2カテゴリ分類困難度の情報理論

鈴木　昇一

Information Theory of Difficulty of Two-Category Classification

Shoichi Suzuki

あらまし

パターン認識の数学的理論(SS理論)では、入力パターンφに対応するパターンモデル$T\varphi$を求め、$T\varphi$から不動点パターンモデルを連想する形で、φの帰属するカテゴリを決定する多段階パターン変換連想形不動点認識法(SS連想形不動点認識法)が考えられている。

このパターン認識法を採用している本研究では、各出力出現確率$q(T\varphi)$($\varphi \in \Phi[t_1, t_2]$)と、出力$T\varphi$が観測された条件の下で各入力$\mathfrak{C}_n[j]$の再現確率$p(\mathfrak{C}_n[j]/T\varphi)$($n \in \{1, 2\}$)とを与え、

$$\mathrm{AMI}(\underline{\mathfrak{C}}[j]; \Phi[t_1, t_2]) = \sum_{n=1}^{2} \sum_{\varphi \in \Phi[t_1, t_2]} q(T\varphi) \cdot p(\mathfrak{C}_n[j]/T\varphi) \cdot \log_e[p(\mathfrak{C}_n[j]/T\varphi)/p(\mathfrak{C}_n[j])]$$

の最大値を求めることに関連し、2カテゴリ分類困難度$\mathrm{DOC}(\varphi, \underline{\mathfrak{C}}[j]; \mathrm{SM})$, 2カテゴリ分類容易度$\mathrm{EOC}(\varphi, \underline{\mathfrak{C}}[j]; \mathrm{SM})$, 曖昧度$H(\underline{\mathfrak{C}}[j]/T\varphi)$に関する解析が展開される。

曖昧度$H(\underline{\mathfrak{C}}[j]/T\varphi)$の、$p(\mathfrak{C}_1[j]/T\varphi)$に関する微係数を2カテゴリ分類困難度$\mathrm{DOC}(\varphi, \underline{\mathfrak{C}}[j]; \mathrm{SM})$と定義することから、解析が始まっているが、SS理論での不動点多段階想起形認識[B3], [B4]は解消される不確定さ$\mathrm{AMI}(\underline{\mathfrak{C}}[j]; \Phi[t_1, t_2])$が大きくなるようなパターン処理法であるとの結論が本研究によって鮮明にされる。

本研究は、SS理論のaxiom 2を満たす類似度関数SMを式(3.63)のように事後確率$p(\mathfrak{C}_1[j]/T\varphi)$と設定している故に(この設定は本研究独創性を確実なものにしている)、得られた研究内容は設定された1つの認識の働きが処理の対象とする問題のパターンφの集合$\Phi[t_1, t_2]$に適切なものであるかどうかを検証する場面で決定的な役割を果たすという意味で、認識システムRECOGNITRON [B3], t[B4] の構成に信頼を与えるものである。

キーワード

パターン認識の数学的理論（SS理論）　モデル構成作用素　類似度関数　大分類関数
不動点連想形多段階認識法　曖昧度　平均相互情報量　シグモイド関数
2カテゴリ分類困難度·容易度　最大類似度認識法

Abstract

A recognition system RECOGNITRON which has been presented in a mathematical theory (i.e.SS theory) of recognizing patterns suggested by S.Suzuki gets a corresponding pattern-model $T\varphi$ of an input pattern φ in question to be recognized, and determines a category to which φ belongs so that a fixed-point pattern-model that appeared on a final stage of a muti-stage structural-fertilization transformation of pattern-models may be recalled in such a way of solving a fixed-point equation of associative reconition about $T\varphi$.

In this recognition method an analysis about an difficulty DOC(φ, $\underline{\mathfrak{C}}[j]$; SM) and an easiness EOC($\underline{\mathfrak{C}}[j]$; SM) of binary classification , and an equivocation H($\underline{\mathfrak{C}}[j]/T\varphi$) is developed in full seeking for the maximum of an average amount

$$\mathrm{AMI}(\underline{\mathfrak{C}}[j];\ \Phi[t_1, t_2]\ )$$
$$=\sum_{n=1}^{2} \sum_{\varphi \in \Phi[t_1, t_2]}$$
$$q(T\varphi)\cdot p(\mathfrak{C}_n[j]/T\varphi)\cdot$$
$$\log_e[p(\mathfrak{C}_n[j]/T\varphi)/p(\mathfrak{C}_n[j])]$$

of mutual information, where

$q(T\varphi)(\varphi \in \Phi[t_1, t_2])$ is a probability of occurrences of the pattern-model $T\varphi$, and $p(\mathfrak{C}_n[j]/T\varphi)(n \in \{1, 2\})$ is a conditional probability of occurrences of the n-th category $\mathfrak{C}_n[j]$ given that $T\varphi$ has occurred at some trial.

The above-mentioned analysis begins with defining DOC(φ, $\underline{\mathfrak{C}}[j]$; SM) as an differential value of the function H($\underline{\mathfrak{C}}[j]/T\varphi$) concerning the variable $p(\mathfrak{C}_1[j]/T\varphi)$.We can conclude that the recognition method proposed by S.Suzuki maximizes AMI($\underline{\mathfrak{C}}[j]$; $\Phi[t_1, t_2]$) which is an amount of uncertainty of $\underline{\mathfrak{C}}[j] \equiv \{\mathfrak{C}_1[j], \mathfrak{C}_2[j]\}$ removed after many observations of $T\varphi$.

In the above analysis we adopt the similarity-measure SM satisfying axiom 2 as the aposteriori probability $p(\mathfrak{C}_1[j]/T\varphi)$, which makes sure of an originality of this study.

The obtained result can play a definite part in verifying whether or not a selected recognition method is suitable for the set $\Phi[t_1, t_2]$ of patterns in question, which therefore gives a reliability to a construction of the recognition system RECOGNITRON [B3], [B4].

Key words : a mathematical theory of recognizing patterns(SS theory)　model-construction operator
similarity-measure function　rough classifier
multi-stage recognition of fixed-point searching type　equivocation
average amount of mutual information　sigmoidal function
difficulty and an easiness of binary classification
recognition method using maximum similarity-measure

1. まえがき

複数のカテゴリが想定される場合、パターンφがその内の1つのカテゴリに帰属するか、しないかを決定することを、**2カテゴリ分類**(binary classification)［A3］という。**ユークリッド空**

間パターン(有限次元実数列)を採用し**単段階パターン変換**を基調とした2カテゴリ分類に関する汎化能力を学習の働き(学習アルゴリズム)によって如何に改善し、獲得するかについてはある程度、詳細な数理解析が可能である [A3]。本研究では、**ヒルベルト空間パターン**(無限次元関数[A5])を採用し、**多段階パターン変換**を基調とした2カテゴリ分類の働きを平均相互情報量の立場から、評価する手段を研究する。学習の問題については、SS理論 [B3], [B4] のaxiom 3を満たす大分類関数の設計問題として論じる。

本研究では、各出力出現確率 $q(T\varphi)(\varphi\in\Phi[t_1,t_2])$ と、出力 $T\varphi$ が観測された条件の下で各入力 $\mathfrak{C}_n[j]$ の再現確率 $p(\mathfrak{C}_n[j]/T\varphi)(n\in\{1,2\})$ とを与え、2カテゴリ分類に関する平均相互情報量

$$\begin{aligned}&\mathrm{AMI}(\underline{\mathfrak{C}}[j];\ \Phi[t_1,t_2])\\&=\sum_{n=1}^{2}\ \sum_{\varphi\in\Phi[t_1,t_2]}\\&q(T\varphi)\cdot p(\mathfrak{C}_n[j]/T\varphi)\cdot\\&\log_e[p(\mathfrak{C}_n[j]/T\varphi)/p(\mathfrak{C}_n[j])]\end{aligned}\tag{1.1}$$

の最大値を求めることに関連し、2カテゴリ分類困難度 $\mathrm{DOC}(\varphi,\underline{\mathfrak{C}}[j];\mathrm{SM})$, 2カテゴリ分類容易度 $\mathrm{EOC}(\varphi,\underline{\mathfrak{C}}[j];\mathrm{SM})$, 曖昧度 $H(\underline{\mathfrak{C}}[j]/T\varphi)$ に関する解析を展開する。

曖昧度 $H(\underline{\mathfrak{C}}[j]/T\varphi)$ の、$p(\mathfrak{C}_1[j]/T\varphi)$ に関する微係数を 2カテゴリ分類困難度 $\mathrm{DOC}(\varphi,\underline{\mathfrak{C}}[j];\mathrm{SM})$ と定義することから、解析が始まっている。このような定義、それに始まる諸解析はこれまで、パターン情報処理分野においては、全くなされていない。シャノン情報理論 [A1], [A4] において、2送信入力、多受信出力として、各々、2カテゴリ、多パターンを想定したことになるのであるが、条件付き確率 $p(\mathfrak{C}_1[j]/T\varphi)$ をSS理論でのaxiom 2を満たす類似度 $\mathrm{SM}(\varphi,\omega_j)$ と設定したこと(式(3.63)を参照)が本研究の新規性の始まりとなっている。

さて、

この入力パターン φ をそのカテゴリの代表パターンのモデルに
多段階にわたってパターン変換することにより、決定するとき
の分類困難度 (1.2)

をどのように定義すればよいのであろうか?

第 $j\in J$ 番目のカテゴリ $\mathfrak{C}j$ の代表パターン ω_j との間の類似度 $\mathrm{SM}(\varphi,\omega_j)$ が計算され得るとしよう。ならば、不等式

$$\mathrm{SM}(\varphi_1,\omega_j)\leqq\mathrm{SM}(\varphi_2,\omega_j)\tag{1.3}$$

が成り立つとき、

第 $j\in J$ 番目のカテゴリ $\mathfrak{C}_j$ に帰属すると決定するのに必要な手続きの煩雑さは、
パターン φ_1 はパターン φ_2 より少なくはない

といえよう。

2カテゴリ分類に関する困難さの程度を上述の定義は簡単に考えているが、対数尤度 $\log_e[\mathrm{SM}(\varphi_1,\omega_j)/\mathrm{SM}(\varphi_2,\omega_j)]$ を導入すれば、

式(1.3)の成立
⇔
$$\log_e[\mathrm{SM}(\varphi_1,\omega_j)/\mathrm{SM}(\varphi_2,\omega_j)]\leqq 0\tag{1.4}$$

が成立することに留意する。

ここで、

$$\mathrm{SM}(\varphi_2,\omega_j)=1-\mathrm{SM}(\varphi_1,\omega_j)\tag{1.5}$$

とおけば、2不等式(1.3), (1.4)は

$$SM(\varphi_1, \omega_j) \leqq 1 - SM(\varphi_1, \omega_j) \tag{1.6}$$

$$\Leftrightarrow$$

$$\log_e [SM(\varphi_1, \omega_j)/\{1 - SM(\varphi_1, \omega_j)\}] \leqq 0 \tag{1.7}$$

と書き換えられる。よって、

パターン φ_1 が第 $j \in J$ 番目のカテゴリ $\mathfrak{C}_j$ を表すように生成されたとき、第 $j \in J$ 番目のカテゴリ $\mathfrak{C}_j$ を表していない程度は、対数尤度

$$\log_e [SM(\varphi_1, \omega_j)/\{1 - SM(\varphi_1, \omega_j)\}] \tag{1.8}$$

であると考えられる。

処理の対象とする問題のパターン φ_1 のパターンモデル $T\varphi_1$ を導入し、初期段階で

$$\varphi_1[0] = T\varphi_1 \tag{1.9}$$

と設定し、パターンモデルの多段階変換

$$\varphi_1[1], \ \varphi_1[2], \ \cdots, \ \varphi_1[t] \tag{1.10}$$

によって、

$$SM(\varphi_1[t], \omega_j) = 1 \tag{1.11}$$

を満たす第t段階(最終段階)のパターンモデル $\varphi_1[1] = T\psi_1[t]$ を求めるSS理論 [B3], [B4] の観点からは、式(1.8)の対数尤度が、

パターン φ_1 を第 $j \in J$ 番目のカテゴリ $\mathfrak{C}_j$ に
帰属すると認識するときの**容易度** (1.12)

を表していると考えられよう。よって、

$$-\log_e [SM(\varphi_1, \omega_j)/\{1 - SM(\varphi_1, \omega_j)\}]$$
$$= \log_e [\{1 - SM(\varphi_1, \omega_j)\}/SM(\varphi_1, \omega_j)] \tag{1.13}$$

は困難度を表していると考えてよいだろう。本研究はこの観点を採用する。

2カテゴリ分類に関する第 $j \in J$ 番目のカテゴリ $\mathfrak{C}_j$ に関する条件エントロピーを、パターンモデル $T\varphi$ が出現したときの $\mathfrak{C}_j$ の条件付き確率と解釈できるような類似度関数 $SM(\varphi, \omega_j)$ ($= SM(T\varphi, \omega_j)$) で定義し、このエントロピーの、$SM(\varphi, \omega_j)$ に関する偏微分係数で、2カテゴリ分類困難度を定義したらどうかというのが、本研究であり、従来のパターン認識研究に類をみない(新規性)。

式(1.8)の対数尤度 loge $[SM(\varphi_1, \omega_j)/\{1 - SM(\varphi_1, \omega_j)\}]$ は、2カテゴリ分類に関する平均情報量(式(3.46)を参照)

$$H(\underline{\mathfrak{C}}[j]/T\varphi)$$
$$= \sum_{n=1}^{2} p(\mathfrak{C}_n[j]/T\varphi) \cdot \mathrm{logep}(\mathfrak{C}_n[j]/T\varphi)$$
$$= -SM(\varphi_1, \omega_j) \cdot \log_e SM(\varphi_1, \omega_j)$$
$$- \{1 - SM(\varphi_1, \omega_j)\} \cdot \log_e \{1 - SM(\varphi_1, \omega_j)\} \tag{1.14}$$

の密度であることが、Gelombの先駆的研究 [A2] から知られ(4式(3.3)～(3.6)を参照)、この事実が本研究の推進力となっている。

SS理論での不動点多段階想起形認識 [B3], [B4] は解消される不確定さ $AMI(\underline{\mathfrak{C}}[j]\,;\Phi[t_1, t_2])$ が大きくなるようなパターン処理法であるとの結論が本研究によって鮮明にされたといえよう(有効性)。

本研究は、SS理論のaxiom 2を満たす類似度関数 SM を式(3.63)のように事後確率 $p(\mathfrak{C}_1[j]/T\varphi)$と設定している故に、得られた研究内容は認識の働きが処理の対象とする問題のパターン φ の集合 $\Phi[t_1, t_2]$ に適切なものであるかどうかを検証する場面に決定的な役割を果たすという意味で、信頼のおけるものである。(新規性·信頼性)。

尚、これまでの文献BでのS.Suzuki諸研究に関連して、付録A～Fが設けられている。

2. パターンモデル Tφと、類似度関数 SM

本章では、axiom 1を満たす対［Φ, T］と、axiom 2を満たす類似度関数 SM が説明される。

2.1 処理の対象とするパターンφの集合 Φ とモデル構成作用素 T の対[Φ,T] の構成

Tφを見たり聞いたりしたならば、φと同じように見えたり聞こえたりするような"パターン$\varphi \in \Phi$に対応するパターンモデル(**同一知覚原理を満たすパターーモデル**)$T\varphi \in \Phi$を出力する"モデル構成作用素"

$$T : \Phi \rightarrow \Phi \tag{2.1}$$

を考えよう。ここに、Φは処理の対象とする問題のパターンΦの集合であり、次の定理2.1の式(2.2)で与えられる。

実は、対［Φ, T］が次のaxiom 1を満たすように構成されるとき、式(2.1)の写像 T はモデル構成作用素(model－construction operator)と呼ばれる［B3]，［B4］：

<u>Axiom 1</u> (パターン集合Φとモデル構成作用素 T との対【Φ, T】の満たすべき公理)

(ｉ)(零元のT－**不動点性**；fixed－point property of zero element under mapping T) $0 \in \Phi \wedge T0 = 0.$

(ii)(**錐性**，正定数倍吸収性；cone property)

$\forall \varphi \in \Phi, a \cdot \varphi \in \Phi \wedge T(a \cdot \varphi) = T\varphi$

for any positive real number a.

