北星短大紀　*Journal of Hokusei Jr. Col. Vol.30 : 87−94, 1994*

知能作用過程のシュミレーション

その５　文字情報の「あいまいさ」に関する理論的考察

Simulation of the Operating-processes of the Human Intellect

Part 5 . Theoritical Study on Fuzzy Entropy in Literal Inforomation

Teraoka Hiroshi and Yabuki Tetuo

寺　岡　　宏　　*矢　吹　哲　夫

In our previous paper, we calculated the amount of information in our literal system, by means of the Shannon entropy. In the calculation, it was assumed tacitly that every word has no obscurity. However there is no single word free from any obscurity. When we practically read some sentence, we are inevitably confronted with many kinds of obscurity in the word which compose the sentence.

In this paper, we assumed that obscurity in the word has a suppressive effect in the reading sentence, and we estimated the obscurity resulted from the multilpe meaning in a word. For this calculation, we used 「fuzzy entropy」which had been introduced by DeLuca and Termini.

The first part of this paper is devoted to the general explanation about the theory of 「fuzzy entropy」, and the second part is devoted to the application of it to our literal system. Finnally we calculate the total amount of fuzzy entropy in our system, and we estimate the total quantity of the infomation whose value should be suppresed by the fuzzy entropy.

本研究シリーズの前報[(1)]において、言語系の持つ情報量についての理論的考察がなされた。さらに文字集団が Zipf の法則[(2)]に基く出現確率で与えられる人工言語系を設定し、その系の情報量を、シャノンの「情報エントロピー」[(3)]を基礎に計算した。前報(１)において、詳述されたように、「情報エントロピー」という概念は、今日情報理論の分野において、不可欠な概念として確立されている。この概念の基礎には、言語の対象に対する二者択一的な指示機能の前提がある。しかし、日常の言語の使用に際しては、この概念だけでは把握し得ない「あいまいさ」に直面する。この「あいまいさ」には、例えば、文

＊北海道大学・理学部・物理学科・大学院博士課程

字の書き損ねという単純なミスによる「あいまいさ」から、一つの文字が本来備える多義性に由来する「あいまいさ」まで、種々のレベルの内容が含まれる。

本論文においては、「あいまいさ」を、文字読取りと認知のモデルの中に導入することを試み、「あいまいさ」を測る尺度として、新たに「ファジーエントロピー」[4]を用いた。

以下において、「ファジーエントロピー」の基礎概念を概説し、これを言語系の計算にいかに応用するかを述べる。最後に人工言語系におけるファジーエントロピーの計算と、それに伴う情報エントロピーの変化について記述する。

A：「あいまいさ」の概念

今日、“不確かさ”という概念は、ふたつのカテゴリーに大別されている。その各々に対して、その概念内容にふさわしい数学的処理の方法が考察され適用されている。

その一つめは、事象対象の生起に関する“不確かさ”で、例えば「明日の天気はどうなるか？」等がその例である。この“不確かさ”は、本質的には、人間の予備知識の欠如に起因するものであり、これに対しては、「確率」という数学的手法が確立し、既に多くの成果をみている。

二つめは、事象対象自体が、備える“不確かさ”であり、事象の生起が完全に把握されても尚残るものであり、一つめの“不確かさ”と峻別する意味で「あいまいさ」と一般に呼ばれている。この「あいまいさ」を解析する新たな数学的道具として考案されたものが、ファジー集合に立脚するファジー概念である。

更に最近になって、「あいまいさ」を数量的に測る新しい尺度が幾つか「ファジー学」において提唱されている。「ファジー学」自体が発展途上にある新しい分野であり、「ファジー」を測る尺度にもいくつかの異なる方法が見られ、必ずしも一意的ではない。本論文ではそのうち最も一般的なものと見られている、「ファジーエントロピー」に立脚し、その定式化と人工言語系への導入を目的とする。

