q + \bar{q}\log_2\bar{q} \\
&= 2 - \mathcal{H}(p) - \mathcal{H}(q)
\end{aligned}
$$

⇒　並列接続されたBSCの通信路容量は，それぞれのBSCの通信路容量の和に等しい

$$
C = 2 - \mathcal{H}(p) - \mathcal{H}(q) = \left(1 - \mathcal{H}(p)\right) + (1 - \mathcal{H}(q))
$$

さまざまな通信路の通信路容量



$$T = \begin{bmatrix} \bar{p} & p & 0 \\ p & \bar{p} & 0 \\ 0 & 0 & 1 \end{bmatrix}$$

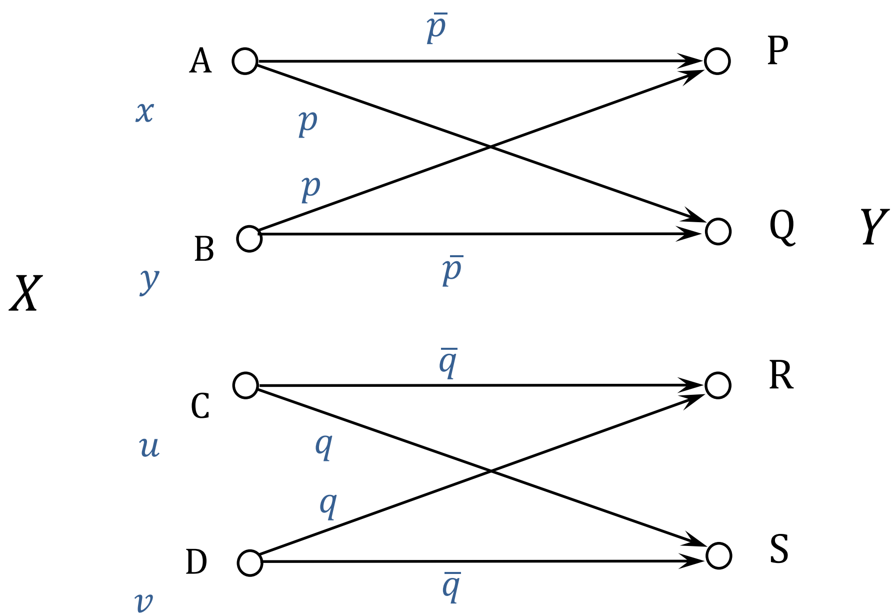
ただし，$\bar{p} \triangleq 1 - p$

A,B,Cの生起確率をx, y, zとする　　$0 \leq \{x, y, z\} \leq 1, x + y + z = 1$

$$\begin{aligned} I(X;Y) &= H(Y) - H(Y|X) \\ &= (\bar{p}x + py)\log_2 \frac{1}{\bar{p}x + py} + (\bar{p}y + px)\log_2 \frac{1}{px + \bar{p}y} + z\log_2 \frac{1}{z} - (x + y)\mathcal{H}(p) \end{aligned}$$

ラグランジュの未定乗数法により，　$C = \log_2\left(1 + 2^{1-\mathcal{H}(p)}\right)$

さまざまな通信路の通信路容量



$$T = \begin{bmatrix} \bar{p} & p & & \\ p & \bar{p} & & \\ & & \bar{q} & q \\ & & q & \bar{q} \end{bmatrix}$$

ただし，$\bar{p} \triangleq 1 - p, \bar{q} \triangleq 1 - q$

A, B, C, Dの生起確率をx, y, u, vとする

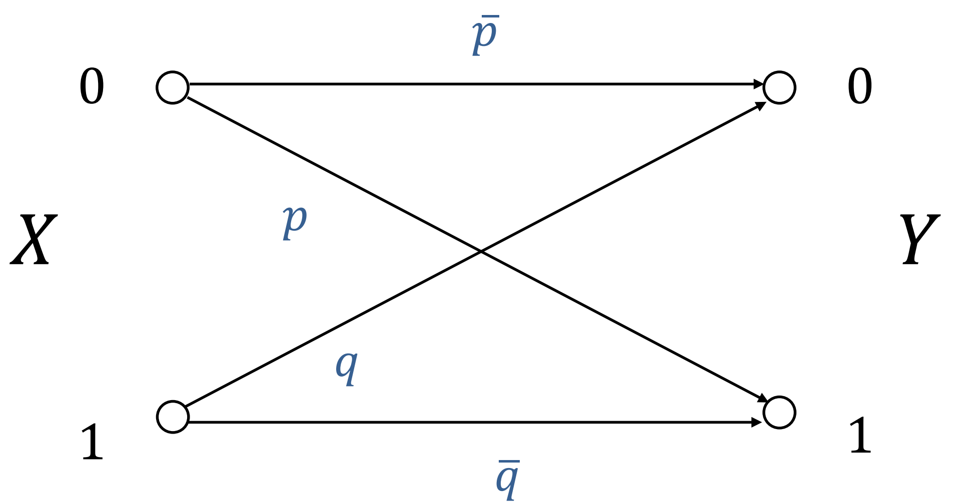
$$0 \le \{x, y, u, v\} \le 1, x + y + u + v = 1$$

