情報の伝達と通信 (その 1)

山本昌志*

2007 年 10 月 24 日

概 要

情報量について説明する．さらに，平均情報量とデータの圧縮限界について，述べる．

1　本日の学習内容

情報量について学ぶ．学習のゴールは以下の通りである．

- 情報の大きさが分かり，情報量の計算ができる．
- 平均情報量とデータ圧縮との関係が分かる．

教科書 [1] の pp.37–45 が本日の範囲である．

2　情報の伝達と情報量

2.1　情報の伝達

諸君は，物やエネルギーの伝達はよく知っている．トラックや鉄道を使った工業製品の輸送 (伝達) は説明するまでもなく，よく見る光景である．電線を使った電力の輸送 (伝達) も理解しているだろう．前者はある定まった物が送られ，後者はエネルギーが送られる．

それに対して，携帯電話のメールで送られるものは何だろうか? 言うまでもなく，送られる物は情報である．これもある程度理解できるであろう．

問題は，送られる情報を定量的に取り扱う方法を見いだすことである．物の伝達であれば，質量や体積で取り扱うことができる．エネルギーであれば，ジュール (Jule) という単位で取り扱うことができる．情報はどのような単位で取り扱えばよりのだろうか?

情報の取り扱いが学問になるためには，それを定量的に表現しなくてはならない．次節から，情報を定量的に取り扱う方法を述べる．

*独立行政法人　秋田工業高等専門学校　電気情報工学科

2.2　情報の大きさ—情報量

2.2.1　情報の大きさに関して

全く同じ**メッセージ**を受け取っても，受け手によって情報量は異なる．例えば，次のような状況を考えよう．

- まじめな，A 君と C 君は講義をよく聞いている．先生の発言「今回の試験では，教科書の 203 ページから 223 ページから出題する」を聞き逃さなかった．
- 一方，B 君は講義中，ほとんど寝ている．
- 学年末試験までに講義で取り扱った内容は，教科書の p.153–253 であった．B 君は 100 ページも勉強しなくてはならないと思い憂鬱な気分であった．
- 昼休憩に A 君と B 君，C 君が話をしていた．図 1 のような場面があった．

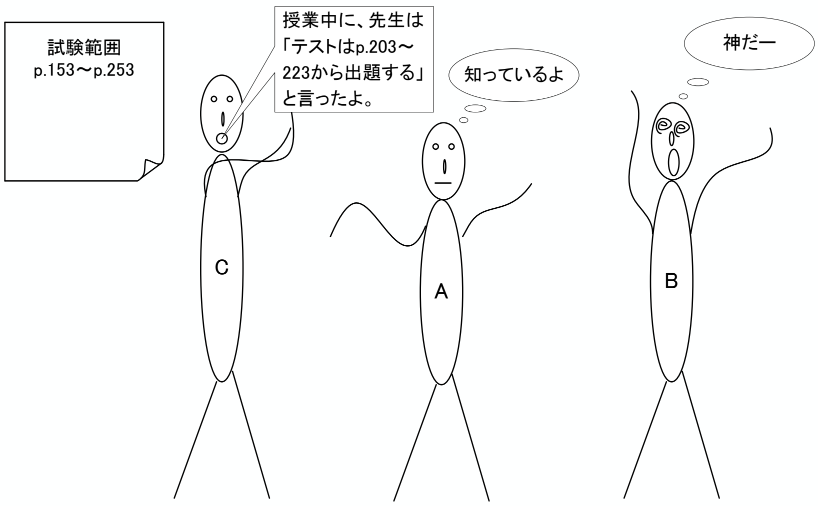

図 1: 同じ情報でも，人により得る情報量が異なる．あらかじめそれを知っていた A 君が得る情報はゼロであるが，知らなかった B 君は多くの情報を得る．

A 君と B 君は，全く同じメッセージを C 君から受け取った．しかし，A 君はすでに知っている内容で，A 君の情報量は全く増加しない．それに対して，このメッセージで B 君はとてつもなく大きな情報を得たことになる．

