誤りのある通信路

山本昌志*

2007 年 11 月 30 日

概 要

誤りのある通信路について，理論的な取り扱いを行う．

1　本日の学習内容

通信の経路に誤りがある場合，情報量を定量的に取り扱う方法を学ぶ．学習のゴールは，以下の通り．

- 通信路ではノイズなどにより符号が変化することがある．そのようなことがイメージできる．
- 2 元対称通信路の確率を使った取り扱いが分かる．
- 結合エントロピーと条件付きエントロピー，相互情報量が分かる．

教科書 [1] の pp.65–68 が範囲である．教科書の通りの内容であるが，詳しく説明する．

2　誤りのある通信路

2.1　2 元対称通信路

図 1 のような状況を考える．送信者はメッセージ[1]を送り，受信者はそれを受け取る．通信路の途中に雑音源があり，メッセージが変わることがある．こういうことは，結構頻繁に生じる．遠距離の通信，例えば宇宙探査船が惑星の画像を送る場合，かなりの確率で誤りが生じる．微弱な電波で長距離を伝送するからである．これほどのことはないが，地上内の通信でも誤りの発生確率をゼロにすることはできない．熱雑音を避けることはできないからである．宇宙線が問題になることもある．

*独立行政法人　秋田工業高等専門学校　電気情報工学科

[1]メッセージは符号化されるので，符号を送るといった方が良い．

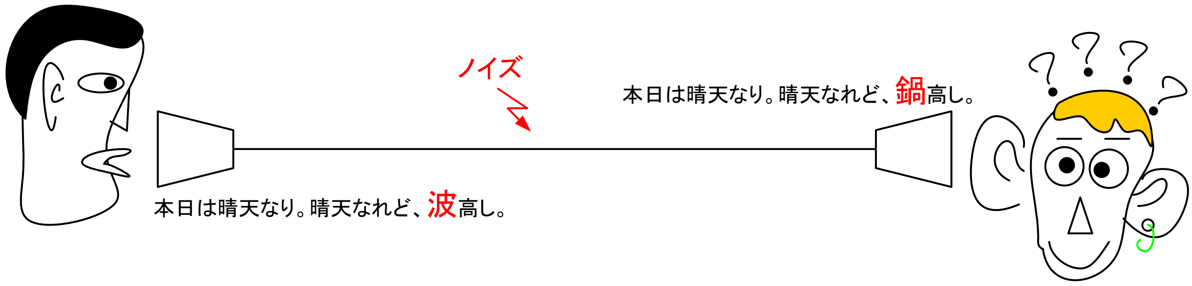

図 1: 誤りのある通信の例．糸電話で通信しているが，途中ノイズで**波**が**鍋**になっている．

これから，誤りのある通信路の理論的な考察を行う．ある程度の式を使うことになるが，意味さえ分かれば難しいことは何もない．

送信者は，メッセージ x_i を送る．もし，アルファベット 26 文字のいずれかを送るならば，$x_0 = \mathrm{a}$, $x_1 = \mathrm{b}$, $x_2 = \mathrm{c}, \cdots, x_{25} = \mathrm{z}$ と考える．また，0 か 1 のバイナリーデータを送るならば，$x_0 = 0$, $x_1 = 1$ である．ここで，送信者が文字 x_i を送る確率を $p(x_i)$ とする．すると，

$$\sum_i p(x_i) = 1 \tag{1}$$

である．送信者は，いずれかの文字を送るので，各々の確率を加えると 1 になるだけである．

メッセージを受け取る受信者の方も同じで，受け取る文字を y_i とする．ここでは，受け取る文字の確率を $p(y_i)$ とする．もちろん，いずれかの文字を受け取るので，その和は 1 になる．

ここで，文字数が多くなると議論が複雑になるので，もっとも単純な二元対称通信路 (binary symeetric chanel) を考える．これは，図 2 のような通信路で以下のように特徴づけられる．