$$\begin{aligned} I = -\big(& (\bar{p}x + py)\log_2(\bar{p}x + py) + (px + \bar{p}y)\log_2(px + \bar{p}y) \\ & + (\bar{q}u + qv)\log_2(\bar{q}u + qv) + (qu + \bar{q}v)\log_2(qu + \bar{q}v)\big) \\ & -\big((x + y)\mathcal{H}(p) + (u + v)\mathcal{H}(q)\big) \end{aligned}$$

ラグランジュの未定乗数法により，$\log_2\left(2^{1-\mathcal{H}(p)} + 2^{1-\mathcal{H}(q)}\right)$

さまざまな通信路の通信路容量

2元通信路一般の場合



$$T = \begin{bmatrix} \bar{p} & p \\ q & \bar{q} \end{bmatrix}$$

さまざまな通信路の通信路容量

- 出力アルファベット 1の発生確率をy とすると，出力側のエントロピーは$H(Y) = \mathcal{H}(y)$である．

- 入力アルファベット1の発生確率をxとすると，$y = (1 - x)p + x\bar{q}$
ゆえに，$x = \frac{y-p}{\bar{q}-p}$.

- xが0~1の間を動くとき，yは$p \sim \bar{q}$を動く（pと$\bar{q}$の大小関係はパラメータ次第）

- この通信路の条件付エントロピー$H(Y|X)$と相互情報量$I(X;\ Y)$は次のようになる．

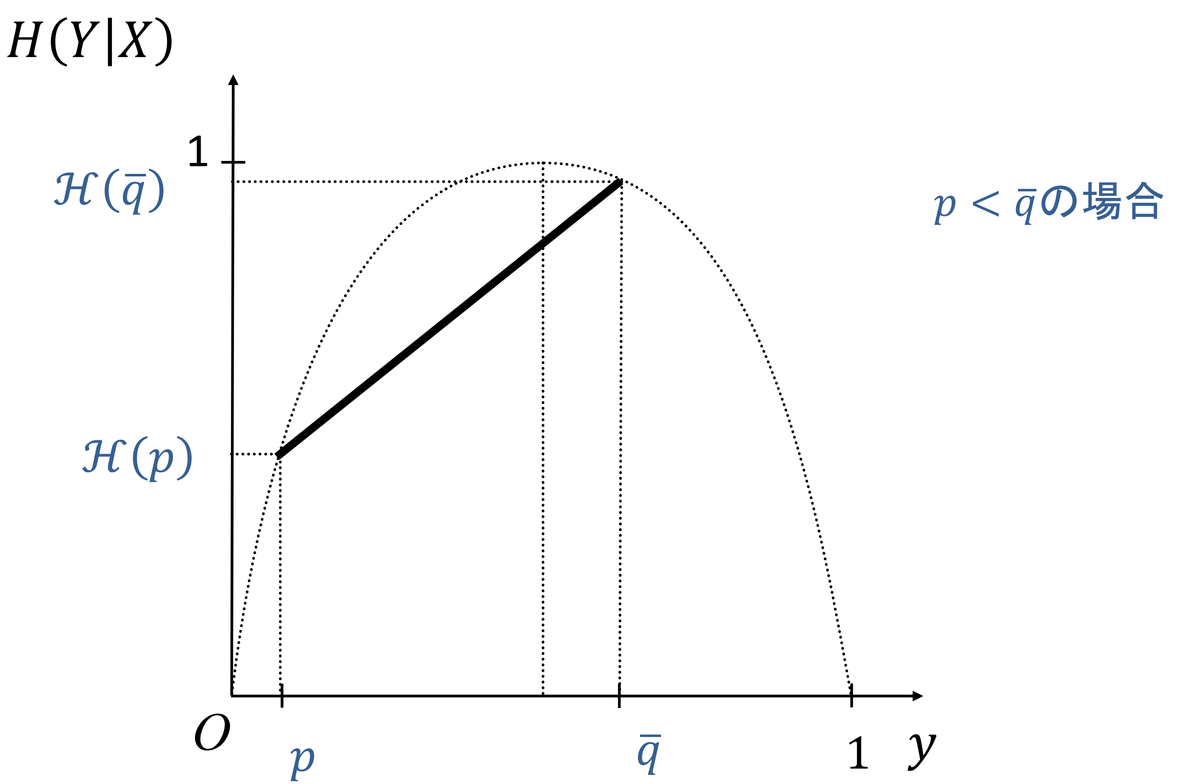
$$H(Y|X) = (1 - x)\mathcal{H}(p) + x\mathcal{H}(\bar{q}) = \frac{\bar{q} - y}{\bar{q} - p}\mathcal{H}(p) + \frac{y - p}{\bar{q} - p}\mathcal{H}(\bar{q})$$