それでは，情報の大きさの大小関係を考えてみよう．先ほどの例から，次のようなことが分かるだろう．

> メッセージを受け取ることにより，何かを知ることになる.「驚き」の多いメッセージほど，大きな情報量がある．

今まで無かった情報を得たときに，人間は驚く．まったく，知識が無いことを得ると驚きは大きくなる．今のところ，情報量を「驚き」というような，およそ学問とは関係なさそうな人間の感情で表している．あとで，分かるがこれを手がかりに情報量を定量的に評価してする．

大小関係が分かったので，次の節から情報量を定量的に定義する．

2.2.2 場合数の変化

このあたりは教科書 [1] の pp.38–41 に沿って説明する．教科書をにらみながら理解せよ．プリントのみでは理解できない．

教科書の p.39 の内容をまとめると，次のようになる．

- 歴史のテストがある．歴史は，日本史と東洋史，西洋史，アメリカ史がある．それぞれの出題確率は，1/4 である．
- 世界史とは，日本史以外，すなわち東洋史と西洋史，アメリカ史のことである．

このような状況の中で，メッセージにより受け取る情報量を考える．

先ほどの例 (A 君と B 君，C 君) から，受け取る情報量は

$$\textbf{受け取る情報量} = \frac{\text{メッセージを受け取る前の勉強する歴史の数}}{\text{メッセージを受け取った後に勉強する歴史の数}} \tag{1}$$

とすることができるだろう．これは，メッセージを受け取ることにより驚きを表していることになる．例えば，次のような例である．明日，歴史のテストがある状況で，教師が次のようなメッセージを発信した．

(A) 明日は歴史のテストがあるので勉強してくるように．

(B) 明日の歴史のテストは，アメリカ史に限る．

式 (1) の定義によると，(A) の場合の受け取る情報量は 1 で，(B) の場合は 4 である．明らかに (B) のメッセージの方が情報が多いので，式 (1) で情報量を定義しても良さそうである．

実は，式 (1) が不便な場合がある．次のように 2 段階にメッセージが発信された場合である．

(i) 明日の歴史のテストは，世界史 (東洋史，西洋史，アメリカ史) を出題する．

(ii) 東洋史と西洋史は出さない．

これで受け取る情報量を計算してみよう．(i) のメッセージでは，4/3 の情報を受け取る．(ii) のメッセージでは，3/1 のの情報を受け取る．

それでは，(i) と (ii) の二つのメッセージで合計，どれだけの情報を受け取ったのであろうか? $4/3+3/1=13/3$ と考えて良いだろうか? これはまずい．なぜならば，(i) と (ii) をあわせたメッセージは，先の (B) のメッセージと同じ内容である[1]．にもかかわらず，情報量が異なる．

解決の方法は，

$$\textbf{受け取る情報量} = \log_2 \frac{\text{メッセージを受け取る前の勉強する歴史の数}}{\text{メッセージを受け取った後に勉強する歴史の数}} \tag{2}$$

[1]明日の歴史のテストはアメリカ史から出題される

と定義すればよい．対数の底は 2 である必要は無い．しかし，コンピューター科学では 0 と 1 の 2 進数を使うので，2 とすると後々便利なことが多い．こうすることにより，

$$\text{(B) のメッセージの情報量} = \log_2\left(\frac{4}{1}\right) = 2 \tag{3}$$

$$\text{(i) と (ii) のメッセージの情報量} = \log_2\left(\frac{4}{3}\right) + \log_2\left(\frac{3}{1}\right) = 2 \tag{4}$$

となり，まことにハッピーな結果が得られる．

ものいい

ちょっと待った! (i) と (ii) の発言により得られる情報量を $4/3 \times 3/1$ とすれば，式 (1) の定義でも良いのではないか? 確かに，2 つの情報を得たときの演算を乗算で定義すれば，情報量を式 (1) のように定義しても矛盾はない．しかし，情報は加えることができる—と言うことを考えると，二つの情報は加算を行う方がしっくりくる．一般に量は，加算ができる方がいろいろと便利である．どのような量も二つの物を合計するときは加算の演算を行う．従って，加法性を受け入れるならば，情報量は式 (2) で定義すべきである．