(iii)(**ベキ等性**，埋込性；idempotency, embeddedness)

$\forall \varphi \in \Phi, T\varphi \in \Phi \wedge T(T\varphi) = T\varphi.$

(iv)(写像 T の**非零写像性**；non-zero mapping property of T) $\exists \varphi \in \Phi, T\varphi \neq 0.$ □

上述のaxiom 1を満たす対［Φ, T］の構成が可能であることは、次の定理2.1［B3]，［B4］で指摘される。

[定理2.1]（**モデル構成作用素 T の基本構成定理**）

写像 T がaxiom 1の(ｉ)，(ii)，(iii)の3後半，並びに、(iv)を満たすとしよう。そして、パターンと判明しているφの集合Φ_Bが与えられたとしよう。ならば、処理の対象とする問題のパターンφの集合Φを

$$\begin{aligned}\Phi &= R^{++}(\Phi_B \cup T \cdot \Phi_B) \\ &\equiv \{r^{++}\varphi \mid \varphi \in \Phi_B, r^{++} \in R^{++}\} \\ &\quad \cup \{r^{++}T\varphi \mid \varphi \in \Phi_B, r^{++} \in R^{++}\} \\ &\quad \text{where } R^{++} \text{ is a set of positive real numbers}\end{aligned} \tag{2.2}$$

の如く設定すれば、

$$\Phi \supset \{0\} \wedge [a \cdot \Phi = \Phi \text{ for any } a \in R^{++}] \wedge$$
$$[T \cdot \Phi = T \cdot \Phi_B \subset \Phi] \tag{2.3}$$
が成立し、axiom 1の(i), (ii), (iii)の3前半を Φ は満たし、結局、対［Φ, T］はaxiom 1を満たす。 □

SS理論［B1］～［B6］では、パターン φ は可分な(separable)**一般抽象ヒルベルト空間**(Hilbert space) $\mathfrak{H}$ の元とする。内積は(φ, η)と表され、ノルムは $\|\varphi\| \equiv \sqrt{(\varphi, \varphi)}$で表される。ここに、$\mathfrak{H}$ が可分とは、稠密な(dense)可算部分集合が $\mathfrak{H}$ に存在することを指す。$\varphi, \eta \in \mathfrak{H}$ 間のノルム距離 $\|\varphi - \eta\| = \sqrt{(\varphi - \eta, \varphi - \eta)}$ に注意しておこう。

理解のためには、例えば、特別な場合として、内積(φ, η)を、
$$(\varphi, \eta) = \int_M dm(x)\, \varphi(x) \cdot \overline{\eta}(x) \tag{2.4}$$
ここに、$\overline{\eta}$ は η の複素共役(a complex conjugate of η)であり、

M ： q次元ユークリッド空間 R^q の可測部分集合 (2.5)

dm(x)：正値Lebesgue-Stieltjes式測度 (2.6)

$$x = \langle x_1, x_2, \cdots, x_q \rangle \in M (\subseteq R^q) \tag{2.7}$$
とする可分なヒルベルト空間 $\mathfrak{H} = L2(M ; dm)$で考えておけばよい［B1］。

2.2 axiom 2と類似度関数 SM

“正常なパターン”(well-formed pattern)は、ある1つのカテゴリ $\mathfrak{C}_j$ (第$j \in J$番目の類概念)のみに帰属しているものとし、このような $\mathfrak{C}_j$ の集まり(有限集合)
$$\underline{\mathfrak{C}} \equiv \{\mathfrak{C}_j \mid j \in J\} \tag{2.8}$$
を想定する。$\mathfrak{C}_j$ の備えている性質を典型的に備えている代表パターン(prototypical pattern) $\omega_j (\neq 0)$を1つ選定する。$\mathfrak{C}_j$ は、典型(prototype)としての代表パターン ω_j を中心とした緩やかなカテゴリであることを仮定したことに注意しておく。ここに、
$$\Omega \equiv \{\omega_j \mid j \in J\} \subset \Phi \tag{2.9}$$
が式(2.8)の全カテゴリ集合 $\underline{\mathfrak{C}}$ に対応する代表パターンの集合である。式(2.9)の系 Ω は、

複素定数 a_j の組 $\{a_j \mid j \in J\}$ について
$$\sum_{j \in J} a_j \cdot \omega_j = 0 \Rightarrow \forall j \in J, a_j = 0 \tag{2.10}$$
が成立しているという意味で、**1次独立**(linearly independent)でなければならない。

axiom 1を満たす式(2.1)のモデル構成作用素 T によって、式(2.9)の代表パターン集合 Ω が変換されて得られる系
$$T \cdot \Omega \equiv \{T\omega \mid \omega \in \Omega\} = \{T\omega_j \mid j \in J\} \tag{2.11}$$
も1次独立であると要請する。このとき、**類似度関数**(similarity-measure function)
$$SM : \Phi \times \Omega \rightarrow \{s \mid 0 \leqq s \leqq 1\} \tag{2.12}$$
を導入し、

$SM(\varphi, \omega_j) = 1, 0$ に従って、パターン $\varphi \in \Phi$は各々、

ω_j と確定的な類似関係、相違関係にあり、

また、$0 < SM(\varphi, \omega_j) < 1$ の場合は、曖昧な類似・相違関係にある (2.13)

と、SM を解釈しよう。関数 SM は次のaxiom 2を満たすように構成されねばならない。Kronecker(クロネッカー)のデルタ記号
$$\delta_{ij} = 1 \text{ if } i = j, = 0 \text{ if } i \neq j \tag{2.14}$$

を導入しておく。

Axiom 2 (**類似度関数 SM の満たすべき公理**)

(i) (規格化**直交性**；orthonormality)

$\forall i, \forall j \in J, SM(\omega_i, \omega_j) = \delta_{ij}.$

(ii) (規格化条件，正規性；probability condition, normality)

$\forall \varphi \in \Phi, \sum_{j \in J} SM(\varphi, \omega_j) = 1.$

(iii) (写像 T の下での**不変性**；invariance under mapping T)

$\forall \varphi \in \Phi, \forall j \in J, SM(T\varphi, \omega_j) = SM(\varphi, \omega_j).$ □

第$j \in J$番目のカテゴリ $\mathfrak{C}_j$ の出現確率 $p(\mathfrak{C}_j)$ を導入しておく。確率性質

$$[\forall j \in J, 0 < p(\mathfrak{C}_j) < 1] \wedge [\sum_{j \in J} p(\mathfrak{C}_j) = 1] \qquad (2.15)$$

を満たしていなければならない。

3. 2カテゴリ分類困難度 DOC (φ, $\underline{\mathfrak{C}}$[j]；SM) の理論

本章では、パターンφをある1つのカテゴリに帰属するかどうかを決定するときの困難度(2カテゴリ分類困難度) DOC(φ, $\underline{\mathfrak{C}}$[j]；SM)を定義するために、
平均相互情報量 AMI($\underline{\mathfrak{C}}$[j]；Φ [1, tmax])、曖昧度 H($\underline{\mathfrak{C}}$[j]/Tφ)を導入し、2カテゴリ分類困難度に関する解析を展開する。

3.1 2分類に関する平均情報量の密度関数

シャノン情報理論 [A1] によれば、確率$(0 \leqq) p (\leqq 1)$で出現する確率事象は、

$$-\log_e p \qquad (3.1)$$

だけの**不確定さ**(uncertainty)を持っている。この入力確率事象に対応して、出力を観測した条件の下でのこの確率事象の条件付き出現確率が 1 で観測された場合(出力を観測することによって、この入力確率事象が生起したことを一意的に知ることが可能で、観測者がこの入力確率事象の持つすべての情報を入手できる場合)、この不確定さ $-\log_e p$ が 0 に解消すると考えている。

この種の入力確率事象に対応する出力を限りなく多数回観測することが繰り返されると、観測者は平均的に、1事象当たりの**平均情報量**(average amount of information)、或いは、エントロピー(entropy)と称される "期待値としての非負量"

$$\begin{aligned} &H(p) \\ &\equiv [\text{その入力確率事象が出現する確率}] \cdot [-\log_e p] \\ &+ [\text{その入力確率事象が出現しない確率}] \cdot [-\log_e (1-p)] \\ &= -p \cdot \log_e p - (1-p) \cdot \log_e (1-p) \qquad (3.2) \end{aligned}$$

を受け取り、H(p) だけの平均的不確定さが1事象当たり解消されることが期待されることになる。

y を変数とするエントロピー関数(entropy function)

$$\begin{aligned} &h(v) \\ &\equiv -v\log_e v - [1-v] \log_e [1-v] \qquad (3.3) \\ &\qquad (0 \leqq v \leqq 1) \end{aligned}$$

を導入すれば、積分公式

$$\int_c^d dv\ \log_e[v/(1-v)]$$
$$=h(c)-h(d), \tag{3.4}$$
where
$$0 \leqq c,\ d \leqq 1 \tag{3.5}$$

が成立ち［A2］、登場している関数

$$\log_e[v/(1-v)]\ (0<v<1) \tag{3.6}$$

は**情報量密度関数**(information density function)と呼ばれてよい。

3.2 平均相互情報量 AMI ($\underline{\mathfrak{C}}$[j]；Φ[1, tmax])と、曖昧度 H($\underline{\mathfrak{C}}$[j]/Tφ)

パターン事例 φ_t の系列

$$\Phi[t_1, t_2] \equiv \{\varphi_t \mid t=t_1, t_1+1, \cdots, t_2\},$$
$$1 \leqq t_1 \leqq t_2 \leqq \mathrm{tmax} \tag{3.7}$$

を導入する。

以後、

$$t_1=1,\ t_2=\mathrm{tmax} \tag{3.8}$$

を採用することが基本となる。同時に、カテゴリ番号 j∈J を1つ任意に選定し、固定することがある。任意のカテゴリ番号 j∈J について、2つのカテゴリ部分集合

$$\mathfrak{C}_1[j] \equiv \{\mathfrak{C}_j\},\ \mathfrak{C}_2[j] \equiv \underline{\mathfrak{C}} - \{\mathfrak{C}_j\} \tag{3.9}$$

を導入し、

$$\underline{\mathfrak{C}}[j] \equiv \{\mathfrak{C}_1[j],\ \mathfrak{C}_2[j]\} \tag{3.10}$$

とおく。

2条件

$$[\forall \varphi \in \Phi[t_1, t_2],\ 0<q(T\varphi)<1] \tag{3.11}$$
$$\wedge\ [\Sigma_{\varphi \in \Phi[t_1, t_2]}\ q(T\varphi)=1] \tag{3.12}$$

を満たす“パターンモデル Tφ の出現確率” q(Tφ) と、2条件

$$\forall \varphi \in \Phi[t_1, t_2],$$
$$[\forall n \in \{1, 2\},\ 0 \leqq p(\mathfrak{C}_n[j]/T\varphi) \leqq 1] \tag{3.13}$$
$$\wedge [\sum_{n=1}^{2} p(\mathfrak{C}_n[j]/T\varphi)=1] \tag{3.14}$$

を満たす“パターンモデル Tφ が出現したときの、カテゴリ $\mathfrak{C}_n$[j]の条件付き出現確率” p($\mathfrak{C}_n$[j]/Tφ) とを考える。

このとき、“カテゴリ $\mathfrak{C}_n$[j]の出現確率” p($\mathfrak{C}_n$[j]) を

$$p(\mathfrak{C}_n[j])$$
$$\equiv \Sigma_{\varphi \in \Phi[t_1, t_2]}\ q(T\varphi) \cdot p(\mathfrak{C}_n[j]/T\varphi),$$
$$n \in \{1, 2\} \tag{3.15}$$

とおくと、確率条件

$$\sum_{n=1}^{2} p(\mathfrak{C}_n[j])=1$$
$$\because\ 2式(3.14),\ (3.12) \tag{3.16}$$

が成り立つ。

処理の対象とする問題のパターン$\varphi \in \Phi[t_1, t_2]$を観測して、$\varphi \in \Phi$ が第 $j \in J$ 番目のカテゴリ$\mathfrak{C}_j \in \mathfrak{C}_1[j]$ に帰属するか、$\mathfrak{C}_2[j]$ 内の、いずれか1つのカテゴリ $\mathfrak{C}_i$ に帰属するかを決定するように、2カテゴリ分類する働きを考えよう。この場合、$\underline{\mathfrak{C}}[j]$の2要素$\mathfrak{C}_1[j]$, $\mathfrak{C}_2[j]$ (入力)は直接には観測できない。ところが、$\Phi[t_1, t_2]$の各元 φ_t(出力)は直接、観測できる。直接観測できる出力φ_tから情報を得ることにより、この出力に対応して2入力$\mathfrak{C}_1[j]$, $\mathfrak{C}_2[j]$ が推論によって決定できる。出力 φ_tから得る情報は十分でないので、この種の推論は**帰納推論**(inductive reasoning)にならざるを得ない。以後、この入出力関係の想定の下で論を組み立てる。

$\underline{\mathfrak{C}}[j]$と$\Phi[t_1, t_2]$との間の平均相互情報量(average amount of mutual information)

$$\begin{aligned}&\mathrm{AMI}(\underline{\mathfrak{C}}[j]\,;\,\Phi[t_1, t_2])\\ &\equiv \Sigma_{\varphi\in\Phi[t_1,t_2]} \sum_{n=1}^{2} q(T\varphi)\cdot p(\mathfrak{C}_n[j]/T\varphi)\\ &\cdot \log_e[p(\mathfrak{C}_n[j]/T\varphi)/p(\mathfrak{C}_n[j])]\end{aligned} \tag{3.17}$$

を導入する。この式(3.17)の $\mathrm{AMI}(\underline{\mathfrak{C}}[j]\,;\,\Phi[t_1, t_2])$は2分類平均相互情報量、或いは、2カテゴリ平均相互情報量と称されてよい。

先ず、次の補助定理3.1の成立に注意する。

［補助定理3.1］（**エントロピー関数の最小性**）

$$x_k > 0 \wedge \sum_{k\in K^+} y_k \leqq \sum_{k\in K^+} x_k \tag{3.18}$$

を満たす添字 k の集合 $K^+(\subseteq K)$ (3.19)

を導入すると、2つの数列$\{x_k\}_{k\in K}$, $\{y_k\}_{k\in K}$に関し、不等式

$$-\sum_{k\in K^+} x_k\cdot\log_e x_k \leqq -\sum_{k\in K^+} x_k\cdot\log_e y_k \tag{3.20}$$