B：ファジーエントロピーの定義とその意味

ファジーエントロピーは、ファジー集合の概念[5]を土台としてその上に定義されている。そこで、まずファジー集合を以下のように定義する。ある全体集合をXとする。このときX上で定義されるファジー集合Aは、Xの各要素xがAに属する度合いを示す所属度関数$\mu_A(x)$で規定される集合である。

もし、この所属度関数$\mu_A(x)$が0か1のどちらかしかとらないときは、Aは通常の集合となり、$\mu_A(x)=1$を満たす要素xの集まりが集合Aであり、$\mu_A(x)=0$をみたす要素xの集まりは、Aの補集合となる。

しかし、一般のファジー集合においては、$\mu_A(x)$は0から1までの連続的な値をとり、集合Aに属するか属さないかの二者択一は成り立たない。この意味で、集合Aはファジー（Aに属するか属さないかがあいまい）となる。

一例として「Man」という全体集合Xを設定するとき、これに属する部分集合として「Infant」「Young age」「Adult」「Old age」などが考えられる。

これらの部分集合にたいし、年齢(x)を変数とした所属度関数$\mu(x)$を設定するとき、各$\mu(x)$はそれぞれ異なるxの範囲において$\mu(x)=0$から$\mu(x)=1$までの連続値を示すことになる。すなわち、上記4つの部分集合はそれぞれに異なる「あいまいさ」を有する「ファジー集合」として定義される。

ファジー集合「Infant」は、他の、３つの集合に比較して、$\mu(x)=0$ から連続的に $\mu(x)=1$ に変化する年齢巾は狭く、それだけ「あいまいさ」は少ないとみなされる。このような「あいまいさ」を数量的に測る尺度とその基本条件について以下考察をする。

今、あるファジー集合Aの「あいまいさ」を測る尺度を d(A) として、d の値域を $d>0$ とする。このとき、d に対して以下の基本条件を課すことは、「あいまいさ」という語のもつ概念に照らしてみて妥当であることが De Luca *et* によって記されている[4]。

(1) Aが通常の集合に一致する場合
　$d(\mathrm{A})=0$ となる。

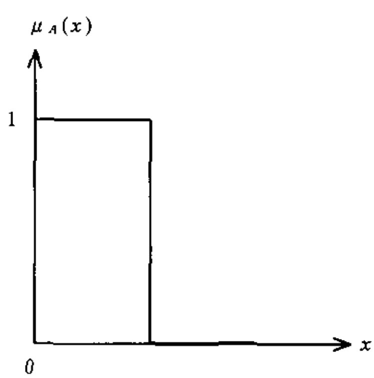

図１

これは図１のように $\mu_A(x)$ の値が０か１しかないときであり、「あいまいさ」は０である。

(2) 任意の要素 x に対して、$\mu_A(x)=\frac{1}{2}$ のとき d(A) は最大となる。

これは、$\mu_A(x)$ のグラフが図２のようになる場合で、すべての x に対してAに所属するかしないかが半々で、まったく判然としない場合である。

このときは、「あいまいさ」d(A) は最大となる。

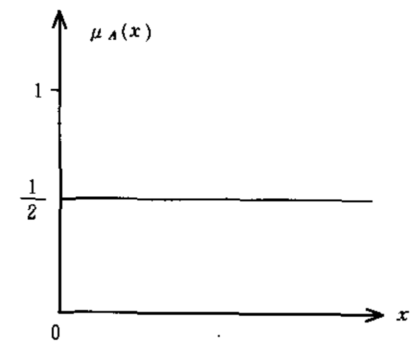

図２

(3) AとA′の二つの集合があり、A′がAより“尖鋭”であるならば
$d(A')\leqq d(\mathrm{A})$

ここでA′がAより“尖鋭”とは $\mu_A(x)\leqq 0.5$ の x に対しては $\mu_{A'}(x)\leqq\mu_A(x)$、$\mu_A(x)>0.5$ の x に対しては $\mu_{A'}(x)\geqq\mu_A(x)$ が成り立つことをいう。（図３参照）