2.2.3 確率による定義

次に述べる確率による定義がもっとも一般的に情報量のとして使われる．確率 p で起きる事実を知ったときに得られる情報量は，

$$\text{情報量} = -\log_2 p \quad \text{[bit]} \tag{5}$$

である．

それでは，先の (B) のメッセージと (i) と (ii) を加えたメッセージで，この確率による情報量を計算してみよう．

- (B) の場合．アメリカ史が出題される確率は 1/4 であった．(B) のメッセージにより，出題内容はアメリカ史に限定された．

$$\text{(B) のメッセージの情報量} = -\log_2\left(\frac{1}{4}\right) = 2 \quad \text{[bit]} \tag{6}$$

- (i) と (ii) の場合．(i) の「世界史を出題する」は 3/4 の確率の事実である．(ii) の「アメリカ史に限る」はそのときは 1/3 の確率である．

$$\text{(i) と (ii) のメッセージの情報量} = -\log_2\left(\frac{3}{4}\right) - \log_2\left(\frac{1}{3}\right) = 2 \quad \text{[bit]} \tag{7}$$

2.3 平均情報量

次にメッセージを受け取ったときに得る情報量を計算しよう．これは，後で述べる情報の圧縮と密接に関わっている．

一つのメッセージを受け取ったときにいつも情報量が同じとは限らない．二つの場合について，メッセージに含まれる平均情報量[2]，すなわちメッセージを受け取ることにより得られる情報量の期待値を

$$\textbf{平均情報量} = -\sum_i p_i \log_2 p_i \tag{8}$$

計算してみよう．ここで，p_i はメッセージ i が生じる確率である．当然，

$$\sum_i p_i = 1 \tag{9}$$

となる．

等確率の場合　受け取る情報が,「日本史」か「東洋史」,「西洋史」,「世界史」の 4 通りでそれぞれ等確率の場合を考える．このとき，受け取るメッセージを次のように符号化する．

$$\text{日本史} \Rightarrow 00 \quad \text{東洋史} \Rightarrow 01 \quad \text{西洋史} \Rightarrow 10 \quad \text{アメリカ史} \Rightarrow 11$$

この場合，メッセージに含まれる平均情報量は

$$\begin{aligned}\text{平均情報量} &= -\frac{1}{4}\log_2\frac{1}{4} - \frac{1}{4}\log_2\frac{1}{4} - \frac{1}{4}\log_2\frac{1}{4} - \frac{1}{4}\log_2\frac{1}{4} \\ &= 2\,[\text{bit}]\end{aligned} \tag{10}$$

となる．符号化の 2 進数の桁数と同じ．

等確率でない場合　受け取る情報が,「日本史」か「世界史 (東洋史・西洋史・アメリカ史)」の二通りで，確率は 1/4 と 3/4 である．受け取るメッセージは，次のように符号化する．

$$\text{日本史} \Rightarrow 0 \quad \text{世界史} \Rightarrow 1$$

この場合，メッセージに含まれる平均情報量は

$$\begin{aligned}\text{平均情報量} &= -\frac{1}{4}\log_2\frac{1}{4} - \frac{3}{4}\log_2\frac{3}{4} \\ &= 0.5 + 0.311 \\ &= 0.811\,[\text{bit}]\end{aligned} \tag{11}$$

となる．平均的に受け取るメッセージは，符号化の 2 進数の桁数よりも小さくなる.「日本史」は 2 [bit] の情報があるが,「世界史」の情報は 0.415 [bit] である．

このように，受け取るメッセージの確率が異なる場合，受け取る平均情報量が 2 進数で符号化した場合の桁数よりも小さくなることがある．一方，確率が等しい場合は，平均情報量と 2 進数で符号化した場合の桁数は同一となる．このあたりは，教科書 [1]pp.43 の図 3.3 を見よ．

[2] シャノンは，平均情報量を「情報のエントロピー」と呼んだ．

コーヒーブレイク

平均情報量を表す式 (8) に関して，ファインマンの教科書 [2] にはおもしろいことが書かれている．この式は，統計力学 (熱力学) のエントロピーの式とほとんど同じである—と言う知識を前提に，以下引用する．

> シャノンはこの平均情報量を「情報のエントロピー」と呼んだが，これを大きな間違いだったと思う人がいる．というのは，これによって多くの人が情報理論と熱力学との関連を強調しすぎることになったからである．