- 通信に使う文字は $0, 1$ の 2 種類である．送信者は，これらの文字のみが送信できる．受信者は，これらの文字のみを受け取る．これが，二元対称通信路の「二元」ということを表している．

- 途中で誤りが発生する確率 q は，$0 \to 1$ も $1 \to 0$ も同じである．この場合，誤りが発生しない;$0 \to 0$ あるいは $1 \to 1$ で正しく通信できる確率は $1 - q$ となる．これが，二元対称通信路の「対称」ということを表している．

いつでもこのような状況が生じるわけではなく，このような理想的な状況を考える—ということを理解しておく必要がある．例えば，'0' あるいは'1' という文字を送ったが，それが途中で失われて，受信者には何も届かないこともある．また，'0' を送るときは 90[%] 正しく通信ができるが，'1' を送るときには 30[%] しか正しく通信できないこともある．これは対称ではない．

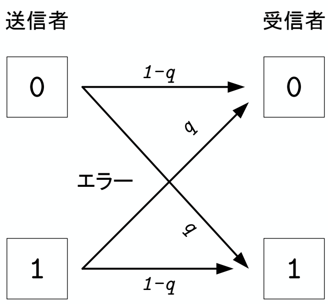

図 2: 二元対称性通信路．送信者，受信者とも $0, 1$ の符号を使う．通信途中で誤りが生じ，符号が入れ替わることがある．通信途中で誤りが生じる確率は q，正しく通信ができる確率 $1-q$ である．

この 2 元対称通信路の場合，もっとも「質の悪い」エラーとはどのようなものだろうか？ 100[%] の確率 ($q=1$) で，エラーが起きる場合か？ この場合，送られるメッセージと受け取るメッセージは，011010 → 100101 のようになる．受信者は，受け取ったメッセージから送ったメッセージが分かる．'0' と'1' を反転させれば良い．エラーが 0[%] の時と同じだけ情報を受け取ることができる．もっとも「質の悪い」エラーは 50[%] の確率 ($q=0.5$) でエラーが発生する場合である．この場合，受け取るメッセージはランダムに'0' と'1' が並んだ数列になり，送信者の送った情報を全く受け取ることができない[2]．

2.2 結合エントロピーと条件付きエントロピー

2.2.1 確率の定義

これから，誤りがある通信路の理論的な考察を行う．そのために，図 3 に示すシステムを考える．このシステムは，送信者 X がある文字 x_i を送ると，受信者は y_j を受け取る．途中，信号にノイズが入り，メッセージが変化する可能性がある．それぞれの確率 (probability) を以下のように定義する．

$p(x_i)$	送信者が文字 x_i を送信する確率.
$p(y_i)$	受信者が文字 y_i を受け取る確率.
$p(x_i, y_i)$	送信者が x_i を送り，受信者が x_i を受け取る確率.

注意　$p()$ は関数ではない．また，添え字の i や j は送信/受信順序を表すものではない．i と j は文字の区別を行っている．

言うまでもないと思うが，全ての確率を加えると 1 になる．したがって，先ほど示したそれぞれの確率は

$$\sum_i p(x_i) = 1 \qquad \sum_j p(y_j) = 1 \qquad \sum_i \sum_j p(x_i, y_j) = 1 \tag{2}$$