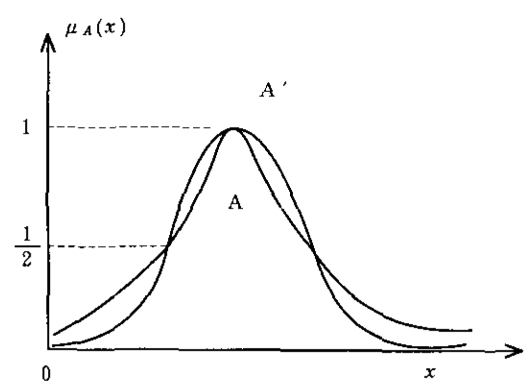

図３

(4) Aの補集合を $\bar{\mathrm{A}}$ とするとき（$\mu_{\bar{A}}(x)=1-\mu_A(x)$ で定義される。）
$d(\bar{\mathrm{A}})=d(\mathrm{A})$

これは、「Aに属するか属さないか」と「Aに属するか属さないか」は互いに表裏の関係であり、両者のあいまいさ d は等しいからである。

上の(１)〜(４)の基本条件を満たす「あ

いまいさ」の尺度 d （A）として Deluca & Termini [4] は $\mu_A(x)$ による次の関数形を与えた。

$$d(A) = -k\int_a^b dx\{\mu_A(x)\log\mu_A(x) - (1-\mu_A(x))\log(1-\mu_A(x))\} \quad \cdots (1)$$

ただし、要素 x が実数空間 $[a,\ b]$ に連続的に分布している場合を考え、

$$k = \frac{1}{\int_a^b dx} = \frac{1}{b-a}$$ で規格化されるものとする。

上式（1）が今日ファジーエントロピーと呼ばれているものである。

その関数形がエントロピーの形と類似していることからこの名が付された。

C：情報量とファジーエントロピー

既述の論文で、我々の設定した人工言語系の有する一語当たりの情報量を情報エントロピーとして計算した。その際、情報素材としての文字又は文字集団一つ一つは正確にアウトプットされ、使用される文脈の中で、一義的な意味を常に付与されていることを前提とした。

しかし、実際には、一つ一つの文字、あるいは文字集団は､多義性の衣をまとっていて、与えられた文脈の中で常に正しくその意味が了解されるとは限らない。

この情報素材のもつ多義性は、情報の製作者（生産者）にとっては、より多くの（広範囲の）情報生成を可能にするもので、その意味では、（潜在的）情報量の増大を促すが、同時にこの多義性は､情報の利用者(消費者)にとっては、与えられた文脈の中で、情報素材に託された意味（情報内容）を正しく了解することへの阻害要因となり､その意味では、情報量の減少をもたらす。

本論文では、後者の立場に立ち、主に多義性に付随する情報素材としての文字または文字集団の「あいまいさ」をファジーエントロピーを用いて、定量的に評価し、その結果もたらされる情報量の減少を算定することにする。

上述した情報素材（文字または文字集団）のもつ「あいまいさ」は使用頻度の高いものほど大きいと仮定した。これは情報の製作者側では「あいまいさ」の高いものほど使いやすい素材であることにもとづく仮定である。そこで素材 Ai に対する、所属度関数 $\mu_{Ai}(x)$ を次図のようにモデル化する。

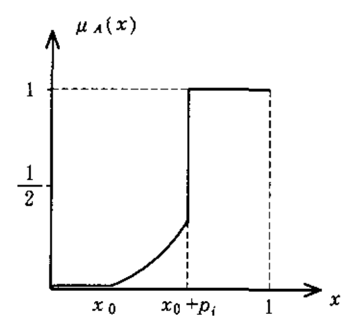

図 4

ここでは簡単のため、図４のように、情報素材 Ai（文字または文字集団）のもつ意味空間を x 軸上の閉空間［０，１］の集合にモデル化し、その中の p_i の幅だけ $\mu_{A_i}(x)$ が０から１までの値を連続的にとることにした。つまり Ai の意味空間全体に対して比率 $p_i(0 \leq p_i \leq 1)$ の部分だけ多義的あいまいさがあると仮定している。ここで p_i は既述の論文その（２）で定義した Ai の使用頻度である。