が成り立つ。ここで、等号は、

$$[\forall k\in K^+, y_k/x_k = 1] \tag{3.21}$$

$$\wedge\,[\sum_{k\in K+} y_k = \sum_{k\in K+} x_k] \tag{3.22}$$

が成立するとき、且つ、その時に限る。 □

上述の補助定理3.1を適用すれば、

$$\begin{aligned}&\mathrm{AMI}(\underline{\mathfrak{C}}[j]\,;\,\Phi[t_1, t_2])\\ &= \Sigma_{\varphi\in\Phi[t_1,t_2]}\, q(T\varphi)\cdot\sum_{n=1}^{2} p(\mathfrak{C}_n[j]/T\varphi)\\ &\cdot \mathrm{loge}\ [p(\mathfrak{C}_n[j]/T\varphi)/p(\mathfrak{C}_n[j])]\\ &\geqq \Sigma_{\varphi\in\Phi[t_1,t_2]}\ q(T\varphi)\cdot 0\\ &\quad \because\ \text{補助定理3.1}\\ &\leqq 0、\end{aligned}$$

を得る。つまり、非負性 (3.23)

$$\forall t_1, \forall t_2, \forall j\in J, 0 \leqq \mathrm{AMI}(\underline{\mathfrak{C}}[j]\,;\,\Phi[t_1, t_2]) \tag{3.24}$$

が成り立つ。特に、零性

$$0 = \mathrm{AMI}(\underline{\mathfrak{C}}[j]\,;\,\Phi[t_1, t_2]) \tag{3.25}$$

が成立するのは、

$$\forall n\in\{1, 2\},\ p(\mathfrak{C}_n[j]) = p(\mathfrak{C}_n[j]/T\varphi) \tag{3.26}$$

の時に限る。ここに、式(3.15)の $p(\mathfrak{C}_n[j])$の表現に注意しておく。

$\underline{\mathfrak{C}}[j]$の平均的不確定さ $H(\underline{\mathfrak{C}}[j])$

$$\equiv -\sum_{n=1}^{2} p(\mathfrak{C}_n[j])\cdot\log_e p(\mathfrak{C}_n[j]) \geqq 0 \tag{3.27}$$

を導入すれば、この非負量 $H(\underline{\mathfrak{C}}[j])$ は、

$\mathfrak{C}_j$ と $\mathfrak{C}-\{\mathfrak{C}_j\}$ の任意のカテゴリとの間が分離できなければできないほど大きい値をとる量であり、その最小値、最大値は各々、

$$\begin{cases} 0 \quad \text{if } p(\mathfrak{C}_1[j])=0 \vee p(\mathfrak{C}_1[j])=1 & (3.28) \\ \log_e 2 & \\ \quad \text{if } \forall n\in\{1, 2\},\ p(\mathfrak{C}_1[j])=p(\mathfrak{C}_2[j])=1/2 & (3.29) \end{cases}$$

であり、不等式

$$\forall j\in J,\ 0\leqq H(\underline{\mathfrak{C}}[j])\leqq \log_e 2 \tag{3.30}$$

が成り立つことが、次の2補助定理3.2, 3.3からわかる。

［補助定理3.2］（**エントロピー関数のとる値の範囲**）

$$x_k>0\wedge 1=\sum_{k\in K+} x_k \tag{3.31}$$

を満たす添字 k の集合 $K^+(\subseteq K)$ (3.32)

を導入すれば、不等式

$$0\leqq -\sum_{k\in K+} x_k\cdot\log_e x_k\leqq \log_e |K^+| \tag{3.33}$$

が成り立つ。ここに、

$$(\mathrm{i})\ 0=-\sum_{k\in K^+} x_k\cdot\log_e x_k \tag{3.34}$$

$$\Leftrightarrow$$

$$\exists k\in K+,\ x_k=1\wedge[\forall \ell\in K-\{k\},\ x_\ell=0] \tag{3.35}$$

$$(\mathrm{ii})\ \log_e |K^+|=-\sum_{k\in K^+} x_k\cdot\log_e x_k \tag{3.36}$$

$$\Leftrightarrow$$

$$\forall k\in K^+,\ x_k=1/|K^+|. \tag{3.37}$$

（証明） 等式

$$0\cdot\log_e 0=1\cdot\log_e 1=0 \tag{3.38}$$

と、不等式

$$\forall x\in\{z\mid 0\leqq z\leqq 1\},\ -\log_e x\geqq 0 \tag{3.39}$$

に注意し、補助定理3.1において、

$$\forall k\in K^+,\ y_k=1/|K^+| \tag{3.40}$$

とおけば、不等式

$$0\leqq -\sum_{k\in K^+} x_k\cdot\log_e x_k\leqq -\sum_{k\in K^+} x_k\cdot\log_e y_k$$
$$=\log_e |K^+| \tag{3.41}$$

の成立がわかる。残りの成立は次の補助定理3.3から、明らかである。 □

［補助定理3.3］（**エントロピーの減少定理**）

確率条件

$$[\forall q\in K,\ 0\leqq x_q\leqq 1]\wedge\sum_{q\in K} x_q=1 \tag{3.42}$$

の下では、

相異なる $k, m\in K$ に対し、ある非負実数 δ が存在して、

$$0\leqq x_k\leqq x_m\leqq 1 \tag{3.43}$$

$$\wedge$$

$$0\leqq x_k'\equiv x_k-\delta\leqq x_m'\equiv x_m+\delta\leqq 1$$
$$\wedge[\forall q\in K-\{k, m\},\ x_q'\equiv x_q] \tag{3.44}$$

であれば、不等式

$$-\sum_{q\in K} x_q' \cdot \log_e x_q' \leqq -\sum_{q\in K} x_q \cdot \log_e x_q \tag{3.45}$$

が成り立つ。等号が成り立つのは、$\delta=0$ のときに限る。 □

2条件式(3.15), (3.16)に注意し、パターンモデル $T\varphi$ を観測した条件の下で残存しているであろう $\underline{\mathfrak{C}}$ [j] の平均的不確定さ、つまり、曖昧度(equivocation)

$$H(\underline{\mathfrak{C}}[j]/T\varphi)$$
$$\equiv -\sum_{n=1}^{2} p(\mathfrak{C}_n[j]/T\varphi)\cdot \log_e p(\mathfrak{C}_n[j]/T\varphi) \tag{3.46}$$

を定義する。この非負量 $H(\underline{\mathfrak{C}}[j]/T\varphi)$は、

パターン φ に対応するパターンモデル $T\varphi$が確保された
条件の下で、$\mathfrak{C}_j$ と $\mathfrak{C}-\{\mathfrak{C}_j\}$の任意のカテゴリとの間が
分離できなければできないほど大きい値をとる量であり、
その最小値、最大値は各々、

$$\begin{cases} 0 \quad \text{if} \quad p(\mathfrak{C}_1[j]/T\varphi)=0 \vee p(\mathfrak{C}_1[j]/T\varphi)=1 & (3.47) \\ \text{loge2} \\ \text{if} \quad \forall n\in\{1, 2\},\ p(\mathfrak{C}_1[j]/T\varphi)=p(\mathfrak{C}_2[j]/T\varphi)=1/2 & (3.48) \end{cases}$$

であり、不等式

$$\forall j\in J,\ 0\leqq H(\underline{\mathfrak{C}}[j]/T\varphi)\leqq \log_e 2 \tag{3.49}$$

が成り立つことが、上述の2補助定理3.2, 3.3からわかる。

更に、パターンモデル集合

$$T\cdot\Phi[t_1, t_2]\equiv\{T\varphi_t \mid t_1\leqq t\leqq t_2\} \tag{3.50}$$

として、式(3.46)の $H(\underline{\mathfrak{C}}[j]/T\varphi)$をパターンモデル $T\varphi$ の、2条件式(3.11), (3.12)の出現確率 $q(T\varphi)$で平均化して得られる非負量

$$H(\underline{\mathfrak{C}}[j]/T\cdot\Phi[t_1, t_2])$$
$$\equiv \Sigma_{\varphi\in\Phi[t_1, t_2]}\ q(T\varphi)\cdot H(\underline{\mathfrak{C}}[j]/T\varphi) \tag{3.51}$$

を定義する。この非負量 $H(\underline{\mathfrak{C}}[j]/T\cdot\Phi[t_1, t_2])$については、次の解釈が可能である：

任意のパターン $\varphi\in\Phi[t_1, t_2]$ に対応するパターンモデル $T\varphi$が
確保された条件の下で、$\mathfrak{C}_j$ と $\mathfrak{C}-\{\mathfrak{C}_j\}$ の任意のカテゴリとの
間が分離できなければできないほど大きい値をとる平均量である。 □

そうすれば、式(3.17)の平均相互情報量 $AMI(\underline{\mathfrak{C}}[j]\ ;\ \Phi[t_1, t_2])$は、

$$AMI(\underline{\mathfrak{C}}[j]\ ;\ \Phi[t_1, t_2])$$
$$=H(\underline{\mathfrak{C}}[j])$$
$$-\Sigma_{\varphi\in\Phi[t_1, t_2]}\ q(T\varphi)\cdot H(\underline{\mathfrak{C}}[j]/T\varphi)$$
$$\because \text{式}(3.15) \tag{3.52}$$
$$=H(\underline{\mathfrak{C}}[j])$$
$$-H(\underline{\mathfrak{C}}[j]/T\cdot\Phi[t_1, t_2])$$
$$\because \text{式}(3.30) \tag{3.53}$$

と再表現される。この最後の式(3.53)からは、

$$AMI(\underline{\mathfrak{C}}[j]\ ;\ \Phi[t_1, t_2])$$

＝【$\underline{\mathfrak{C}}$ [j]の各元 $\mathfrak{C}_1[j]$, $\mathfrak{C}_2[j]$が平均的に持っている不確定さ】

－【$\Phi[t_1, t_2]$ の任意の元 φ が確保された後、
$\underline{\mathfrak{C}}[j]$ の各元 $\mathfrak{C}_1[j]$, $\mathfrak{C}_2[j]$ が平均的に持って
いる不確定さ】 (3.54)
＝$\Phi[t_1, t_2]$ の任意の元 φ が確保されたことが原因となって、
$\underline{\mathfrak{C}}[j]$ の各元 $\mathfrak{C}_1[j]$, $\mathfrak{C}_2[j]$ が平均的に持っている不確定さの内、
解消された非負量 (3.55)
＝$\Phi[t_1, t_2]$ の任意の元 φ が確保されたことが
原因となって、$\underline{\mathfrak{C}}[j]$ の各元 $\mathfrak{C}_1[j]$, $\mathfrak{C}_2[j]$ に
関し、平均的に得られた情報量 (3.56)

という解釈が可能である。

尚、積分公式(3.4)を適用すれば、式(3.17)の平均相互情報量 $\mathrm{AMI}(\underline{\mathfrak{C}}[j]\,;\,\Phi[t_1,t_2])$ については、式(3.16)の情報量密度関数によって、その積分表示式

$$\mathrm{AMI}(\underline{\mathfrak{C}}[j]\,;\,\Phi[t_1, t_2])$$
$$=\Sigma_{\varphi\in\Phi[t_1,t_2]}\ q(T\varphi)\cdot$$
$$\int_a^b dv\ \log_e[v/(1-v)] \tag{3.57}$$

where

$$a=p(\mathfrak{C}_1[j]) \tag{3.58}$$
$$b=p(\mathfrak{C}_1[j]/T\varphi) \tag{3.59}$$

が可能である。

3.3 2カテゴリ分類困難度 DOC $(\varphi, \underline{\mathfrak{C}}[j]\,;\,\mathrm{SM})$

以後、正条件

$$q_t>0 \ \text{ for any } (t_1\leqq)t(\leqq t_2) \tag{3.60}$$

を満たす各 p_t を選び、2条件式(3.11), (3.12)を満たす確率 $q(T\varphi)$ を、

$$q(T\varphi_t)=q_t/\sum_{s=t_1}^{t_2} qs$$
$$(t_1\leqq t\leqq t_2) \tag{3.61}$$

とおく。例えば、

$$q_t$$
$$=[2\pi\sigma(t)^2]^{-1/2}\cdot\exp[-\|T\varphi_t\|^2/\{2\cdot\sigma(t)^2\}]$$
$$>0, \text{ where } \sigma(t)>0 \text{ for any } t \tag{3.62}$$

とおけばよい。

以後、axiom 2を満たす式(2.12)の類似度関数 SM を1つ、採用し、2条件式(3.13), (3.14)を満たす各 $p(\mathfrak{C}n[j]/T\varphi)$ を

$$p(\mathfrak{C}_1[j]/T\varphi)=\mathrm{SM}(\varphi, \omega_j) \tag{3.63}$$
$$p(\mathfrak{C}_2[j]/T\varphi)=1-\mathrm{SM}(\varphi, \omega_j)$$
$$=\sum_{i\in J-\{j\}} \mathrm{SM}(\varphi, \omega_j) \tag{3.64}$$

とおく。

2式(3.63), (3.64)の設定は、

“モデル $T\varphi$ を見たり聞いたりしたならば、原パターン φ と
同じように見えたり、聞こえたりする (3.65)

という **"同一知覚原理"** の下で、

確率 $SM(\varphi, \omega_j)$ でパターン φ が第 $j \in J$ 番目のカテゴリ $\mathfrak{C}_j$ に帰属し、
確率 $1-SM(\varphi, \omega_j)$ でパターン φ が第 $j \in J$ 番目のカテゴリ $\mathfrak{C}_j$ に
帰属しない (3.66)

という **"binary recognitionの働き"** を想定していることになる。

2式(3.63), (3.64)のこの設定の下で、式(3.46)の $H(\underline{\mathfrak{C}}[j]/T\varphi)$ を、

$$y \equiv SM(\varphi, \omega_j) \tag{3.67}$$

の関数 $f(y)$ と考えると、

$$f(y) \equiv H(\underline{\mathfrak{C}}[j]/T\varphi) \tag{3.68}$$

とみなせることに注意して、**2カテゴリ分類困難度**(difficulty of binary classification) $DOC(\varphi, \underline{\mathfrak{C}}[j]; SM)$ を $SM(\varphi, \omega_j)$ に関する微分係数として、

$$DOC(\varphi, \underline{\mathfrak{C}}[j]; SM) \equiv df(y)/dy \tag{3.69}$$

と定義しよう。2カテゴリ分類困難度 $DOC(\varphi, \underline{\mathfrak{C}}[j]; SM)$ は無論、

パターンモデル $T\varphi \in \Phi[t_1, t_2]$ が確保された条件の下で、
$\mathfrak{C}_j$ と $\mathfrak{C}-\{\mathfrak{C}_j\}$ の任意のカテゴリとの間が分離できなければ
できないほど大きい値をとる量 $f(y)$ の増分 $\triangle f(y)$ と y の増分
$\triangle y$ との比の, $\triangle y \to 0$ での極限値である (3.70)

ということになる。

式(3.70)での極限値は次の定理3.1で決定され、次の3事項が成立している：

(イ) $SM(\varphi, \omega_j) \to 0$ につれ、
$DOC(\varphi, \underline{\mathfrak{C}}[j]; SM) \to \infty.$

(ロ) $SM(\varphi, \omega_j) \to 1/2$ につれ、
$DOC(\varphi, \underline{\mathfrak{C}}[j]; SM) \to 0.$

(ハ) $SM(\varphi, \omega_j) \to 1$ につれ、
$DOC(\varphi, \underline{\mathfrak{C}}[j]; SM) \to -\infty.$ □

[定理3.1] **(2カテゴリ分類困難度 $DOC(\varphi, \underline{\mathfrak{C}}[j]; SM)$ の表現定理)**

$\forall \varphi \in \Phi[t_1, t_2]$,

$$DOC(\varphi, \underline{\mathfrak{C}}[j]; SM)$$
$$= -\log_e[SM(\varphi, \omega_j)/\{1-SM(\varphi, \omega_j)\}] \tag{3.71}$$
$$= \log_e[\{1-SM(\varphi, \omega_j)\}/SM(\varphi, \omega_j)]. \tag{3.72}$$