となる．また，

$$p(x_i) = \sum_j p(x_i, y_j) \qquad p(y_j) = \sum_i p(x_i, y_j) \tag{3}$$

[2]”情報を送った”という情報を受け取ることができるが，これは別の話．

という関係も直ちに分かる．

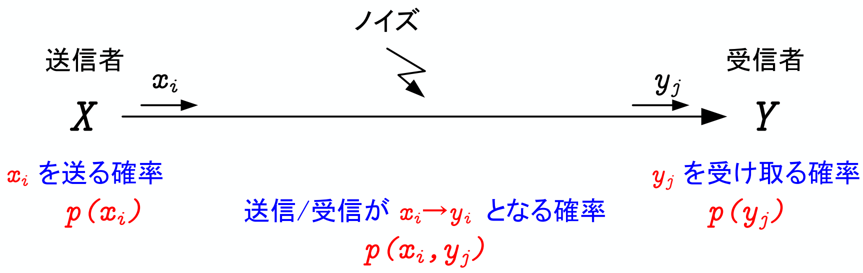

図 3: X から Y へメッセージを送る場合の確率．$p(x_i)$ は送信者が x_i を送る確率．$p(y_j)$ は受信者が y_j を受け取る確率．$p(x_i, y_j)$ は，メッセージ $x_i \rightarrow y_i$ となる確率．

ここで，$p(x_i, y_j) = p(x_i)p(x_j)$ と考える人がいるかもしれない．それは間違いである．次のような二元対称通信路 (図 4) を考えれば，明らかである．

1. 誤りが全く生じない場合を考える．送信者が'0' を送れば，受信者は必ず'0' を受け取る．反対に'1' を送れば，'1' を受け取る．

2. 送信者が'0' および'1' を送る確率は，それぞれ 0.5 である．すなわち，$p(x = 0) = p(x = 1) = 0.5$ となる．

3. 途中で誤りが全く生じないので，受信者が'0' および'1' を受け取る確率はそれぞれ 0.5 である．すなわち，$p(y = 0) = p(y = 1) = 0.5$ となる．

4. 従って，$p(x = 0, y = 0) = p(x = 1, y = 1) = 0.5$，$p(x = 0, y = 1) = p(x = 1, y = 0) = 0$ となる．

この場合，明らかに $p(x_i, y_j) = p(x_i)p(x_j)$ は成立しない．この関係式が成り立つのは，入出力がお互いに全く無関係の場合のみである．このような通信路は全く役に立たない!! 2 元対称通信路の場合，誤りの確率が $q = 0.5$ のときである．

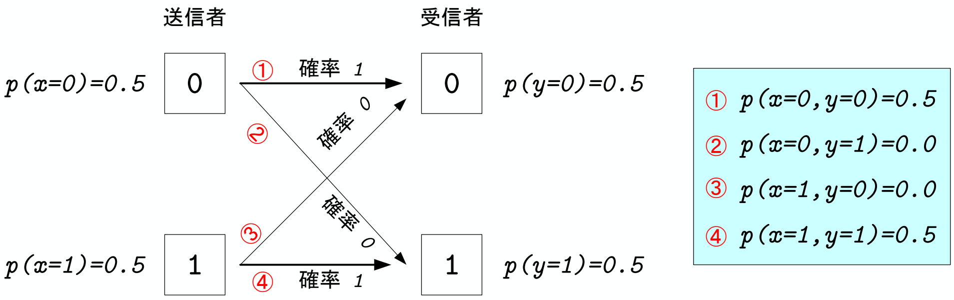

図 4: 通信路の途中で誤りが発生しない 2 元対称通信路の確率．この場合，明らかに $p(x_i, y_j) \neq p(x_i)p(x_j)$ である．

元に戻って，今後のために一般的な 2 次元対通信路の確立の関係を示しておく．全て，直感的に理解できる式である．

$$p(x=0)+p(x=1)=1 \qquad p(y=0)+p(y=1)=1 \tag{4}$$

$$p(x=0,y=0)+p(x=0,y=1)+p(x=1,y=0)+p(x=1,y=1)=1 \tag{5}$$

$$p(x=0)=p(x=0,y=0)+p(x=0,y=1) \qquad p(x=1)=p(x=1,y=0)+p(x=1,y=1) \tag{6}$$

$$p(y=0)=p(x=0,y=0)+p(x=1,y=0) \qquad p(y=1)=p(x=0,y=1)+p(x=1,y=1) \tag{7}$$

これらの式は，式 (2) や式 (3) に対応したものである．

2.2.2　システムのエントロピー

神様になった気持ちでシステム全体を見渡して，エントロピー (平均情報量) を考える．このシステムを観測することにより得られる情報は，送信者の文字 x_i と受信者の文字 y_j である．送受信が (x_i, y_j) となる確率は $p(x_i, y_j)$ なので，エントロピー $H(X,Y)$ は