このときの Ai のもつ「あいまいさ」を先述したファジーエントロピーの定義式（１）で評価してみる。

ただし本論文においては、定義式（1）中の対数の底は、前報その（2）同様に2とする。

規格化定数kは

$$k=\frac{1}{\int_0^1 dx}=\frac{1}{1}=1$$ となる。

ゆえに、ファジーエントロピーは、

$$d(Ai)=-\int_0^1 dx\{\mu_{Ai}(x)\log\mu_{Ai}(x)+(1-\mu_{Ai}(x))\log(1-\mu_{Ai}(x))\} \cdots (2)$$

（$\log$ の底＝2）

となるが、x 軸上の $\mu_{Ai}(x)=0$ 及び $\mu_{Ai}(x)=1$ の部分は $d(A_i)$ への寄与は0であるから、

$d(Ai)=\int_{x_0}^{x_0+pi} dxF_{Ai}(x)$ と書ける…(3)

ただし、$F_{Ai}(x)=-\{\mu_{Ai}(x)\log\mu_{Ai}(x)+(1-\mu_{Ai}(x))\log(1-\mu_{Ai}(x))\}$ である。

ここで、積分変数を x から $x+x_0$ へ変換（平行移動）すると、

$d(Ai)=\int_0^{pi} dxF_{Ai}(x+x_0)$ となる…（4）

このとき、

$0\leq F_{Ai}(x+x_0)\leq 1$ …(5)が成り立つから、

$\mu_{Ai}(x+x_0)=0\ or\ 1$ のとき $F_{Ai}=0$ で、

$\mu_{Ai}(x+x_0)=\frac{1}{2}$ のとき $F_{Ai}=1$ である。

（5）式の両辺に積分 $\int_0^{pi} dx$ を施すと

$0\leq\int_0^{pi} F_{Ai}(x+x_0)dx\leq p_i$ を得る。

よって（4）式より

$d(Ai)$ は

$0\leq d(Ai)\leq p_i$ …（6）

を得る。

そこで、全ファジーエントロピー S_F を

$S_F=\sum_{i=1}^{m} d(Ai)$ …（7）で定義すると、

（m は素材の数である。）

（6）式より

$$\sum_{i=1}^{m} 0\leq\sum_{i=1}^{m} d(Ai)\leq\sum_{i=1}^{m} pi$$

$\therefore 0\leq S_F\leq 1$ が成り立つ。

このことから、全ファジーエントロピー S_F による情報エントロピーの減少効果を次式の形でとり入れることが可能となる。

$Stotal=S_I(1-S_F)$ …（8）

ここで S_I は、あいまいさ（＝全ファジーエントロピー S_F）がないときの情報エントロピー(Shannon のエントロピー)であり、Stotal はあいまいさ（＝全ファジーエントロピー S_F）による減少効果を加味した全情報エントロピーである。

次に $\mu_{Ai}(x)$ の区間（x_0, x_0+pi）における具体的な関数形を次のガウス分布（正規分布）で与えることにする

$$\mu_{Ai}(x)=exp\left[-\frac{(x-x_0)}{2\sigma^2}\right] \cdots (9)$$

このときファジーエントロピー $d(A_i)$ は、

$$\begin{aligned} d(Ai)&=\int_0^{pi} dx\ F_{Ai}(x+x_0)\\ &=-\int_0^{pi} dx\{\mu_{Ai}(x+x_0)\cdot\\ &\quad\cdot\log\mu_{Ai}(x+x_0)+(1-\mu_{Ai}(x+x_0)\cdot\\ &\quad\cdot\log(1-\mu_{Ai}(x+x_0))\}\\ &=-\int_0^{pi} dx\{exp(-\frac{x^2}{2\sigma^2})\cdot\\ &\quad\cdot\log exp(-\frac{x^2}{2\sigma^2})+\\ &\quad(1-exp(-\frac{x^2}{2\sigma^2}))\cdot\\ &\quad\cdot\log(1-exp(-\frac{x^2}{2\sigma^2}))\} \cdots (10) \end{aligned}$$