$$I(X;Y) = H(Y) - H(Y|X) = \mathcal{H}(y) - \left(\frac{\bar{q} - y}{\bar{q} - p}\mathcal{H}(p) + \frac{y - p}{\bar{q} - p}\mathcal{H}(\bar{q})\right)$$

さまざまな通信路の通信路容量

$$H(Y|X) = (1-x)\mathcal{H}(p) + x\mathcal{H}(\bar{q}) = \frac{\bar{q}-y}{\bar{q}-p}\mathcal{H}(p) + \frac{y-p}{\bar{q}-p}\mathcal{H}(\bar{q})$$

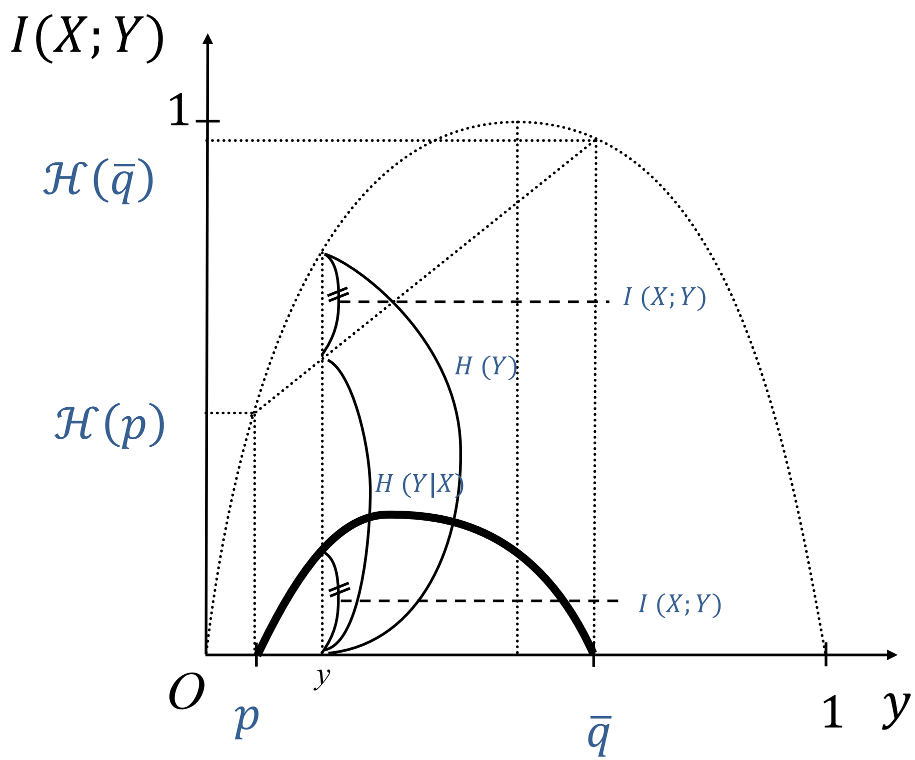
はyの関数　⇒　どのような関数？



さまざまな通信路の通信路容量

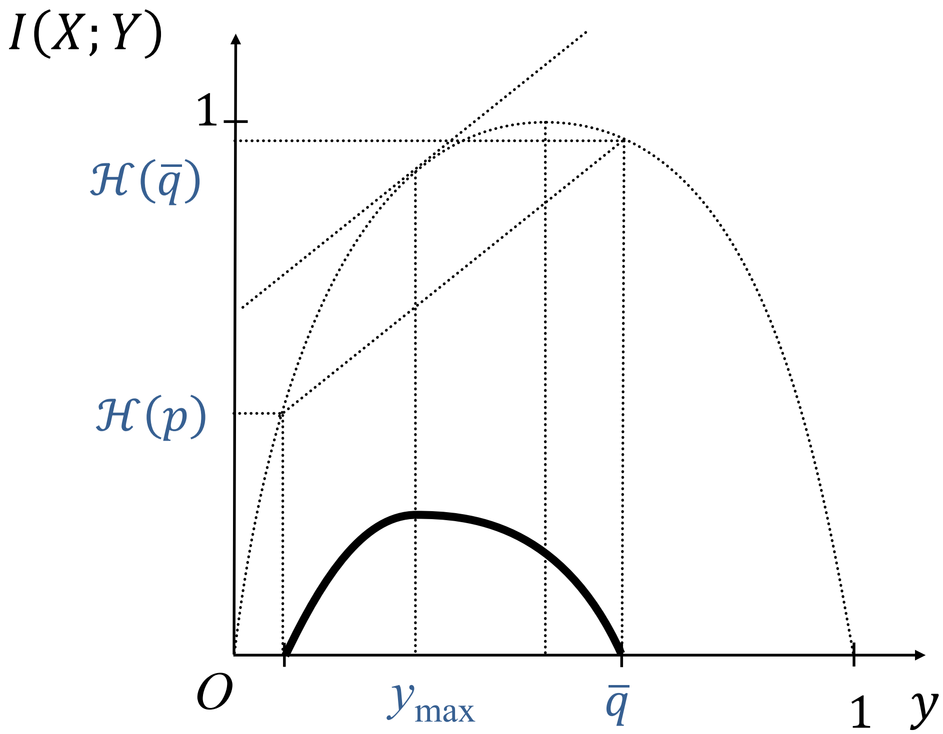
$$I(X;Y) = H(Y) - H(Y|X) = \mathcal{H}(y) - \left(\frac{\bar{q}-y}{\bar{q}-p}\mathcal{H}(p) + \frac{y-p}{\bar{q}-p}\mathcal{H}(\bar{q})\right)$$