$$H(X,Y) = -\sum_i \sum_j p(x_i, y_j) \log_2 p(x_i, y_j) \tag{8}$$

である．これを，結合エントロピーと呼ぶ．また，システムエントロピーと呼ばれることもある．

これは，ある一つの通信を行ったとき，送信/受信の両方の文字を知ることにより，得ることができる平均情報量である．あるいは，このシステムの通信前の情報の平均的な情報の曖昧 (不確かさ) の度合いを表していると言っても良い．

1 回通信を行い，送信/受信の文字を知ることにより，情報が確定する; 曖昧さがなくなる．確率の低い情報ほど大きな情報を得ることができる．確率と情報量の関係は，

$$\text{情報量} = -\log_2 p \tag{9}$$

であった．ここで，p はその情報が生じる確率である．底の 2 は大きな意味は無く，他でも良いがバイナリーデータを扱うことが多いので，その方が都合が良い．これから，結合エントロピーは，得られる情報量の期待値であることが分かる．これは，以前学習したとおり．

2.3 条件付きエントロピー

次に受信者側を考える．この場合でも，受け取る文字に関するエントロピー (平均情報量) は，

$$H(Y) = -\sum_j p(y_j)\log_2 p(y_j) \tag{10}$$

となる．受信者が文字 y_j を受け取ると，平均的に $H(Y)$ の情報を得る．

つぎに，文字 y_j を受け取ったときに，送信者が送った文字を推定することを考える．文字 y_j を受け取ったときに，送信文字 x_i に関してどれだけの情報を得ることができるか?—と言う問題である．

いま，受信者が文字'0' を受け取ったとする．このとき，

$$\begin{aligned}
\text{送信文字が'0' である確率} &= \frac{p(x=0,y=0)}{p(x=1,y=0)+p(x=0,y=0)} \\
&\quad \text{式 (7) を使うと} \\
&= \frac{p(x=0,y=0)}{p(y=0)}
\end{aligned} \tag{11}$$

となる．これを

$$p(x=0|y=0) = \frac{p(x=0,y=0)}{p(y=0)} \tag{12}$$

と表し，条件付き確率と呼ぶ．文字'0' を受け取ったとき，送信文字が'0' である確率である．同様に，文字'0' を受け取ったときに，送信文字が'1' である確率は，$p(x=1|y=0) = p(x=1,y=0)/p(y=0)$ となる．したがって，文字'0' を受け取った場合，入力文字のあいまいさ (エントロピーあるいは平均情報量) は

$$H(X|y=0) = -\sum_i p(x_i|y=0)\log_2 p(x_i|y=0) \tag{13}$$

となる．$p(x_i|y=0)$ は受け取った文字が'0' のとき，送信した文字が x_i である確率である．

以上のことより，受け取った文字が y_j の場合の，送信文字に関するエントロピー $H(X|Y)$ は

$$H(X|Y) = -\sum_j p(y_j)H(X|y=y_j) \tag{14}$$

となる．$p(y_j)$ は受け取る文字が y_j となる確率である．$H(X|y=y_j)$ はそのときのエントロピーである．したがって，$H(X|Y)$ は文字 y_j を受け取った後の送信文字に関する平均的なあいまいさ (条件付きエントロピー) となっている．