(証明) 微分公式

$$d/dp[-p \cdot \log_e p - (1-p) \cdot \log_e(1-p)]$$
$$= -\log_e p - 1 + \log_e(1-p) + 1$$
$$= \log_e\{(1-p)/p\}, \text{ where } 0 \leqq p \leqq 1 \tag{3.73}$$

が成り立つから、

$$DOC(\varphi, \underline{\mathfrak{C}}[j]; SM)$$
$$\equiv df(y)/dy \quad \because \text{ 式(3.69)}$$
$$= d/dy\ [-y \cdot \log_e y - (1-y) \cdot \log_e(1-y)]$$
$$\because \text{ 4式(3.46), (3.63), (3.64), (3.67)}$$
$$= \log_e\{(1-y)/y\} \tag{3.74}$$

を得、証明が終わったことがわかる。 □

実変数 x の関数

$$y=1/【1+\exp[-x]】 \tag{3.75}$$

は**シグモイド関数**(sigmoidal function)と称されるが、次の定理3.2は、変数 DOC(φ, $\underline{\mathfrak{C}}$[j]；SM)] のシグモイド関数として類似度 SM(φ, ω_j)が表現されることを指摘したものである。

［定理3.2］ **(類似度関数 SM(φ, ω_j)と2カテゴリ分類困難度 DOC(φ, $\underline{\mathfrak{C}}$[j]；SM)との間の関係定理)**

$$\forall \varphi \in \Phi[t_1, t_2],\ \ \forall j \in J,$$

$$SM(\varphi,\ \omega_j)=1/【1+\exp[z]】 \tag{3.76}$$

の解 z は、

$$z=DOC(\varphi, \underline{\mathfrak{C}}[j]; SM). \tag{3.77}$$

である。式(3.76)から、式(3.75)の成立もいえ、結局、

$$SM(\varphi,\ \omega_j)=1/【1+\exp[DOC(\varphi, \underline{\mathfrak{C}}[j]; SM)]】 \tag{3.78}$$

が成り立つ。

(証明)　変換公式

$$y=1/【1+\exp[z]$$

$$\Leftrightarrow y\cdot\exp[z]=1-y$$

$$\Leftrightarrow \exp[z]=(1-y)/y$$

$$\Leftrightarrow z=\log_e\{(1-y)/y\} \tag{3.79}$$

が成り立つから、式(3.67)の如く、変数 y の値を設定すれば、式(3.72)より、本定理3.2が成立することがわかる。 □

式(3.78)から(イ)，(ロ)，(ハ)の逆(イ′)，(ロ′)，(ハ′)が成立する：

(イ′)DOC(φ, $\underline{\mathfrak{C}}$[j]；SM)→∞ につれ、

SM(φ, ω_j)→0.

(ロ′)DOC(φ, $\underline{\mathfrak{C}}$[j]；SM)→0 につれ、

SM(φ, ω_j)→1/2.

(ハ′)DOC(φ, $\underline{\mathfrak{C}}$[j]；SM)→−∞ につれ、

SM(φ, ω_j)→1. □

更に、変換公式

$$1/\{1+\exp(a)\} \leqq 1/\{1+\exp(b)\}$$

$$\Leftrightarrow \{1+\exp(b)\} \leqq \{1+\exp(a)\}$$

$$\Leftrightarrow a \geqq b \tag{3.80}$$

を適用すれば、式(3.78)より、不等式

$$SM(\varphi,\ \omega_j) \leqq SM(\eta,\ \omega_j)$$

$$\Leftrightarrow DOC(\varphi, \underline{\mathfrak{C}}[j]; SM)] \geqq DOC(\eta, \underline{\mathfrak{C}}[j]; SM)] \tag{3.81}$$

が成立し、正に、式(3.72)の DOC(φ, $\underline{\mathfrak{C}}$[j]；SM)］がbinary recognitionにおける困難度であることを意味付けている。

次の補助定理3.4は、カテゴリ総数｜J｜が 2 以上であるという設定

$$|J| \geqq 2 \tag{3.82}$$

から当然、成り立つ。

［補助定理3.4］（**類似度関数 SM の唯一 1/2 大存在定理**）

不等式

$$SM(\varphi, \omega_j) > 1/2 \tag{3.83}$$

を満たすカテゴリ番号 $j \in J$ が存在するとすれば、ただ1つしかなくて、このとき、

$$\forall i \in J - \{j\},\ SM(\varphi, \omega_i) < 1/2] \tag{3.84}$$

が成り立つ。

（証明）　先ず、$j \neq i$ として、2つの不等式

$$SM(\varphi, \omega_j) > 1/2 \tag{3.85}$$

$$SM(\varphi, \omega_i) > 1/2 \tag{3.86}$$

が成立するとすれば、

$$\begin{aligned} 1 &= \sum_{j \in J} SM(\varphi, \omega_j) \quad \because \text{ axiom 2の(ii)} \\ &= SM(\varphi, \omega_j) + SM(\varphi, \omega_i) \\ &+ \sum_{k \in J - \{i, j\}} SM(\varphi, \omega_k) \\ &> \sum_{k \in J - \{i, j\}} SM(\varphi, \omega_k) \quad \because \text{ 式(3.83)} \\ &> 1 \end{aligned} \tag{3.87}$$

を得、これは矛盾である。よって、不等式(3.80)を満たすカテゴリ番号 $j \in J$ が存在するとすれば、ただ1つしかない。次に、式(3.84)の成立を示そう。

$$\begin{aligned} 1 &= \sum_{j \in J} SM(\varphi, \omega_j) \quad \because \text{ axiom 2の(ii)} \\ &= SM(\varphi, \omega_j) \\ &+ \sum_{k \in J - \{i\}} SM(\varphi, \omega_k) \\ &> 1/2 + \sum_{k \in J - \{i\}} SM(\varphi, \omega_k) \quad \because \text{ 式(3.80)} \\ \therefore\ & 1/2 > \sum_{k \in J - \{i\}} SM(\varphi, \omega_k) \geqq SM(\varphi, \omega_i) \\ & \text{for any } i \in J - \{j\} \end{aligned} \tag{3.88}$$

も得、式(3.81)が成立する。 □

次の定理3.3は、上述の補助定理3.4を適用して証明されるものであり、2カテゴリ分類困難度 DOC の効用の1つを明らかにしている。

［定理3.3］（**2カテゴリ分類困難度 DOC の唯一負存在定理**）

固定した $\varphi \in \Phi[t_1, t_2]$ について、不等式

$$DOC(\varphi, \underline{\mathfrak{C}}[j]; SM) < 0 \tag{3.89}$$

を満たすカテゴリ番号 $j \in J$ が存在するとすれば、ただ1つしかなくて、このとき、

$$\forall i \in J - \{j\}, DOC(\varphi, \underline{\mathfrak{C}}[i]; SM) > 0 \tag{3.90}$$

が成り立つ。

（証明）　一先ず、

$$0 < SM(\varphi, \omega_j) \leqq 1 \tag{3.91}$$

を仮定すれば、

式(3.89)の成立

$$\Leftrightarrow \log_e[\{1 - SM(\varphi, \omega_j)\} / SM(\varphi, \omega_j)] < 0$$

$$\because \text{ 式(3.72)}$$

$\Leftrightarrow \{1-\mathrm{SM}(\varphi,\ \omega_j)\}/\mathrm{SM}(\varphi,\ \omega_j)<1$

$\Leftrightarrow 1-\mathrm{SM}(\varphi,\ \omega_j)<\mathrm{SM}(\varphi,\ \omega_j)$

$\Leftrightarrow 1/2<\mathrm{SM}(\varphi,\ \omega_j)$ (3.92)

を得、よって、補助定理3.4を適用すれば、不等式(3.89)を満たすカテゴリ番号 $j \in J$ が存在するとすれば、ただ1つしかない。不等式(3.92)に補助定理3.4を適用すれば

$\forall i \in J-\{j\},\ 1/2>\mathrm{SM}(\varphi,\ \omega_i)$ (3.93)

を得、よって、不等式(3.91)を仮定して、

式(3.93)の成立

$\Leftrightarrow \forall i \in J-\{j\},\ 1-\mathrm{SM}(\varphi,\ \omega_j)>\mathrm{SM}(\varphi,\ \omega_i)$

$\Leftrightarrow \forall i \in J-\{j\},\ \{1-\mathrm{SM}(\varphi,\ \omega_j)\}/\mathrm{SM}(\varphi,\ \omega_i)>1$

$\Leftrightarrow \forall i \in J-\{j\},$

$\log_e[\{1-\mathrm{SM}(\varphi,\ \omega_j)\}/\mathrm{SM}(\varphi,\ \omega_j)]>0$

$\Leftrightarrow$式(3.90)の成立 $\because$ 式(3.72) (3.94)

が得られ、証明が終わる。 □

次の定理3.4は、不等式(3.95)の条件の下で、定理3.1で求められた2カテゴリ分類困難度 $\mathrm{DOC}(\varphi, \underline{\mathfrak{C}}[j]\ ;\mathrm{SM})$ の下界、上界を求めたものである。

［定理3.4］(**2カテゴリ分類困難度 DOC の下界・上界定理**)

$\forall \varphi \in \Phi[t_1, t_2],$

$0<\mathrm{SM}(\varphi,\ \omega_j)<1/2$ (3.95)

であるカテゴリ番号 $j \in J$ について、

$\mathrm{DOC}(\varphi, \underline{\mathfrak{C}}[j]\ ;\mathrm{SM})>0$ (3.96)

が成立し、このとき、

$[1-2\cdot\mathrm{SM}(\varphi,\ \omega_j)]/[1-\mathrm{SM}(\varphi,\ \omega_j)]$

$\leqq \mathrm{DOC}(\varphi, \underline{\mathfrak{C}}[j]\ ;\mathrm{SM})$ (3.97)

$\leqq [1-2\cdot\mathrm{SM}(\varphi,\ \omega_j)]/\mathrm{SM}(\varphi,\ \omega_j)$ (3.98)

(証明) 先ず、

$1/2<1-\mathrm{SM}(\varphi,\ \omega_j)<1$

$\Leftrightarrow$不等式(3.95)の成立

$\Leftrightarrow 0<1-2\cdot\mathrm{SM}(\varphi,\ \omega_j)<1$ (3.99)

に注意すれば、不等式

$\{1-\mathrm{SM}(\varphi,\ \omega_j)\}/\mathrm{SM}(\varphi,\ \omega_j)>1$

が成立し、よって、式(3.72)から、不等(3.96)が成立する。

任意の $x>0$ について、不等式［A4］

$\log_e x \leqq x-1$(等号は $x=1$ の時に限る) (3.100)

が成立することを適用すれば、

$\mathrm{DOC}(\varphi, \underline{\mathfrak{C}}[j]\ ;\mathrm{SM})$

$=\log_e[\{1-\mathrm{SM}(\varphi,\ \omega_j)\}/\mathrm{SM}(\varphi,\ \omega_j)]$

$\because$ 式(3.72)

$\leqq \{1-\mathrm{SM}(\varphi,\ \omega_j)\}/\mathrm{SM}(\varphi,\ \omega_j)-1$

$\because$ 式(3.99)

$\leqq \{1-2\cdot SM(\varphi,\ \omega_j)\}/SM(\varphi,\ \omega_j)$

を得、式(3.98)が示された。また、

任意の $x>0$ について、不等式

$$-\log_e x \geqq 1-x \text{(等号は } x=1 \text{ の時に限る)} \quad (3.101)$$

を適用すれば、

$DOC(\varphi, \underline{\mathfrak{C}}[j]\ ;\ SM)$

$= -\mathrm{loge}\ [SM(\varphi,\ \omega_j)/\{1-SM(\varphi,\ \omega_j)\}]$

∵ 式(3.71)

$\geqq 1-SM(\varphi,\ \omega_j)/[1-SM(\varphi,\ \omega_j)]$

∵ 式(3.98)

$= [1-2\cdot SM(\varphi,\ \omega_j)]/\{1-SM(\varphi,\ \omega_j)\}$

を得、式(3.97)が示された。 □

次の定理3.5は、$\exp[DOC(\varphi, \underline{\mathfrak{C}}[j]\ ;\ SM)]$ の下界を評価したものである。

［定理3.5］ **(2カテゴリ分類困難度 DOC の指数関数の下界定理)**

不等式

$\forall \varphi \in \Phi[t_1, t_2],\ \ \forall j \in J,$

$[1-SM(\varphi,\ \omega_j)]^2 \leqq [\mathrm{logeSM}(\varphi,\ \omega_j)]^{-2}$

$$\leqq \exp[DOC(\varphi, \underline{\mathfrak{C}}[j]\ ;\ SM)]. \quad (3.102)$$

が成り立つ。

（証明） 不等式［A4］

$1-x \leqq -\log_e x \leqq [(1-x)/x]^{1/2}$

(等号は $x=1$ のときに限る) for any $(0<x\leqq 1)$ (3.103)

を適用すれば、

$\forall \varphi \in \Phi[t_1, t_2],\ \ \forall j \in J,$

$1-SM(\varphi,\ \omega_j) \leqq -\log_e SM(\varphi,\ \omega_j)$

$\leqq\ [(1-SM(\varphi,\ \omega_j))/SM(\varphi,\ \omega_j)]^{1/2}$

for $0<SM(\varphi,\ \omega_j)\leqq 1$ (3.104)

が成り立つ。$SM(\varphi,\ \omega_j)=0$ のときも成り立つと見てよい。不等式(3.104)の自然対数をとれば、

$\log_e[1-SM(\varphi,\ \omega_j)] \leqq \log_e \log_e SM(\varphi,\ \omega_j)^{-1}$

$\leqq (1/2)\cdot\log_e[(1-SM(\varphi,\ \omega_j))/SM(\varphi,\ \omega_j)]$

$=(1/2)\cdot DOC(\varphi, \underline{\mathfrak{C}}[j]\ ;\ SM)$

を得るが、2 をかけ、exp をとれば、不等式(3.99)が成り立つ。 □

次の定理3.6は、式(3.46)の曖昧度 $H(\underline{\mathfrak{C}}[j]/T\varphi)$ の下界・上界を求めたものである。

［定理3.6］ **(曖昧度 $H(\underline{\mathfrak{C}}[j]/T\varphi)$ の下界・上界定理)**

$\forall \varphi \in \Phi[t_1, t_2],\ \ \forall j \in J,$

$\min\{SM(\varphi,\ \omega_j),\ 1-SM(\varphi,\ \omega_j)\} \leqq H(\underline{\mathfrak{C}}[j]/T\varphi)$

$$\leqq 2\cdot[SM(\varphi,\ \omega_j)\cdot\{1-SM(\varphi,\ \omega_j)\}]^{1/2} \quad (3.105)$$

が成立ち、等号は $SM(\varphi,\ \omega_j)\in\{0, 1\}$ のときに限る。

（証明） 不等式［A4］

$\min\{p, 1-p\}$

$$\leqq -p\cdot\log_e p-(1-p)\cdot\log_e(1-p)$$
$$\leqq 2\cdot[p(1-p)] \text{（等号は } p=1 \text{ の時に限る）} \tag{3.106}$$
において、$p=SM(\varphi,\ \omega_j)$とおき、3式(3.46)，(3.63)，(3.64)を考慮すればよい。 □