はyの関数　⇒　どのような関数？



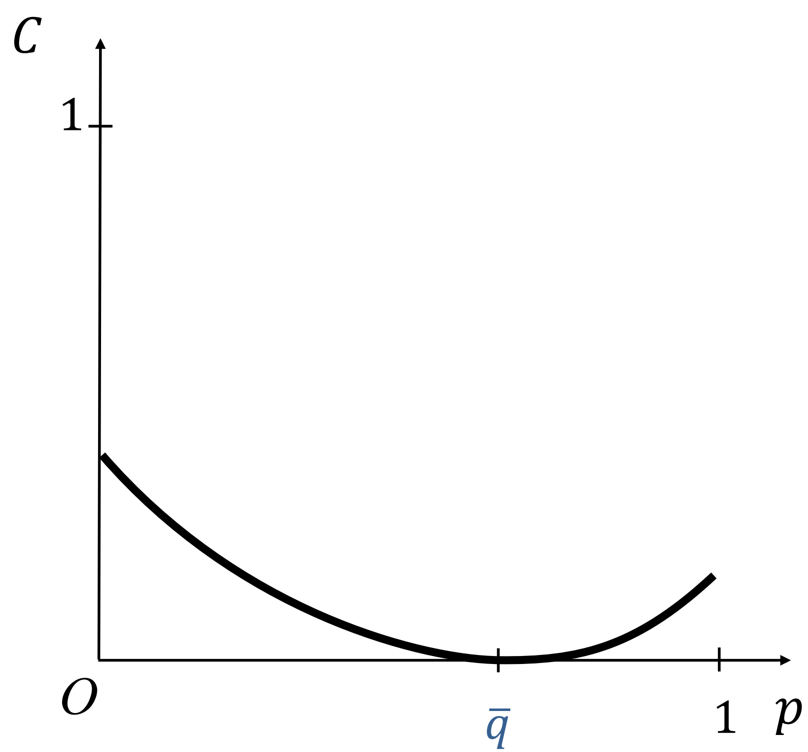
さまざまな通信路の通信路容量

$I(X;Y)$の最大値と，それを与えるyの値



さまざまな通信路の通信路容量

通信路容量Cとpの関係

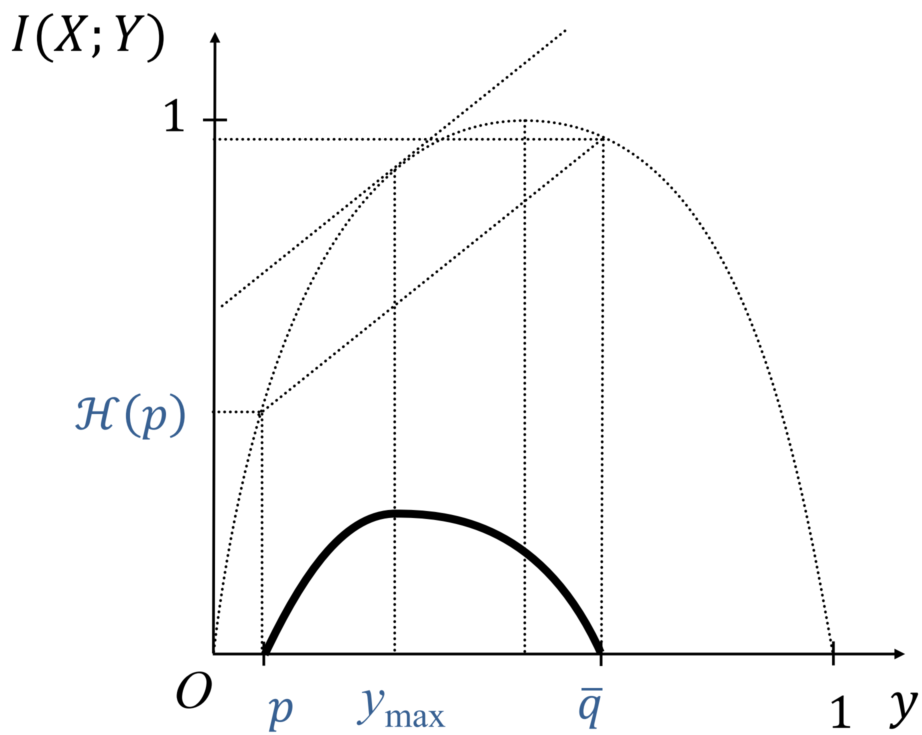


さまざまな通信路の通信路容量

- $p < \bar{q}$であれば，$p \le y \le \bar{q}$であり，ある$p \le y_{\max} \le \bar{q}$で最大値：

$$\mathcal{H}(y_{\max}) - \left(\frac{\bar{q} - y_{\max}}{\bar{q} - p} \mathcal{H}(p) + \frac{y_{\max} - p}{\bar{q} - p} \mathcal{H}(\bar{q}) \right)$$

をとる．この値が通信路容量となる．



さまざまな通信路の通信路容量

通信路 $T = \begin{bmatrix} \bar{p} & p \\ q & \bar{q} \end{bmatrix}$ $(\bar{p} \triangleq 1-p, \bar{q} \triangleq 1-q)$の通信路容量（定量解）

入力シンボルの生起確率をx, y $(0 \le \{x, y\} \le 1, x+y=1)$とすると，相互情報量は

$$I(X;Y) = -(\bar{p}x + qy)\log_2(\bar{p}x + qy) - (px + \bar{q}y)\log_2(px + \bar{q}y) - x\mathcal{H}(p) - y\mathcal{H}(q)$$

$\max\limits_{0 \le \{x,y\} \le 1, x+y=1} I(X;Y)$を求めるために，ラグランジュの未定乗数法により，

$$J = I(X;Y) - \lambda(x+y-1)$$