この式をもう少し変形すると，次のようになる．

$$
\begin{aligned}
H(X|Y) &= -\sum_j p(y_j) H(X|y=y_j) \\
&= -\sum_j p(y_j) \sum_i p(x_i|y=y_j) \log_2 p(x_i|y=y_j) \\
&= -\sum_i \sum_j p(y_j) \frac{p(x=x_i, y=y_j)}{p(y=y_j)} \log_2 \frac{p(x=x_i, y=y_j)}{p(y=y_j)} \\
&\qquad p(x=x_i, y=y_j) = p(x_i, y_j), \quad p(y=y_j) = p(y_j) \text{ となるので} \\
&= -\sum_i \sum_j p(x_i, y_j) \log_2 \frac{p(x_i, y_j)}{p(y_j)} \\
&= -\sum_i \sum_j p(x_i, y_j) \log_2 p(x_i, y_j) + \sum_i \sum_j p(x_i, y_j) \log_2 p(y_j) \\
&\qquad \text{右辺第一項は，式 (8) より} \\
&= H(X, Y) + \sum_j \left[\sum_i p(x_i, y_j) \right] \log_2 p(y_j) \\
&\qquad \text{右辺第二項は，式 (3) より} \\
&= H(X, Y) + \sum_j p(y_j) \log_2 p(y_j) \\
&\qquad \text{右辺第二項は，式 (10) より} \\
&= H(X, Y) - H(Y)
\end{aligned} \tag{15}
$$

これは，出力文字を知った後での，入力文字のあいまいさを表している．

これと全く同じ議論が，逆の場合でも成り立つ[3]．文字 x_i を送ったとき，受信者が受け取る文字 y_j に関するあいまいさでる．式で表すと

$$
H(Y|X) = H(X, Y) - H(X) \tag{16}
$$

となる．ここで，$H(Y|X)$ は文字を送信した後に受信者が受け取る文字のあいまいさ (条件付きエントロピー) である．$H(X)$ は元々の送信する側の文字のあいまいさである．$H(X, Y)$ の X と Y が入れ替わらない理由? これはシステムのエントロピーであり，送信者側や受信者側とは関係ない，全体に関することである．神様が見ているようなもので，逆の議論でも関係ない．

2.4 相互情報量

定義　文字 y_j 受け取ったとき，送信文字 x_i について分かったことを考える．文字を受け取る前の受信者が送信文字 x_i に関して持っているあいまいさ (エントロピー) は，$H(X)$ であった．ここで y_j 受け取る (観

[3]送信/受信の状況をビデオ撮影して，逆回転させてみよ．受信者が y_j と言う文字を送り，送信者が x_i を受け取るように見える．正転と逆転の映像の区別は付かないだろう．

測) とあいまいさは，$H(X|Y)$ となった．文字 Y_j を受け取ることにより，あいまいさが $H(X) \to H(X|Y)$ と変化した．この変化した量が受け取った情報量 $I(X;Y)$ である．したがって，

$$I(X;Y) = H(X) - H(X|Y) \tag{17}$$

である．この $I(X;Y)$ を**相互情報量**と呼ぶ．式 (16) と式 (15) から，

$$I(X;Y) = H(X) - H(X|Y) = H(Y) - H(Y|X) = I(Y;X) \tag{18}$$

と書き直すことができる．この式の言っていることは少し不思議な感じがする．つぎの 2 つが全く同一となる．

- $I(Y;X)$ は，文字 x_i を送った後，受信者が受け取るであろう文字 y_j のあいまいさの減少を表す．
- $I(X;Y)$ は，文字 y_j を受け取ると，送信者が送ったであろう文字 x_i のあいまいさの減少を表す．

グラフ化　それでは，具体的に $H(X|Y)$ と $I(X;Y)$ をプロットしてみよう．もっとも単純な 2 次元対称線路を考える．図 5 に示すように，送信者が'0' を送る確率 $p(x_i = 0) = \alpha$ とし，誤りの確率を q とする．