その値が1より大きくない非負実数値変数 y を
$$y\equiv SM(\varphi,\ \omega_j) \tag{3.107}$$
とおき、関数
$$f(y)=-y\cdot\log_e y-(1-y)\cdot\log_e(1-y) \tag{3.108}$$
を定義すれば、
$$H(\underline{\mathfrak{C}}[j]/T\varphi)=f(y) \tag{3.109}$$
である。次の定理3.7は、類似度 $SM(\varphi,\ \omega_j)$に関する式(3.46)の $H(\underline{\mathfrak{C}}[j]/T\varphi)$の2次微係数を求めたものである。

［定理3.7］（**曖昧度 $H(\underline{\mathfrak{C}}[j]/T\varphi)$の2次微係数定理**）
$$d^2H(\underline{\mathfrak{C}}[j]/T\varphi)/dSM(\varphi,\ \omega_j)^2$$
$$=-1/SM(\varphi,\ \omega_j)-1/[1-SM(\varphi,\ \omega_j)]$$
$$\leqq \min\{-1/SM(\varphi,\ \omega_j),\ -1/(1-SM(\varphi,\ \omega_j))\} \tag{3.110}$$
$$-SM(\varphi,\ \omega_j)\cdot d^2H(\underline{\mathfrak{C}}[j]/T\varphi)/dSM(\varphi,\ \omega_j)^2$$
$$=1+\exp[-DOC(\varphi,\underline{\mathfrak{C}}[j];SM)] \tag{3.111}$$
が成立する。

（証明） $df(y)/dy$
$$=\log_e[(1-y)/y] \quad \because \text{ 式(3.70)}$$
$$=\log_e(1-y)-\log_e y \tag{3.112}$$
であるから、直ちに、等式，不等式
$$d^2f(y)/dy^2$$
$$=-1/y-1/[1-y] \tag{3.113}$$
$$\leqq \min\{-1/y,\ -1/(1-y)\} \tag{3.114}$$
が得られ、式(3.107)を勘案すれば、式(3.110)の証明が終わる。不等式(3.114)は、等式(3.113)において、
$$-1/y,\ -1/\ [1-y]\ \leqq 0$$
を考慮したものである。

y を式(3.113)の両辺にかければ、
$$y\cdot d^2f(y)/dy^2$$
$$=-1-y/[1-y]$$
$$\therefore \quad -y\cdot d^2f(y)/dy^2$$
$$=1+y/[1-y] \tag{3.115}$$
が得られる。ここで、等式
$$x=\exp[\log_e x]\ (0<x\leqq 1) \tag{3.116}$$
を考慮すれば、
$$y/[1-y]=\exp[\log_e\{y/(1-y)\}]$$
$$=\exp[-[-\log_e\{y/(1-y)\}]] \tag{3.117}$$
であるから、4式(3.107)，(3.115)，(3.117)，(3.71)を考慮すれば、式(3.111)が成立することになる。 □

4. 2カテゴリ分類容易度 EOC(φ, $\underline{\mathfrak{C}}$[j]；SM)と、認識手続き煩雑さ COR(φ, $\underline{\mathfrak{C}}$[j]；SM)

本章では、2カテゴリ分類容易度 EOC(φ, $\underline{\mathfrak{C}}$[j]；SM)を定義し、3章に引き続いて、認識手続き煩雑さ COR(φ, $\underline{\mathfrak{C}}$[j]；SM)などに関連した解析を行う。

4.1 2カテゴリ分類容易度 EOC(φ, $\underline{\mathfrak{C}}$[j]；SM)

2式(3.71), (3.72)の2カテゴリ分類困難度 DOC(φ, $\underline{\mathfrak{C}}$[j]；SM)の負として、2カテゴリ分類容易度(easiness of binary classification)EOC(φ, $\underline{\mathfrak{C}}$[j]；SM)を、

$$\begin{aligned}&\mathrm{EOC}(\varphi, \underline{\mathfrak{C}}[j];\mathrm{SM})\\ &\equiv -\mathrm{DOC}(\varphi, \underline{\mathfrak{C}}[j];\mathrm{SM})\end{aligned} \tag{4.1}$$

と定義すれば、式(3.71)から、表現

$$\begin{aligned}&\mathrm{EOC}(\varphi, \underline{\mathfrak{C}}[j];\mathrm{SM})\\ &=\log_e[\mathrm{SM}(\varphi,\ \omega_j)/\{1-\mathrm{SM}(\varphi,\ \omega_j)\}]\end{aligned} \tag{4.2}$$

が成り立つ。明らかに、0になる等式

$$\begin{aligned}&\forall\varphi\in\Phi[t_1, t_2],\ \ \forall j\in J,\\ &\mathrm{DOC}(\varphi, \underline{\mathfrak{C}}[j];\mathrm{SM})+\mathrm{EOC}(\varphi, \underline{\mathfrak{C}}[j];\mathrm{SM})=0\end{aligned} \tag{4.3}$$

が成り立つ。

次の定理4.1は、axiom 2の(ii)の別表現式(4.4)が得られ、2カテゴリ分類困難度 DOC(φ, $\underline{\mathfrak{C}}$[j]；SM)を変数とするシグモイド関数による類似度 SM(φ, ωj)の表現式(3.76), (3.77)に対し、2カテゴリ分類容易度 EOC(φ, $\underline{\mathfrak{C}}$[j]；SM)を変数とするシグモイド関数による類似度の、カテゴリ番号 i∈J−{j} にわたる総和 $\sum_{i\in J-\{j\}}$ SM(φ, ω_i)の表現式(4.5)が成り立つことを指摘している。

［定理4.1］(**2カテゴリ分類困難・容易度の総和1のシグモイド関数の両和定理**)

等式

$$\begin{aligned}&\forall\varphi\in\Phi[t_1, t_2],\ \ \forall j\in J,\\ &1/【1+\exp[\mathrm{DOC}(\varphi, \underline{\mathfrak{C}}[j];\mathrm{SM})]】\\ &+1/【1+\exp[\mathrm{EOC}(\varphi, \underline{\mathfrak{C}}[j];\mathrm{SM})]】=1\end{aligned} \tag{4.4}$$

が成立し、よって、

$$\begin{aligned}&\sum_{i\in J-\{j\}}\mathrm{SM}(\varphi,\ \omega_i)=1-\mathrm{SM}(\varphi,\ \omega_j)\\ &=1/【1+\exp[\mathrm{EOC}(\varphi, \underline{\mathfrak{C}}[j];\mathrm{SM})]】\end{aligned} \tag{4.5}$$

が成り立つ。

(証明) 1−1/[1+exp(−z)] を計算すれば、

$$\begin{aligned}&1-1/[1+\exp(-z)]\\ &=\exp(-z)/[1+\exp(-z)]\\ &=1/[1+\exp(z)]\end{aligned}$$

を得、その和が 1 になる等式

$$\begin{aligned}&\therefore\ \ 1/[1+\exp(z)]+1/[1+\exp(-z)]=1\\ &\qquad\text{for any }-\infty<z<+\infty\end{aligned} \tag{4.6}$$

が成立する。よって、2式(3.77), (4.1)を考慮すれば、式(4.4)が成り立つ。

更に、

$\sum_{i \in J-\{j\}}$ SM(φ, ω_i)
$=1-$SM(φ, ω_j) ∵ axiom 2の(ii)
$=1-1/$【1+exp［DOC$(\varphi, \underline{\mathfrak{C}}[j]$；SM)］】
∵ 2式(3.76), (3.77)
$=1/$【1+exp［EOC$(\varphi, \underline{\mathfrak{C}}[j]$；SM)］】
∵ 式(4.4)

を得、式(4.5)が成立する。 □

次の定理4.2は、【1/{1+exp(DOC$(\varphi, \underline{\mathfrak{C}}[j]$；SM))}】,【1/{1+exp(EOC$(\varphi, \underline{\mathfrak{C}}[j]$；SM))}】の和式(4.4)に対応して、その積式(4.7)が類似度 SM の、2カテゴリ分類容易度 EOC$(\varphi, \underline{\mathfrak{C}}[j]$；SM)に関する微係数になることを指摘している。

［定理4.2］(**類似度関数 SM の2カテゴリ分類容易度 EOC に関する微係数定理**)

$\forall \varphi \in \Phi[t_1, t_2]$, $\forall j \in J$,
dSM(φ, ω_j)/dEOC$(\varphi, \underline{\mathfrak{C}}[j]$；SM)
$=$【1/{1+exp(DOC$(\varphi, \underline{\mathfrak{C}}[j]$；SM))}】.
【1/{1+exp(EOC$(\varphi, \underline{\mathfrak{C}}[j]$；SM))}】. (4.7)

(証明) w$=$EOC$(\varphi, \underline{\mathfrak{C}}[j]$；SM) (4.8)

として、

dSM(φ, ω_j)/dEOC$(\varphi, \underline{\mathfrak{C}}[j]$；SM)
$=$d［1/{1+exp(−w)}］/dw
∵ 3式(3.76), (3.77), (4.1)
$=(-1)\cdot\{1+\exp(-w)\}^{-2}\cdot\exp(-w)\cdot(-1)$
$=[1/\{1+\exp(-w)\}]\cdot[\exp(-w)/\{1+\exp(-w)\}]$
$=[1/\{1+\exp(-w)\}]\cdot[1/\{1+\exp(w)\}]$ (4.9)

を得、式(4.1)を考慮すれば証明されたことがわかる。 □

次の定理4.3は、3式(3.17), (3.52), (3.53)の平均相互情報量 AMI$(\underline{\mathfrak{C}}[j]$；$\Phi[t_1, t_2])$を類似度 SM$(\varphi, \omega_j)$に関し偏微分した結果を示し、式(4.1)の2カテゴリ分類容易度 EOC$(\varphi, \underline{\mathfrak{C}}[j]$；SM)が線形的に関係しており、EOC の定数倍に敏感であることがわかる。

［定理4.3］(**平均相互情報量 AMI の、類似度 SM に関する偏微分定理**)

∂AMI$(\underline{\mathfrak{C}}[j]$；$\Phi[t_1, t_2])/\partial$SM$(\varphi, \omega_j)$
$=$q$(T\varphi)\cdot$【EOC$(\varphi, \underline{\mathfrak{C}}[j]$；SM)
$-\log_e[p(\mathfrak{C}_1[j])/\{1-p(\mathfrak{C}_1[j])\}]$】. (4.10)

(証明) 式(3.52)の AMI$(\underline{\mathfrak{C}}[j]$；$\Phi[t_1, t_2])$は、

AMI$(\underline{\mathfrak{C}}[j]$；$\Phi[t_1, t_2])$
$=-[\sum_{\varphi \in \Phi[t_1, t_2]} q(T\varphi)\cdot SM(\varphi, \omega_j)]\cdot\log_e$
$[\sum_{\varphi \in \Phi[t_1, t_2]} q(T\varphi)\cdot SM(\varphi, \omega_j)]$
$-[\sum_{\varphi \in \Phi[t_1, t_2]} q(T\varphi)\cdot\{1-SM(\varphi, \omega_j)\}]\cdot\log_e$
$[\sum_{\varphi \in \Phi[t_1, t_2]} q(T\varphi)\cdot\{1-SM(\varphi, \omega_j)\}]$
$-\sum_{\varphi \in \Phi[t_1, t_2]} q(T\varphi)\cdot$
$[-SM(\varphi, \omega_j)\cdot\log_e SM(\varphi, \omega_j)$
$-\{1-SM(\varphi, \omega_j)\}\cdot\log_e\{1-SM(\varphi, \omega_j)\}]$

$\because$　5式(3.15), (3.63), (3.64), (3.27), (3.46)　(4.11)

と表現されるから、

$$\begin{aligned}
&\partial\,\mathrm{AMI}(\underline{\mathfrak{C}}\,[j]\,;\,\Phi[t_1,t_2])/\partial\,\mathrm{SM}(\varphi,\ \omega_j)\\
&=-q(T\varphi)\cdot\log_e\\
&\quad[\Sigma_{\varphi\in\Phi[t_1,t_2]}\ q(T\varphi)\cdot\mathrm{SM}(\varphi,\ \omega_j)]\\
&-q(T\varphi)\\
&+q(T\varphi)\cdot\log_e\\
&\quad[\Sigma_{\varphi\in\Phi[t_1,t_2]}\ q(T\varphi)\cdot\{1-\mathrm{SM}(\varphi,\ \omega_j)\}]\\
&+q(T\varphi)\\
&+q(T\varphi)\cdot\log_e\mathrm{SM}(\varphi,\ \omega_j)+q(T\varphi)\\
&-q(T\varphi)\cdot\log_e\{1-\mathrm{SM}(\varphi,\ \omega_j)\}-q(T\varphi)
\end{aligned}\tag{4.12}$$

$$\begin{aligned}
&=q(T\varphi)\cdot\log_e[\mathrm{SM}(\varphi,\ \omega_j)/\{1-\mathrm{SM}(\varphi,\ \omega_j)\}]\\
&-q(T\varphi)\cdot\\
&\log_e【[\Sigma_{\varphi\in\Phi[t_1,t_2]}\ q(T\varphi)\cdot\mathrm{SM}(\varphi,\ \omega_j)]\\
&\qquad/\ [\Sigma_{\varphi\in\Phi[t_1,t_2]}\ q(T\varphi)\cdot\{1-\mathrm{SM}(\varphi,\ \omega_j)\}]】
\end{aligned}\tag{4.13}$$

$$\begin{aligned}
&=q(T\varphi)\cdot【\mathrm{EOC}(\varphi,\underline{\mathfrak{C}}\,[j]\,;\,\mathrm{SM})-\log_e\ [p(\mathfrak{C}_1[j])\\
&\quad/\ \{1-p(\mathfrak{C}_1[j])\}]】
\end{aligned}$$

$\because$　5式(3.71), (4.1), (3.15), (3.63), (3.12)

を得、証明が終わる。 □

次の定理4.4は、平均相互情報量 $\mathrm{AMI}(\underline{\mathfrak{C}}\,[j]\,;\,\Phi[t_1,t_2])$の、類似度 $\mathrm{SM}(\varphi,\ \omega_j)$に関する2次微分を求めたものである。

［定理4.4］（**平均相互情報量 AMI の、類似度 SM に関する2次偏微分定理**）

$$\begin{aligned}
&\partial^2\mathrm{AMI}(\underline{\mathfrak{C}}\,[j]\,;\,\Phi[t_1,t_2])/\partial\,\mathrm{SM}(\varphi,\ \omega_j)^2\\
&=q(T\varphi)\cdot【1/\mathrm{SM}(\varphi,\ \omega_j)+1/\{1-\mathrm{SM}(\varphi,\ \omega_j)\}\\
&-q(T\varphi)\cdot[1/\{\Sigma_{\varphi\in\Phi[t_1,t_2]}\ q(T\varphi)\cdot\mathrm{SM}(\varphi,\ \omega_j)\}\\
&+1/\{\Sigma_{\varphi\in\Phi[t_1,t_2]}\ q(T\varphi)\cdot\{1-\mathrm{SM}(\varphi,\ \omega_j)\}\}]】
\end{aligned}\tag{4.14}$$