となる。

次節で、この評価式（10）を用いることにより、我々の人工言語系のもつ全ファジーエントロピー S_F を具体的に計算する。

D：人工言語系におけるファジーエントロピー

前報その（2）と同様、本論文においも情報素材（前述のAi）を単語とする。このとき、前節（7）式中の全情報素材の数mは言語系において用いられる全単語数となる。さらに日常言語においては、単語の使用頻度はZipf の法則に従うことが知られている。本論文の人工言語系においても、この事実に基づき、単語の出現頻度（pi）は前報その2と同様 Zipf の法則に基づき与えられるものとする。その結果ファジーエントロピーは、

$$S_F(\sigma)=\sum_{i=1}^{m} d(Ai,\ \sigma)\cdots\cdots\cdots\cdots (11)$$

ただし

$$d(Ai,\ \sigma)=-\int_0^{pi} dx\,\{\,exp(-\frac{x^2}{2\sigma^2})\cdot$$
$$\cdot\,log\,exp(-\frac{x^2}{2\sigma^2})$$
$$-(1-exp(-\frac{x^2}{2\sigma^2}))\cdot$$
$$\cdot\,log(1-exp(-\frac{x^2}{2\sigma^2}))\}$$

よってファジーエントロピー

$S_F(\sigma)=\sum_{i=1}^{m} d\,(Ai,\sigma)$ は単語数mと分散σの各値に応じて定まり、mとσの2変数関数とみなすことができる。そこで以下ファジーエントロピーを$S_F(m,\sigma)$で表すこととする。

我々は、種々なmとσの値に対する$S_F(m,\sigma)$の値をコンピュータによる数値計算で求めた。以下その結果を記し、考察を加える。まず$S_F(m,\sigma)$のmに対する依存性を知るために、$\sigma=\frac{1}{10}$で固定してmを50から600までの範囲で変えたときの各$S_F(m,\sigma)$の値が次の表1である。

表（1）の計算結果から、所属度関数の分散σが一定ならば､単語数mの増加に伴って、ファジーエントロピーS_Fは減少することが読み取れる。

表1　単語数とファジーエントロピーの関係
（$\sigma=\frac{1}{10}$）

単語数m	ファジーエントロピーS_F
50	0.2860
100	0.2354
200	0.1950
300	0.1752
400	0.1627
500	0.1537
600	0.1468

次に、$S_F(m,\ \sigma)$のσに対する依存性を知るために、$m=100$で固定して、σの1から$\frac{1}{180}$までの種々の値に対して$S_F(m,\sigma)$を求めた結果が次の表である。

表2　分散とファジーエントロピーの関係
（m＝100）

分散σ	ファジーエントロピーS_F
1	0.01295
1／64	0.3040
1／80	0.3053
1／90	0.305532
1／92	0.305544
1／93	0.305546
1／94	0.305545
1／95	0.305542
1／100	0.305487
1／120	0.304742
1／150	0.302500
1／180	0.299365

この表（2）の結果から、ファジーエントロピーS_Fは分散σに単調に依存しておらず、$\sigma=\frac{1}{93}$前後に最大値を有していることが分かる。次の表（3）は、このS_Fの最大値をより正確に知るために、$\sigma=\frac{1}{93}$の付近を更に細かく計算した結果である。

表3　ファジーエントロピー最大値付近の分散値　（m＝100）

σ	S_F
1／92.8	0.30554589
1／93.0	0.30554600
1／93.2	0.30554613
1／93.3	0.30554612
1／93.5	0.30554602