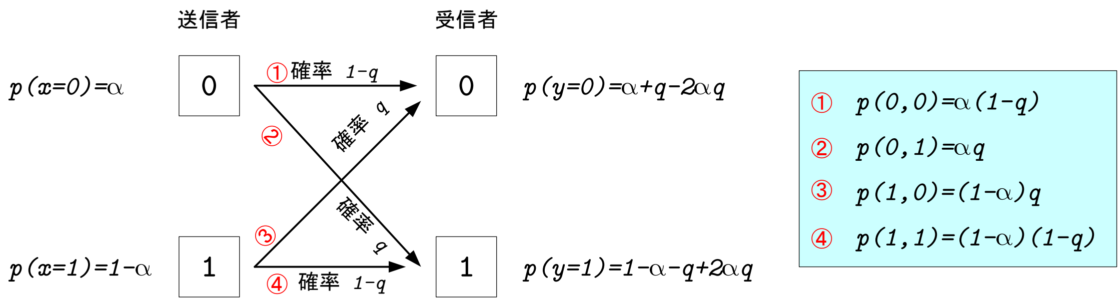

図 5: 一般的な 2 元対称通信路．送信者が'0' を送る確率を α，通信路の途中で誤りが生じる確率を q としている．

式 (15) を使うと，

$$\begin{aligned} H(X|Y) &= -p(0,0)\log_2\frac{p(0,0)}{p(y=0)} - p(0,1)\log_2\frac{p(0,1)}{p(y=1)} - p(1,0)\log_2\frac{p(1,0)}{p(y=0)} - p(1,1)\log_2\frac{p(1,1)}{p(y=1)} \\ &= -\alpha(1-q)\log_2\frac{\alpha(1-q)}{\alpha+q-2\alpha q} - \alpha q\log_2\frac{\alpha q}{1-\alpha-q+2\alpha q} \\ &\qquad -(1-\alpha)q\log\frac{(1-\alpha q)}{\alpha+q-2\alpha q} - (1-\alpha)(1-q)\log_2\frac{(1-\alpha)(1-q)}{1-\alpha-q+2\alpha q} \end{aligned} \tag{19}$$

が得られる．さらに，式 (10) を使うと

$$\begin{aligned} H(X) &= -p(x=0)\log_2 p(x=0) - p(x=1)\log_2 p(x=1) \\ &\quad -\alpha\log_1\alpha - (1-\alpha)\log_2(1-\alpha) \end{aligned} \tag{20}$$

が得られる．以上より，$H(X|Y)$ と $H(X)$ の計算でき，図 6 のグラフ (教科書 [1] の p.68 の図 3.23 と同じ) が得られる．

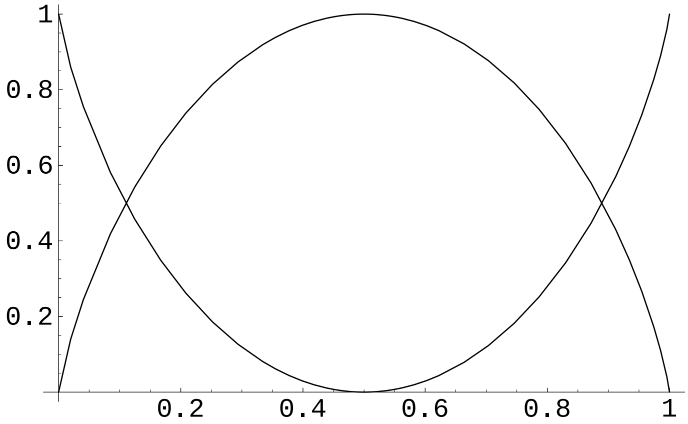

図 6: 2 元対称線路の $H(X|Y)$ と $I(X;Y)$ のグラフ．上向きの凸が $H(X|Y)$ で，凹が $I(X;Y)$ である．横軸は通信路のエラーの確率 q である．$p(x=0)=0.5$ の場合である．