（証明）　式(4.12)に、$\partial/\partial\,\mathrm{SM}(\varphi,\ \omega_j)$を作用させれば、

$$\begin{aligned}
&\partial^2\mathrm{AMI}(\underline{\mathfrak{C}}\,[j]\,;\,\Phi[t_1,t_2])/\partial\,\mathrm{SM}(\varphi,\ \omega_j)^2\\
&=-q(T\varphi)^2/[\Sigma_{\varphi\in\Phi[t_1,t_2]}\ q(T\varphi)\cdot\mathrm{SM}(\varphi,\ \omega_j)]\\
&-q(T\varphi)^2/[\Sigma_{\varphi\in\Phi[t_1,t_2]}\ q(T\varphi)\cdot\\
&\qquad\qquad\{1-\mathrm{SM}(\varphi,\ \omega_j)\}]\\
&+q(T\varphi)/\mathrm{SM}(\varphi,\ \omega_j)+q(T\varphi)/[1-\mathrm{SM}(\varphi,\ \omega_j)]
\end{aligned}\tag{4.15}$$

$$\begin{aligned}
&=q(T\varphi)\cdot【1/\mathrm{SM}(\varphi,\ \omega_j)+1/\{1-\mathrm{SM}(\varphi,\ \omega_j)\}\\
&-q(T\varphi)\cdot[1/\{\Sigma_{\varphi\in\Phi[t_1,t_2]}\ q(T\varphi)\cdot\mathrm{SM}(\varphi,\ \omega_j)\}\\
&+1/\{\Sigma_{\varphi\in\Phi[t_1,t_2]}\ q(T\varphi)\cdot\\
&\{1-\mathrm{SM}(\varphi,\ \omega_j)\}\}]】
\end{aligned}$$

を得、証明が終わる。 □

4.2　最大類似度認識法における誤認識確率 error-prob$\{\varphi,\ \mathfrak{C}_j\}$の下界・上界

処理の対象とする問題のパターン $\varphi\in\Phi$ の帰属するカテゴリについての、 | J | 個の仮説

Pattern φ belongs to the jth category $\mathfrak{C}_k$

$$, k=1\sim|J| \tag{4.16}$$

の内、どの1つが真であるかを推定する方法の1つとして、**最尤認識法**(recognition-method of maximum likelihood)がある。確率条件

$$\forall \varphi \in \varphi, [\forall k \in J, 0 \leqq p(\mathfrak{C}_k/T\varphi) \leqq 1]$$
$$\wedge \sum_{k \in J} p(\mathfrak{C}_k/T\varphi)=1 \tag{4.17}$$

を満たすような“$\varphi \in \Phi$に対応し、パターンモデル $T\varphi \in \Phi$が確保された条件の下で、第 $k \in J$ のカテゴリ $\mathfrak{C}_k$ が出現する条件つき**事後確率**(aposteriori conditional probability) $p(\mathfrak{C}_k/T\varphi)$” の系

$$p(\mathfrak{C}_k/T\varphi), \quad k \in J \tag{4.18}$$

を導入すると、最尤認識法とは、パターンモデル $T\varphi \in \Phi$ が観測されたとき、事後確率の最も大きい仮説が真であると採用する認識法であり、

$$\max_{k \in J} p(\mathfrak{C}_k/T\varphi)=p(\mathfrak{C}_j/T\varphi) \tag{4.19}$$
$$\Rightarrow$$

Pattern φ belongs to the jth category $\mathfrak{C}_j$ (4.20)

と記述される。

パターンモデル $T\varphi \in \Phi$が第 $k \in J$ のカテゴリ $\mathfrak{C}_k$ に帰属する事後確率 $p(\mathfrak{C}_k/T\varphi)$として、2式(3.9), (3.63)の設定から、

$$p(\mathfrak{C}_k/T\varphi)=SM(\varphi, \omega_k), \quad k \in J \tag{4.21}$$

を採用している。よって、この場合、最尤推定法は、下記の**最大類似度法**といわれるものになる。

処理の対象とする問題のパターン$\varphi \in \Phi$の帰属するカテゴリとしては、類似度が最も大になるカテゴリが尤もらしいという観点を採用する単段階認識法とは、

$$\max_{k \in J} SM(\varphi, \omega_k)=SM(\varphi, \omega_j) \tag{4.22}$$
$$\Rightarrow$$

Pattern φ belongs to the jth category $\mathfrak{C}_j$ (4.23)

と記述される最大類似度認識法(recognition-method of maximum similarity-measure)であるが、この認識法での**誤認識確率**(error-probability)

$$\text{error-prob}\{\varphi, \mathfrak{C}_j\}$$
$$\equiv 1-SM(\varphi, \omega_j) \tag{4.24}$$
$$=1-\max_{k \in J} SM(\varphi, \omega_k) \tag{4.25}$$
$$\text{where } j=\text{argmax}_{i \in J}\, SM(\varphi, \omega_i) \tag{4.26}$$

を考えれば、

$$\text{error-prob}\{\varphi, \mathfrak{C}_j\}$$
$$=1/【1+\exp[EOC(\varphi, \underline{\mathfrak{C}}[j]; SM)]】 \tag{4.27}$$

が成り立つ。

次の定理4.5は、誤認識確率 error-prob$\{\varphi, \mathfrak{C}_j\}$を式(4.1)の2カテゴリ分類容易度 $EOC(\varphi, \underline{\mathfrak{C}}[j]; SM)$の指数関数を用いて、その下界・上界で評価したものである。

[定理4.5] (**誤認識確率 error-prob$\{\varphi, \mathfrak{C}_j\}$ の下界・上界定理**)

$$\forall \varphi \in \Phi[t_1, t_2], \quad \forall j \in J,$$
$$1-\exp[EOC(\varphi, \underline{\mathfrak{C}}[j]; SM)]$$

$\leqq$ error-prob $\{\varphi, \mathfrak{C}_j\}$ (4.28)

$\leqq 1-\exp[\mathrm{EOC}(\varphi, \underline{\mathfrak{C}}[j]; \mathrm{SM})]$

$+\exp[2\cdot \mathrm{EOC}(\varphi, \underline{\mathfrak{C}}[j]; \mathrm{SM})]$ (4.29)

（証明） 任意の実数 w について、2公式

$$1/[1+\exp(w)]-[1-\exp(w)]$$
$$=\exp(2w)/[1+\exp(w)]$$
$$=\exp(w)\cdot 【1/[1+\exp(-w)]】\geqq 0 \quad (4.30)$$
$$1-\exp(w)+\exp(2w)-1/[1+\exp(w)]$$
$$=\exp(3w)/[1+\exp(w)]$$
$$=\exp(2w)\cdot 【1/[1+\exp(-w)]】\geqq 0 \quad (4.31)$$

において、式(4.8)の如くおけば、error-prob $\{\varphi, \mathfrak{C}_j\}$ の表現式(4.27)から明らかである。 □

4.3 認識するときの手続きの煩雑さ COR $(\varphi, \underline{\mathfrak{C}}[j]; \mathrm{SM})$

4.3.1 カテゴリ帰属知識集合の包摂

認識システム RECOGNITRON がパターン φ に対し持つカテゴリ帰属知識(categorical-membership knowledge)を

$$\langle \varphi, \gamma \rangle \in \langle \Phi, 2^J \rangle \quad (4.32)$$

と表す。$\langle \varphi, \gamma \rangle$ はパターン $\varphi \in \Phi$ がカテゴリ $\mathfrak{C}_j$, $j \in J$ の何れか1つに帰属している可能性があることの表現である。ここに、Φ, 2^J は各々、処理の対象とする問題のパターン φ の集合、カテゴリ番号集合 J のすべての部分集合のなす集合であり、$\langle \Phi, 2^J \rangle$ はカテゴリ帰属知識空間と呼ばれ、その代数的･幾何学的･解析的諸性質が S.Suzuki によって既に解明されている [B3], [B4]。

カテゴリ帰属知識空間 $\langle \Phi, 2^J \rangle$ から $\langle \Phi, 2^J \rangle$ への写像(構造受精変換)

$$\mathrm{TA}(\mu)\mathrm{T} : \langle \Phi, 2^J \rangle \to \langle \Phi, 2^J \rangle \quad (4.33)$$

を導入すれば [B3], [B4]、カテゴリ帰属知識のなす2つの部分集合

$$\langle \Psi, \Gamma \rangle \equiv \{\langle \varphi, \gamma \rangle \mid \varphi \in \Psi, \gamma \in \Gamma\} \quad (4.34)$$
$$\langle \Psi', \Gamma' \rangle \equiv \{\langle \varphi', \gamma' \rangle \mid \varphi' \in \Psi, \gamma' \in \Gamma\} \quad (4.35)$$

について、

$$\exists \mu \in 2^J, \mathrm{TA}(\mu)\mathrm{T}\cdot \langle \Psi, \Gamma \rangle$$
$$\equiv \{\mathrm{TA}(\mu)\mathrm{T}\cdot \langle \varphi, \gamma \rangle \mid \langle \varphi, \gamma \rangle \in \langle \Psi, \Gamma \rangle\} \quad (4.36)$$
$$=\langle \Psi', \Gamma' \rangle \quad (4.37)$$

が成立するとき、$\langle \Psi, \Gamma \rangle$ は $\langle \Psi', \Gamma' \rangle$ を包摂する(subsume)といい、

$$\langle \Psi, \Gamma \rangle \to_{ss} \langle \Psi', \Gamma' \rangle \quad (4.38)$$

と表す。認識システム RECOGNITRON が処理の対象とする問題の入力パターン $\varphi \in \Phi$ に関し無知でカテゴリ帰属知識 $\langle \varphi, J \rangle$ しか持っていない場合、初期条件

$$\psi_0 \equiv \mathrm{T}\varphi, \ \lambda_0 \equiv J \quad (4.39)$$

を設定し、カテゴリ番号集合 μ_s の列

$$\mu_s \in 2^J, s=0, 1, 2, \cdots, t-1, t \quad (4.40)$$

を見つけ、多段階カテゴリ帰属知識変換

$$\langle \psi_s, \lambda_s \rangle = \mathrm{TA}(\mu_{s-1})\mathrm{T}\cdot \langle \psi_{s-1}, \lambda_{s-1} \rangle,$$
$$s=1, 2, \cdots, t \quad (4.41)$$

を実行し、第 t 認識段階で不動点方程式(終了基準)

$$\langle \psi_t,\ \lambda_t \rangle = \mathrm{TA}(\mu_t)\mathrm{T}\cdot\langle \psi_t,\ \lambda_t \rangle \tag{4.42}$$

が成立するような不動点認識計算で示される不動点連想形認識の働きが、多段階包摂過程

$$\exists j \in J,\ \langle \Phi, 2^J \rangle\ (\supset \{\langle \varphi, J \rangle\})\rightarrow ss\ \langle \Psi_1,\ \Gamma_1 \rangle$$
$$\rightarrow ss \langle \Psi_2,\ \Gamma_2 \rangle \rightarrow \cdots \rightarrow ss\ \{\langle \omega_j,\ [j] \rangle\} \subset \langle \Omega, J \rangle \tag{4.43}$$

を生成する場合があることを、文献［B4］の定理2.2の(ｉ)が保証している。この保証は、文献［B4］の定理3.4からわかるように、文献［B4］のA14.1節のSM－ミックスチュア条件と文献［B4］の4.2.2項のSM－直交条件とを満たす文献［B4］の式(A2.5)の類似度関数 SM によって得られる。

4.3.2 多段階パターン変換における認識手続き煩雑さ COR(φ, $\underline{\mathfrak{C}}$ [j]；SM)

S.Suzuki の提案した前項4.3.1の多段階想起形認識法では、処理の対象とする問題のパターン $\varphi \in \Phi$ を

$$\psi_0 \equiv \mathrm{T}\varphi \rightarrow \psi_1 \rightarrow \psi_2$$
$$\rightarrow \cdots \rightarrow \psi_t = \mathrm{T}\,\omega_j \tag{4.44}$$

と多段階パターン変換していって、最終的には、

$$\exists j \in J,\ \mathrm{SM}(\psi_t,\ \omega_j) = 1 \tag{4.45}$$

なるカテゴリ番号を発見して、

Pattern φ is associated with a recalled
pattern-model $\mathrm{T}\omega_j$ and gets classified into jth category $\mathfrak{C}_j$ (4.46)

と認識する。

このとき、

$$v_j \equiv v_j(\varphi) \equiv \mathrm{SM}(\varphi,\ \omega_j) \tag{4.47}$$

とおくと、2つの解釈

①$(+\infty \geqq) - \log_e(1 - v_j)\ (\geqq 0)$

：Tφ の内に $\mathrm{T}\omega_j$ が含まれている程度を表す情報量 (4.48)

②$(\log_e 2 \geqq)\,[-v_j \cdot \log_e v_j$

$\quad - (1 - v_j)\cdot\log_e(1 - v_j)]\ (\geqq 0)$

：Tφ が $\mathrm{T}\omega_j$ を「表しているか、表していないどうか」についての
不確定さの程度を表している情報量 (4.49)

が可能であり、**多段階想起形認識手続き煩雑さ** COR(φ, $\underline{\mathfrak{C}}$ [j]；SM) (complexity of multi-stage associative-recognition)が、

$$\mathrm{COR}(\varphi,\ \underline{\mathfrak{C}}\,[j]\ ;\ \mathrm{SM})$$
$$\equiv p(\mathfrak{C}_1[j]/\mathrm{T}\varphi)\cdot$$
$$\log_e [p(\mathfrak{C}_1[j]/\mathrm{T}\varphi)/\{1 - p(\mathfrak{C}_1[j]/\mathrm{T}\varphi)\}] \tag{4.50}$$
$$= \mathrm{SM}(\varphi,\ \omega_j)\cdot \mathrm{EOC}(\varphi\ ;\ \underline{\mathfrak{C}}\,[j]\ ;\ \mathrm{SM})$$
$$\because\ 3式(3.63),\ (4.1),\ (3.71) \tag{4.51}$$
$$= v_j \cdot \log_e \{v_j/(1 - v_j)\}$$
$$\because\ 式(4.47) \tag{4.52}$$
$$= -\log_e(1 - v_j)$$
$$-[-v_j\cdot\log_e v_j - (1 - v_j)\cdot\log_e(1 - v_j)]$$
$$= ① - ② \tag{4.53}$$

$= T\varphi$ の内に $T\omega_j$ が含まれている程度を表す情報量から、
$T\varphi$ の内に $T\omega_j$ が含まれているか、含まれていないかに
ついての不確定さを取り除いて得られた情報量 (4.54)

と導入できる。明らかに、

$$\mathrm{COR}(\varphi, \underline{\mathfrak{C}}[j]; \mathrm{SM}) = +\infty \quad \text{if} \quad v_j = 1$$
$$\mathrm{COR}(\varphi, \underline{\mathfrak{C}}[j]; \mathrm{SM}) = 0 \quad \text{if} \quad v_j = 1/2$$
$$\mathrm{COR}(\varphi, \underline{\mathfrak{C}}[j]; \mathrm{SM}) = 0 \quad \text{if} \quad v_j = 0 \tag{4.55}$$

であるし、

$$\mathrm{COR}(\varphi, \underline{\mathfrak{C}}[j]; \mathrm{SM}) \geqq 0 \quad \text{if} \quad 1/2 \leqq v_j < 1$$
$$\mathrm{COR}(\varphi, \underline{\mathfrak{C}}[j]; \mathrm{SM}) \leqq 0 \quad \text{if} \quad 0 < v_j \leqq 1/2 \tag{4.56}$$