と置いて，

$$\frac{\partial J}{\partial x} = \frac{\partial J}{\partial y} = \frac{\partial J}{\partial \lambda} = 0$$

を満たすx, yを求めると，

$$x = \frac{q - \bar{q} \cdot 2^{\frac{\mathcal{H}(p)-\mathcal{H}(q)}{q-\bar{p}}}}{(p-\bar{q})\left(2^{\frac{\mathcal{H}(p)-\mathcal{H}(q)}{q-\bar{p}}} + 1\right)}, \qquad y = \frac{p \cdot 2^{\frac{\mathcal{H}(p)-\mathcal{H}(q)}{q-\bar{p}}} - \bar{p}}{(p-\bar{q})\left(2^{\frac{\mathcal{H}(p)-\mathcal{H}(q)}{q-\bar{p}}} + 1\right)}$$

そして，

$$C = \max_{0 \le \{x,y\} \le 1, x+y=1} I(X;Y) = \log_2\left(2^{\frac{\bar{q}\mathcal{H}(p)-p\mathcal{H}(q)}{q-\bar{p}}} + 2^{\frac{-q\mathcal{H}(p)+\bar{p}\mathcal{H}(q)}{q-\bar{p}}}\right)$$

まとめ

- 通信路は入力の確率分布を変化させたときの相互情報量の最大値であり，通信路に固有の値である.
- 一様性があるときや構造的な特徴があるときは，通信路容量は簡単に計算できる.
- 一般的には，最適化問題の一種として扱われ，ラグランジュの未定乗数法などで解けることもある.
- 記憶のない2元定常通信路一般の通信路容量は幾何学的に理解することができる.
- 記憶のない2元定常通信路一般の通信路容量は解析的にも求められる.