図 6 のグラフから，次のことが分かる．

- 条件付きエントロピー $H(X|Y)$ は，文字 y_j を受け取ったときに送信文字 x_i のあいまいさを表している．受信者が送信文字を何かの方法で知ったとき; 送信文字が確定したときに得られる平均情報量 (エントロピー) と言ってもよい．グラフより以下のことが分かる．
  - 通信路でのエラーの確率 q が 0 または 1 の時は，条件付きエントロピーはゼロである．この場合，即座に送信文字は確定できる．したがって，あいまいさは全くない．
  - 通信路でのエラーの確率 q が 0.5 のとき，条件付きエントロピーは最大になる．この場合，文字を受け取っても，全く送信文字の推定ができないからである．
- 相互情報量 $I(X;Y)$ は，文字 y_j を受け取ったときに，送信文字 x_i に関して得る情報量である．グラフより以下のことが分かる．
  - 通信路でのエラーの確率 q が 0 または 1 の時は，相互情報量エントは最大になる．この場合，即座に送信文字は確定できる．したがって，送信文字に関して多くの情報を得たことになる．
  - 通信路でのエラーの確率 q が 0.5 のとき，相互情報量はゼロになる．この場合，文字を受け取っても，全く送信文字に関する情報を得ることができないからである．

付録 A　$x \log_2 x$ の関数の形

エントロピーの計算では，

$$f(x) = -x \log_2 x \tag{21}$$

の関数がしばしば表れる．情報科学の分野では，変数 x の部分には確率になることが多い．そのため，$[0, 1]$ の間で，この関数の形を覚えておくと良いだろう．この関数をプロットすると，図 7 のようになうる．

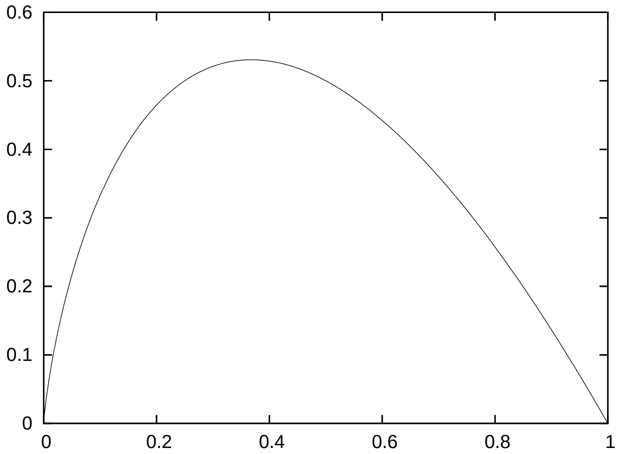

図 7: $x \log_2 x$ の関数の形.

ここで，$f(0) = 0$ なのか?—という疑問が湧く．$x \to 0$ の極限では，x はゼロに近づくが，$\log_2 x$ は $-\infty$ となる．$x = 0$ の部分をちゃんと調べなくてはならない．そこで，$x = 0$ の極限の値を調べてみることにする．

$$f(0) = -\lim_{x \to 0} x \log_2 x \tag{22}$$

この手に極限の問題は，ロピタルの定理を使う．

$$\begin{aligned} f(0) &= -\lim_{x \to +0} x \log_2 x \\ &= -\frac{1}{\log_e 2} \lim_{x \to +0} \frac{\log_e x}{1/x} \\ &= -\frac{1}{\log_e 2} \lim_{x \to +0} \frac{1/x}{-1/x^2} \\ &= \frac{1}{\log_e 2} \lim_{x \to +0} x \\ &= 0 \end{aligned} \tag{23}$$

これで，めでたしめでたしである．この関数では，$f(0)$ は不定になるであろう．$x = 0$ のところではかなり特異な性質がある．ただし，$f(+0) = 0$ となる．

それでは，質問．$f(x) = x^x$ とする．$f(0)$ はどうなるか?

参考文献

[1] 河合慧（編）．情報. 東京大学出版会, 2006.