である。

この $\mathrm{COR}(\varphi, \mathfrak{C}[j]; \mathrm{SM})$ については、

パターン φ が第 $j \in J$ 番目のカテゴリ $\mathfrak{C}_j$ に帰属すると多段階認識するときの手続きの煩雑さという解釈がなされ、この値 $\mathrm{COR}(\varphi, \underline{\mathfrak{C}}[j]; \mathrm{SM})$ が大きいければ大きいほど、パターン φ が第 $j \in J$ 番目のカテゴリ $\mathfrak{C}_j$ に帰属すると多段階認識するときの手続きの煩雑さは簡単である、すなわち、例えば、認識が終了するまでの、2式(4.41), (4.44)の認識段階数 t は小さいだろう。

4.3.3 2カテゴリ分類性能、認識性能

Ⅰ. パターン $\varphi \in \Phi[t_1, t_2]$ を固定した場合の、認識手続き煩雑さ

不等式

$$\mathrm{COR}(\varphi, \underline{\mathfrak{C}}[i]; \mathrm{SM})$$
$$\leqq \mathrm{COR}(\varphi, \underline{\mathfrak{C}}[j]; \mathrm{SM}) \tag{4.57}$$

が成立していれば、

パターン φ が第 $i \in J$ 番目のカテゴリ $\mathfrak{C}_i$ に帰属すると認識するときの手続きの煩雑さよりも、パターン φ が第 $j \in J$ 番目のカテゴリ $\mathfrak{C}_j$ に帰属すると認識するときの手続きの煩雑さの方が簡単である。

Ⅱ. $\Phi[t_1, t_2]$ を固定した場合の、2カテゴリ分類性能 POC

$\mathrm{COR}(\varphi, \underline{\mathfrak{C}}[j]; \mathrm{SM})$ の、2条件式(3.11), (3.12)を満たす出現確率 $q(T\varphi)$ によるその平均値

$$\mathrm{POC}(\underline{\mathfrak{C}}[j], \Phi[t_1, t_2]; \mathrm{SM})$$
$$\equiv \Sigma_{\varphi \in \Phi[t_1, t_2]}\, q(T\varphi) \cdot \mathrm{COR}(\varphi, \underline{\mathfrak{C}}[j]; \mathrm{SM})$$
$$= \Sigma_{\varphi \in \Phi[t_1, t_2]}\, q(T\varphi) \cdot p(\mathfrak{C}_1[j]/T\varphi) \cdot$$
$$\mathrm{EOC}(\varphi; \underline{\mathfrak{C}}[j]; \mathrm{SM})$$
$$\because \quad 3式(3.9), (3.63), (4.50) \tag{4.58}$$

の値が大きいければ大きいほど、パターン φ が第 $j \in J$ 番目のカテゴリ $\mathfrak{C}_j$ に帰属すると認識するときの手続きの煩雑さは簡単であるようなパターン φ から $\Phi[t_1, t_2]$ はなっている。

式(4.58)の $\mathrm{POC}(\underline{\mathfrak{C}}[j], \Phi[t_1, t_2]; \mathrm{SM})$ は類似度関数 SM を採用している認識システム RECOGNITRON の、パターン集合 $\Phi[t_1, t_2]$ についての認識性能(performance of classification)の指標であろう。

Ⅲ.パターン集合 $\Phi[t_1, t_2]$ を変えた場合の、認識手続き煩雑さ

2つのパターン集合 $\Phi_1[t_1, t_2]$, $\Phi_2[t_1, t_2]$ について、不等式

$$\mathrm{POC}(\underline{\mathfrak{C}}[j], \Phi_1[t_1, t_2]; \mathrm{SM})$$

$\leqq$POC($\underline{\mathfrak{C}}$[j], Φ_2[t_1, t_2]；SM) (4.59)

が成立していれば、

Φ_2[t_1, t_2] は Φ_1[t_1, t_2] よりも、第 j∈J 番目のカテゴリ $\mathfrak{C}_j$ に帰属すると認識するときの手続きの煩雑さは簡単である。

Ⅳ. 類似度関数 SM を変えた場合の、認識性能

axiom 2を満たす2つの類似度関数 SM_1, SM_2 について、不等式

POC($\underline{\mathfrak{C}}$[j], Φ[t_1, t_2]；SM_1)

$\leqq$POC($\underline{\mathfrak{C}}$[j], Φ[t_1, t_2]；SM_2) (4.60)

が成立していれば、

SM_2 を採用している認識システム RECOGNITRON(2)は SM_1 を採用している認識システム RECOGNITRON(1)よりも、第 j∈J 番目のカテゴリ $\mathfrak{C}_j$ に帰属すると認識するときの性能は悪くないと言えるであろう。

4.3.4 認識手続き煩雑さ COR(φ, $\underline{\mathfrak{C}}$[j]；SM) の最小値

式(4.52)の認識手続き煩雑さ COR(φ, $\underline{\mathfrak{C}}$[j]；SM)は

EOC(φ, $\underline{\mathfrak{C}}$[j]；SM)

$=-$DOC(φ, $\underline{\mathfrak{C}}$[j]；SM) ∵ 式(4.1)

$=\log_e\{SM(\varphi, \omega_j)/(1-SM(\varphi, \omega_j))\}$

∵ 式(3.71) (4.61)

を考慮すれば、次の補助定理4.1から、式(4.47)の下で、方程式(4.64)を満たす $SM^*(\varphi, \omega_j)$ を用いて、最小値 $-\exp$ [EOC(φ, $\underline{\mathfrak{C}}$[j]；SM*)] を持つことがわかる。更に、$SM^*(\varphi, \omega_j)$ は、

$$0<SM^*(\varphi, \omega_j)<1/2 \quad (4.62)$$

の範囲に存在することが2式(4.55), (4.56)からわかる。

[補助定理4.1]

変数 $(0<)v_j(<1)$ の実数値関数

$$f(v_j)=v_j\cdot\log_e\{v_j/(1-v_j)\} \quad (4.63)$$

について、

$$\min_{0<v_j<1} f(v_j)$$

$$=-\exp[\log_e\{v_j^*/(1-v_j^*)\}]. \quad (4.64)$$

が成り立つ。ここに、v_j^* は、方程式

$$0=f(v_j)+\exp[\log_e\{v_j/(1-v_j)\}] \quad (4.65)$$

の解である。

(証明) 先ず、

$$d/dv_j[v_j\cdot\log_e\{v_j/(1-v_j)\}]$$

$$=d/dv_j[v_j\cdot\log_e v_j-v_j\cdot\log_e(1-v_j)]$$

$$=\log e v_j+1-\log_e(1-v_j)+v_j/(1-v_j)$$

$$=\log_e[v_j/(1-v_j)]+1/(1-v_j)$$

$$=0 \quad (4.66)$$

が得られ、よって、式(4.64)が成り立つ。ところで、

$$d^2/dv_j^2[v_j\cdot\log_e\{v_j/(1-v_j)\}]$$

$$=1/v_j+1/(1-v_j)+1/(1-v_j)^2$$

$$=1/[v_j(1-v_j)^2]>0 \quad (4.67)$$

であるから、等式(4.64)を満たすv_j* は関数 $f(v_j)$に最小値を与えることがわかる。 □

5. 平均歪み $dtr(p(T\varphi/\mathfrak{C}_n[j])$；$n=1, 2)$ が指定された値 d 以下の場合、平均相互情報量 $AMI(\underline{\mathfrak{C}}[j]$；$\Phi[t_1, t_2])$の、$p(T\varphi/\mathfrak{C}_n[j])$を変えた場合の極値性

一般に、2式(4.41)，(4.44)の多段階パターン変換を採用し、最終段階のパターン ψ_t の帰属するカテゴリを処理の対象とする問題の入力パターン $\varphi\in\Phi$ の帰属するカテゴリであるとする多段階想起認識法では、**当初のパターン φ を出来るだけ変形しないで最終段階のパターン ψ_t が得られること**が望ましい。というのは、変形が少ない方が例えば、多段階パターン変換 $|t|$ が可能な限り小になるなどの利点が生まれ、$\varphi\in\Phi$ の認識に関する手続きの、4.3.2項でいう煩雑さが少なくてすむからである。

本章では、歪み測度 dtr を導入し、式(3.17)の平均相互情報量 AMI の下界を評価し、認識に関する歪みが最小限、どの程度存在するかを明らかにしてみよう。

5.1 axiom 2の類似度関数 SM の構成2例

axiom 2を満たす式(2.12)の類似度関数 SM を用いて、2式(3.63)，(3.64)のように、2つの事後出現確率 $p(\mathfrak{C}_1[j]/T\varphi)$，$p(\mathfrak{C}_2[j]/T\varphi)$を設定したとき、式(3.17)の平均相互情報量 $AMI(\underline{\mathfrak{C}}[j]$；$\Phi[t_1, t_2])$は、2式(3.15)，(3.61)を用いると実際に計算され得る。このとき、重要な2種類の SM として、次の定理5.1で指摘されるものがある。

[定理5.1]（**類似度関数 SM の構成定理**）

式(2.11)の代表パターンモデル系 $T\cdot\Omega$ が1次独立であり、非一致条件

$$\forall j\in J,\ \forall i\in J-\{j\},\ \|T\omega_i-T\omega_j\|>0 \quad (5.1)$$

が成立しているとしよう。このとき、

①，②のように定義される式(2.12)の関数 SM はaxiom 2を満たす：

①(指数関数系による構成)

$$\begin{aligned}&SM(\varphi,\ \omega_j)\\&=p(\mathfrak{C}_j)\cdot\exp[a\cdot\|T\varphi-T\omega_j\|^{-2}]\\&/\sum_{i\in J}p(\mathfrak{C}_i)\cdot\exp[a\cdot\|T\varphi-T\omega_i\|^{-2}\\&,a>0\end{aligned}\tag{5.2}$$

②(指数関数系，1次関数系による構成)

式(5.2)の $SM(\varphi,\ \omega_j)$を $SM'(\varphi,\ \omega_j)$と表記して

不等式

$$0\leqq\varepsilon_0(j)<\varepsilon_1(j)\leqq1\tag{5.3}$$

を満たす2つの閾値系 $\varepsilon_0(j)$，$\varepsilon_1(j)$ $(j\in J)$を選定して、

$$S(\varphi,\ \omega_j)=\begin{cases}1 & \cdots\varepsilon_1(j)\leqq SM'(\varphi,\ \omega_j)\leqq1\ \text{の場合}\\ [SM'(\varphi,\ \omega_j)-\varepsilon_0(j)]/[\varepsilon_1(j)-\varepsilon_0(j)] & \cdots\varepsilon_0(j)<SM'(\varphi,\ \omega_j)<\varepsilon_1(j)\text{の場合}\\ 0 & \cdots0\leqq SM'(\varphi,\ \omega_j)\leqq\varepsilon_0(j)\text{の場合}\end{cases}\tag{5.4}$$

として、

$$\mathrm{SM}(\varphi,\ \omega_j)=\mathrm{S}(\varphi,\ \omega_j)/\sum_{i\in J}\mathrm{S}(\varphi,\ \omega_i). \tag{5.5}$$

（証明）　容易に証明される。　□

5.2　p($\mathfrak{C}_n$[j]/Tφ)を変えた場合の、平均相互情報量 AMI($\underline{\mathfrak{C}}$[j]；Φ[t_1, t_2])の最小値

式(3.17)の AMI($\underline{\mathfrak{C}}$[j]；Φ[t_1, t_2])において 式(3.63)のp($\mathfrak{C}_1$[j]/Tφ)を変え、その極値を求めるために、先ず、次の補助定理5.1を証明する。

クロネッカー(Kronecker)の δ 記号

$$\delta_{jk}=1 \text{ if } j=k,\ =0 \text{ if } j\neq k \tag{5.6}$$

を導入しておく。

［補助定理5.1］（**多カテゴリ平均相互情報量 AMI(z_{jt}；j∈J, t∈{t_1, t_1+1, …, t_2})の、各事後確率 p($\mathfrak{C}_j$/Tφ_t)(j∈J) を変えた場合の最小値定理**）

$$y_t\equiv q(\mathrm{T}\varphi_t),\ t\in\{t_1, t_1+1, \cdots, t_2\} \tag{5.7}$$

$$z_{jt}\equiv p(\mathfrak{C}_j/\mathrm{T}\varphi_t),$$
$$j\in J,\ t\in\{t_1, t_1+1, \cdots, t_2\} \tag{5.8}$$

として、

$$x_j\equiv p(\mathfrak{C}_j)$$
$$=\sum_{t=t_1}^{t_2} p(\mathfrak{C}_j/\mathrm{T}\varphi_t)\cdot q(\mathrm{T}\varphi_t)$$
$$=\sum_{t=t_1}^{t_2} z_{jt}\cdot y_t \tag{5.9}$$

が成り立つが、正条件

$$\forall t\in\{t_1, t_1+1, \cdots, t_2\},\ y_t>0 \tag{5.10}$$

の下で、各 z_{jt} を変えて得られる多カテゴリ平均相互情報量と称されてよい非負量

$$\mathrm{AMI}\equiv\mathrm{AMI}(z_{jt}\ ;\ j\in J,\ t\in\{t_1, t_1+1, \cdots, t_2\})$$
$$\equiv\sum_{j\in J}\sum_{t=t_1}^{t_2} p(\mathfrak{C}_j/\mathrm{T}\varphi_t)\cdot q(\mathrm{T}\varphi_t)\cdot$$
$$\log_e[p(\mathfrak{C}_j/\mathrm{T}\varphi_t)/p(\mathfrak{C}_j)] \tag{5.11}$$
$$=\sum_{j\in J}\sum_{t=t_1}^{t_2} z_{jt}\cdot y_t\cdot\log_e[z_{jt}/x_j] \tag{5.12}$$
$$=\sum_{j\in J}\sum_{t=t_1}^{t_2} z_{jt}\cdot y_t\cdot\log_e[z_{jt}/\sum_{s=t_1}^{t_2} z_{js}\cdot y_s] \tag{5.13}$$
$$=\sum_{j\in J}\sum_{t=t_1}^{t_2} z_{jt}\cdot y_t\cdot\log_e z_{jt}$$
$$-\sum_{j\in J}\sum_{t=t_1}^{t_2} z_{jt}\cdot y_t\cdot\log_e\sum_{s=t_1}^{t_2} z_{js}\cdot y_s \tag{5.14}$$

の最小値は 0 であり、この最小値 0 は

$$\forall j\in J,\ z_{jt}=x_j \text{ for any } t\{t_1, t_1+1, \cdots, t_2\} \tag{5.15}$$

のときに生じる。

（証明）　AMI は多変数 z_{jt}, j∈J, t∈{t_1, t_1+1, …, t_2} の関数であり、極値を取るものとすれば、

$$\forall k\in J,\ \forall u\in\{t_1, t_1+1, \cdots, t_2\},$$
$$0=\partial\,\mathrm{AMI}/\partial z_{ku} \tag{5.16}$$