また，次のようなことも書いてある．

> 言い伝えによれば，シャノンはこの用語を，数学者フォン・ノイマンの助言によって採用した．フォン・ノイマンは「ともかく誰も実際にエントロピーが何なのかがわからないので $\cdots$ この用語によってシャノンは論争に優位に立つだろう」と公言したということである．

2.4　符号化と情報量

2.4.1　符号化

先に述べたように，メッセージは 2 進数で符号化できる．現代のデジタル化した情報社会では，この符号化したデータを送受信することにより，情報のやりとりを行っている．

送受信のコストを下げるためには，できるだけ少ないデータ;2 進数の桁数が小さいで，多くの情報を送らなくては成らない．ここでの問題は，平均情報量が分かった場合，必要な 2 進数符号の桁数は? どのようにすれば 2 進数符号の桁数を最小にできるか?である．

2.4.2　データの圧縮

先ほどの二つの問いは，データ圧縮の問題そのものである．データ圧縮は，最小のデータで最大の情報を表す方法である．

メッセージを 2 進数で符号化する．i 種類目のメッセージ m_i の符号の桁数を ℓ_i とする．この場合，平均符号長は，

$$\text{平均符号長} = \sum_i p_i \ell_i \tag{12}$$

となる．この平均符号長さは，そのメッセージの平均情報量は N ビットを超えることができない．なぜならば，N 桁の 2 進数; 符号長で表すことのできる最大情報量は N ビットであるからである．

これがデータ圧縮の限界を表す．教科書の例はハフマン符号化を行い，平均符号長を平均情報量に近づけている．

ハフマン符号化については，省略する．ただし，課題にそれがあるので調べよ．

2.5　2 進数の桁数と情報量の関係

0 と 1 が同じ確率で発生するメッセージの場合，2 進数の桁数と情報量が同じであることを示そう．0 と 1 の発生確率が同じ N 桁の 2 進数の場合，ある特定のパターンの 2 進数が発生する確率 p は，

$$p = \frac{1}{2^N} \tag{13}$$

である．従って，どのようなパターンのメッセージを受け取っても，得られる情報は

$$\begin{aligned} \textbf{情報量} &= -\sum_{i=1}^{2^N} p \log_2 p \\ &= -\sum_{i=1}^{2^N} \frac{1}{2^N} \log_2 \frac{1}{2^N} \\ &= \sum_{i=1}^{2^N} \frac{N}{2^N} \\ &= 2^N \frac{N}{2^N} \\ &= N\,[\mathrm{bit}] \end{aligned} \tag{14}$$

となる．この場合，2 進数の桁数と情報量が同一となる．

3　課題

3.1　課題内容

以下の課題を実施し，レポートとして提出すること．

[問 1] (復予) 教科書 [1]pp.37–68 を 2 回読み，重要な部分には赤線でアンダーラインを入れよ．レポートには「2 回読んだ」と書け．

[問 2] (復)4 つの整数 (0,1,2,3) の 6 桁で符号化してメッセージを作成する．例えば，041023 や 112403 等である．この符号一つで，送ることのできる最大情報量は，いくらか? また 12 桁のひとつの符号と，6 桁二つの符号では情報量は異なるか?

[問 3] (復) 教科書 [1] 章末問題 (p.69) の [3.1]

[問 4] (復) 教科書 [1] 章末問題 (p.69) の [3.2]

[問 5] (復) 教科書 [1] 章末問題 (p.69) の [3.3]

3.2 レポート 提出要領

期限	11 月 17 日 (金) AM 8:45
用紙	A4 のレポート用紙．左上をホッチキスで綴じて，提出のこと．
提出場所	山本研究室の入口のポスト
表紙	表紙を 1 枚つけて，以下の項目を分かりやすく記述すること． 授業科目名「情報理論」 課題名「課題　情報の伝達と通信 (その 1)」 提出日 5E　学籍番号　氏名
内容	2 ページ以降に問いに対する答えを分かりやすく記述すること．

参考文献

[1] 河合慧（編）．情報．東京大学出版会, 2006.

[2] A. ヘイ, R. アレン（編）．ファインマン 計算機科学．岩波書店, 2002．物理学者が情報科学に関して，講義を行った時の講義ノート．非常にユニークな視点で書かれており，一読を勧める．