この表から$\sigma=\frac{1}{93.2}$のときの値$S_F=0.30554613$がほぼS_Fの最大値とみなすことができる。

S_Fをσの関数をみたとき、S_Fが最大値をもつことは、S_Fの定義式（11）の形から以下のように理解することができる。

式（11）で与えられた$d(Ai,\ \sigma)$は、$\sigma=0$、$\sigma=\infty$の両方に対して0となる。即ち$d(Ai,0)=d(Ai,\infty)=0$

よって　$S_F(m,0)=S_F(m,\infty)=0\cdots$（12）

更に、$S_F(m,\sigma)$がσの微分可能な関数であることより、ロルの定理から$\frac{\partial S_E}{\partial\sigma}=0$を満たす$\sigma$が存在する。これと$S_F>0$であることより$S_F$F は$\sigma$のある値で最大値をもつ。

上述の(12)式を、所属度関数$\mu(x)$に戻って考えてみると、次のような定性的理解が得られる。

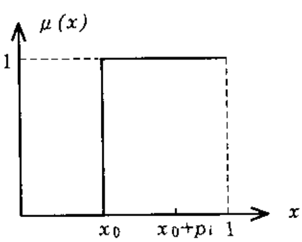

図5

まず$\sigma=0$のときは、区間$(x_0,\ x_0+pi)$における正規分布の中心$x=x_0$からの分散が0であるから図5のようになる。逆に$\sigma=\infty$のときは、正規分布の中心$x=x_0$からの分散が∞であるから図6のようになる。

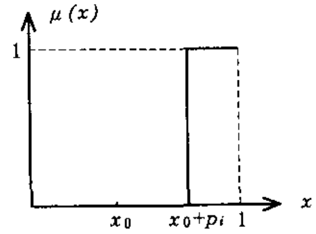

図6

両図とも所属度関数$\mu(x)$は0か1の二者択一的な値しかとらず、このことから、あいまいさが全くないことが明らかとなる。

最後に、単語数m＝200、分散$\sigma=\frac{1}{10}$で計算したファジーエントロピー$S_F(200,\ \frac{1}{10})$を用いて、全情報エントロピー

$Stotal=S_I(1-S_F)$を求める。

単語数m＝200の場合、$S_I=5.987$（前報その（2））、$S_F(200,\ \frac{1}{10})=0.1950$

よって、$Stotal=5.987\times(1-0.1950)$

$\fallingdotseq 4.818$

この結果は、各単語の意味空間［0，1］における所属度関数$\mu(x)$を区間幅piの中で分散$\sigma=\frac{1}{10}$の正規分布と設定しさらに、各単語のもつファジーエントロピーが使用頻度と相関をもつと仮定した全情報エントロピーの値である。

前述したように、情報の製作者（送り手）と利用者（受け手）という立場の相異によりファジー性の意味も異なる。本論文では、利用者の立場でのファジーの情報量への効果を定量的に扱かった。

上の計算をコンピュータで行なうに際しては、プログラム言語として Lattice C が用いられた。Step 数については、20～100の値についてあらかじめモニターした結果、信頼性を有する値を得るものとして最終的に Step 数＝80が計算に用いられた。

〈参考文献〉

(1) 寺岡　宏，矢吹哲夫　1993　知能作用過程のシュミレーション　その2　文字情報に関する理論的考察　北星短大紀要：29　67－73

(2) G. K. Zipf. 1949 Human bebavior and the princile of the least effort. Addison, Wesley Press Inc.

(3) C. E. Shannon. 1948 Mathematical theory of communication. Bell system Tech J. 27 379 and 623

(4) De Luca, A and S. Termini 1972 A definition of a nonprobabilistic entropy in the setting of fuzzy sets theory. In formation and Control, 20 301－312

(5) George J. Kliv and Tina A. Folger 本多中二訳：1993　ファジー情報学　日刊工業社