でなければならない。∂AMI/∂z_{jt} を計算すれば、

$$\partial\,\mathrm{AMI}/\partial z_{ju}$$

$$=\partial[z_{ku}\cdot y_u\cdot \log_e z_{ku}]/\partial z_{ku}$$
$$-y_u\cdot \log_e \sum_{s=t_1}^{t_2} z_{ks}\cdot y_s$$
$$-\sum_{j\in J}\sum_{t=t_1}^{t_2} z_{jt}\cdot y_t\cdot$$
$$\partial \log_e \sum_{s=t_1}^{t_2} z_{js}\cdot y_s/\partial z_{ku}$$
$$\because \quad 式(5.14)$$
$$=y_u\cdot[\log_e z_{ku}+1]$$
$$-y_u\cdot \log_e \sum_{s=t_1}^{t_2} z_{ks}\cdot y_s$$
$$-\sum_{j\in J}\sum_{t=t_1}^{t_2} z_{jt}\cdot y_t\cdot$$
$$[\sum_{s=t_1}^{t_2} z_{js}\cdot y_s]^{-1}\cdot \delta_{jk}\cdot y_u$$
$$=y_u\cdot[\log_e z_{ku}+1]$$
$$-y_u\cdot \log_e \sum_{s=t_1}^{t_2} z_{js}\cdot y_s$$
$$-\sum_{t=t_1}^{t_2} z_{jt}\cdot y_t\cdot$$
$$[\sum_{s=t_1}^{t_2} z_{js}\cdot y_s]^{-1}\cdot y_u$$
$$=y_u\cdot[\log_e z_{ku}+1]$$
$$-y_u\cdot \log_e \sum_{s=t_1}^{t_2} z_{js}\cdot y_s$$
$$-\sum_{t=t_1}^{t_2} z_{jt}\cdot y_t\cdot$$
$$[\sum_{s=t_1}^{t_2} z_{js}\cdot y_s]^{-1}\cdot y_u$$
$$=y_u\cdot[\log_e z_{ku}+1]$$
$$-y_u\cdot \log_e \sum_{s=t_1}^{t_2} z_{js}\cdot y_s-y_u$$
$$=y_u\cdot[\log_e z_{ku}/\sum_{s=t_1}^{t_2} z_{ks}\cdot y_s] \tag{5.17}$$

であるから、正条件式(5.10)を考慮すれば、式(5.16)から、

$$\forall k\in J,\ \forall u\in\{t_1, t_1+1, \cdots, t_2\},$$
$$1=z_{ku}/\sum_{s=t_1}^{t_2} z_{ks}\cdot y_s$$
$$=z_{ku}/x_k \quad \because \quad 式(5.14) \tag{5.18}$$

を得、一般に、定理E1での不等式(E.44)から直ちに判明する AMI の非負性

$$\mathrm{AMI}\geqq 0 \tag{5.19}$$

を考慮し、AMI を計算すれば、

$$\mathrm{AMI}$$
$$=\sum_{j\in J}\sum_{t=t_1}^{t_2} z_{jt}\cdot y_t\cdot \log_e x_j$$
$$-\sum_{j\in J}\sum_{t=t_1}^{t_2} z_{jt}\cdot y_t\cdot \log_e x_j$$
$$\because \quad 2式(5.14),\ (5.9)$$
$$=\sum_{j\in J} x_j\cdot \log_e x_j$$
$$-\sum_{j\in J} x_j\cdot \log_e x_j$$
$$\because \quad 式(5.9)$$
$$=0 \tag{5.20}$$

が得られ、証明が終わる。 □

上述の補助定理5.1を適用すれば、**2カテゴリ平均相互情報量**と称されてよい式(3.17)の AMI において、各類似度 SM(φ_t, ω_j) $(t\in\{t_1, t_1+1, \cdots, t_2\})$が変化した場合の次の定理5.2が成り立つ。

［定理5.2］（**式(3.17)の2カテゴリ平均相互情報量 AMI の最小値定理**）

1実多変数

$$y_t = \mathrm{SM}(\varphi_t,\ \omega_j),\quad t \in \{t_1, t_1+1, \cdots, t_2\} \tag{5.21}$$

の実数値関数とみなされた場合の、式(3.17)の2カテゴリ平均相互情報量

$$g(y_t) = \mathrm{AMI}(\underline{\mathfrak{C}}[j]\ ;\ \Phi[t_1, t_2]) \tag{5.22}$$

が最小値 0 をとるのは、

$$0 < p(\mathfrak{C}_n[j]/T\varphi_t)\,(=\mathrm{SM}(\varphi_t,\ \omega_j)) = q(\mathfrak{C}_1[j]) < 1,\ t \in \{t_1, t_1+1, \cdots, t_2\} \tag{5.23}$$

の成立のときに限る。 □

式(5.14)の AMI の第(j, t)成分

$$\mathrm{AMI}[j, t] \equiv p(\mathfrak{C}_j/T\varphi_t)\cdot q(T\varphi_t)\cdot \log_e[p(\mathfrak{C}_j/T\varphi_t)/p(\mathfrak{C}_j)] \tag{5.24}$$

$$= z_{jt}\cdot y_t\cdot \log_e[z_{jt}/x_j] \quad \left(= z_{jt}\cdot y_t\cdot \log_e[z_{jt}/\sum_{s=t_1}^{t_2} z_{js}\cdot y_s] \quad \because\ \text{式}(5.9)\right) \tag{5.25}$$

$$= y_t\cdot[-z_{jt}\cdot \log_e x_j - \{-z_{jt}\cdot \log_e z_{jt}\}] \tag{5.26}$$

$$\leqq y_t\cdot[-z_{jt}\cdot \log_e x_j] \tag{5.27}$$

について、式(5.26)から、次の(i)～(iii)が成り立つ：

(i) $z_{jt}=0$ であれば、

$$\mathrm{AMI}[j, t] = 0. \quad \because\ 0\log_e 0 = 0 \tag{5.28}$$

(ii) $z_{jt}=e^{-1}$ であれば、

$$\mathrm{AMI}[j, t] = y_t\cdot e^{-1}\cdot[-\log_e x_j{}^{-1}]. \tag{5.29}$$

(iii) $z_{jt}=2^{-1}$ であれば、

$$\mathrm{AMI}[j, t] = y_t\cdot 2^{-1}\cdot[-\log_e x_j - \log_e 2]. \tag{5.30}$$

(iv) $z_{jt}=1$ であれば、

$$\mathrm{AMI}[j, t] = -y_t\cdot \log_e x_j. \tag{5.31}$$

□

尚、上述の(ii)においては、$-z_{jt}\cdot\log_e z_{jt}$は最大値 e^{-1} をとっていることは、次の補助定理5.2からわかる。

［補助定理5.2］（**エントロピー関数 $-x\cdot\log_e x$ の最小値・最大値定理**）

$$f(x) = -x\cdot\log_e x \geqq 0\ (0 \leqq x \leqq 1) \tag{5.32}$$

について、

$$x \in \{0, 1\}\ \text{のとき、最小値}\ f(0) = f(1) = 0\ \text{をとる} \tag{5.33}$$

と、

$$x = e^{-1} \fallingdotseq 0.366\ \text{のとき、最大値}\ f(e^{-1}) = e^{-1}\text{をとる} \tag{5.34}$$

とが成り立つ。

（証明） 先ず、

$$df(x)/dx = -\log_e x - 1 = 0$$

$$\therefore \quad x = e^{-1} \tag{5.35}$$

であり、

$$f(e^{-1}) = e^{-1}$$

$$df(x)/dx = -1/x \leqq 0 \quad \text{for any} \quad x(0 \leqq x \leqq 1) \tag{5.36}$$

であることより、明らか。 □

確率条件

$$\sum_{j \in J} z_{js} = 1 \text{ for any } s \in \{t_1, t_1+1, \cdots, t_2\} \tag{5.37}$$

の下で、

$$\text{AMI}[j, t]/y_t \quad \because \quad 式(5.10)$$

$$= z_{jt} \cdot \log_e [z_{jt}/x_j] \tag{5.38}$$

$$= z_{jt} \cdot \log_e z_{jt} - z_{jt} \cdot \log_e \sum_{s=t_1}^{t_2} z_{js} \cdot y_s$$

$$\because \quad 2式(5.9), (5.25) \tag{5.39}$$

を極値ならしめる各変数 z_{jt} の値を求めよう。それは、次の定理5.3で与えられ、補助定理5.1の局所化に相当する。

［定理5.3］ **（式(5.24)の AMI の第(j, t)成分 AMI[j, t]）を式(5.7)の y_t で除した成分 AMI[j, t]/y_t の最小値定理）**

式(5.24)の AMI の第(j, t)成分 AMI[j, t] と、式(5.7)の y_t に関し、正条件式(5.10)の下で、

$$\forall t \in \{t_1, t_1+1, \cdots, t_2\},$$

$$z_{jt} = x_j \quad \text{for} \quad j \in J \tag{5.40}$$

のとき、AMI[j, t]/y_t は最小値 0 をとる。

（証明） 未定定数 λ を導入し、ラグランジュ(Lagrange)の未定乗数法を適用する。AMI[j, t]/y_t が極値を取るものとすれば、関数

$$F$$

$$\equiv z_{jt} \cdot \log_e z_{jt} - z_{jt} \cdot \log_e \sum_{s=t_1}^{t_2} z_{js} \cdot y_s$$

$$+ \lambda \cdot \left(\sum_{j \in J} z_{jt} - 1\right)$$

$$\because \quad 2式(5.37), (5.39) \tag{5.41}$$

について、

$$\forall j \in J, \forall t \in \{t_1, t_1+1, \cdots, t_2\},$$

$$0 = \partial F / \partial z_{jt} \tag{5.42}$$

が成立する。$\partial F/\partial z_{jt}$ を計算すれば、

$$\partial F/\partial z_{jt}$$

$$= \log_e z_{jt} + 1 - \log_e \sum_{s=t_1}^{t_2} z_{js} \cdot y_s$$

$$- z_{jt} \cdot \left[\sum_{s=t_1}^{t_2} z_{js} \cdot y_s\right]^{-1} \cdot y_t + \lambda \tag{5.43}$$

であるから、

$$\forall j \in J, 0 = \sum_{t=t_1}^{t_2} \partial F / \partial z_{jt}$$

$$= \sum_{t=t_1}^{t_2} \log_e z_{jt} + (t_2 - t_1 + 1)$$

$$- (t_2 - t_1 + 1) \cdot \log_e \sum_{s=t_1}^{t_2} z_{js} \cdot y_s$$

$$-1$$

$$+\lambda\cdot(t_2-t_1+1)$$
$$=\log_e\prod_{t=t_1}^{t_2} z_{jt}+\log_e e^{(t_2-t_1)}$$
$$-\log_e\left[\sum_{s=t_1}^{t_2} z_{js}\cdot y_s\right]^{t_2-t_1+1}$$
$$+\log_e e^{\lambda\cdot(t_2-t_1+1)}$$
$$=\log_e\left[\left\{\prod_{t=t_1}^{t_2} z_{jt}\right\}\cdot e^{(t_2-t_1)}e^{\lambda\cdot(t_2-t_1+1)}\right.$$
$$\left./\left\{\sum_{s=t_1}^{t_2} z_{js}\cdot y_s\right]^{t_2-t_1+1}\right\}\Big] \tag{5.44}$$
$$\therefore$$
$$\left\{\prod_{t=t_1}^{t_2} z_{jt}\right\}\cdot e^{(t_2-t_1)}e^{\lambda\cdot(t_2-t_1+1)}$$
$$/\left\{\sum_{s=t_1}^{t_2} z_{js}\cdot y_s\right]^{t_2-t_1+1}\Big\}=1 \tag{5.45}$$
$$\therefore$$
$$\prod_{t=t_1}^{t_2} z_{jt}$$
$$=e\cdot[x_j/e]^{t_2-t_1+1}\cdot e^{-\lambda\cdot(t_2-t_1+1)}$$
$$\because \quad \text{式}(5.9)$$
$$=x_j^{t_2-t_1+1}\cdot e^{1-(t_2-t_1+1)-\lambda\cdot(t_2-t_1+1)} \tag{5.46}$$

を得る。ここで、λ を

$$0=1-(t_2-t_1+1)-\lambda\cdot(t_2-t_1+1)$$
$$\therefore \quad \lambda=-(t_2-t_1)/(t_2-t_1+1) \tag{5.47}$$

ととれば、

$$\prod_{t=t_1}^{t_2} z_{jt}$$
$$=x_j^{t_2-t_1+1} \tag{5.48}$$

を得、この等式は、(5.40)のごとく、各 z_{jt} をとれば、満たされる。

このとき、

$$\mathrm{AMI}[j,t]/y_t$$
$$=z_{jt}\cdot\log_e[z_{jt}/x_j] \quad \because \quad \text{式}(5.38)$$
$$=0 \tag{5.49}$$

が得られ、証明が終わる。 □

5.3 誤り確率の評価法

$\Phi[t_1, t_2]$ 内のパターン列 $\varphi_{t_1}, \varphi_{t_1}+1, \cdots, \varphi_{t_2}$ が1つのパターン多段階認識で発生されるような式(4.44)のパターン列の場合、不等式

$$\mathrm{DOC}(\varphi_s, \underline{\mathfrak{C}}[j]; \mathrm{SM}) \geqq \mathrm{DOC}(\varphi_{s+1}, \underline{\mathfrak{C}}[j]; \mathrm{SM})$$
$$\text{for any } s\in\{0, 1, 2, \cdots, t\} \tag{5.50}$$

が成立し、2カテゴリ分類困難度 DOC は減少している筈である。

さて、2つの非負整数 N(j), n(j)を

N(j)：第 $j\in J$ 番目のカテゴリ $\mathfrak{C}_j$ に帰属するように生成された

$\Phi[t_1, t_2]$ 内のパターン $\varphi_t(t_1\leqq t\leqq t_2)$ の個数 $(0\leqq N(j)\leqq t_2-t_1+1)$ (5.51)

n(j)：N(j)個のパターン φ_t の内、不等式

$$\mathrm{SM}(\varphi_t, \omega_j)\leqq 1-\mathrm{SM}(\varphi_t, \omega_j)$$

を満たすパターン φ_t の個数 $(0\leqq n(j)\leqq N(j))$ (5.52)

と定義し、更に、1実変数 y

$$y=\log_e[\{1-SM(\varphi_t,\ \omega_j)\}/SM(\varphi,\ \omega_j)] \quad \because \ 式(3.72) \tag{5.53}$$

の、2値関数 f(y)を

$$f(y)\equiv \begin{cases} 1 \text{ if } y\geqq 0 \\ 0 \text{ if } y<0 \end{cases} \tag{5.54}$$

と定義すると、不等式

$$\forall y\in R(実数全体の集合),\quad f(y)\leqq \exp[by] \text{ for any } b\geqq 0 \tag{5.55}$$

が成り立つ。

確率分布関数

$$\mathrm{Prob}\{y\leqq z\}\equiv \int_{-\infty}^{z} dy\ h(y) \tag{5.56}$$

の近似値を、パターン集合 $\Phi[t_1, t_2]$ から出現頻度として求める。つまり、確率密度関数 h(y)を近似的に求めておく。

$\Phi[t_1, t_2]$ の統計的性質を反映している"変数 b の関数"

$$g(b)\equiv\int_{-\infty}^{+\infty} dy\ h(y)\cdot\exp[by] = \text{"}\exp[by]\text{ の期待値" for any } b\in R \tag{5.57}$$

を定義すれば、3性質

$$(イ)\ g(0)=\int_{-\infty}^{+\infty} dy\ h(y)=1 \tag{5.58}$$

$$(ロ)\ \forall b\in R,\quad dg(b)/db=\int_{-\infty}^{+\infty} dy\ h(y)\cdot y\cdot\exp[by] \tag{5.59}$$

$$(ハ)\ dg(b)/db\ |\ b=0=\int_{-\infty}^{+\infty} dy\ h(y)\cdot y = \text{"y の期待値"} \tag{5.60}$$

が成立することが直ちに、わかる。積分値 g(b)は誤認識確率の **Chernoff bound** と呼ばれる。

［定理5.4］（**誤認識確率定理**）

（i）$\Phi[t_1, t_2]$ 内の、第 $j\in J$ 番目のカテゴリ $\mathfrak{C}_j$ に帰属するように生成された各パターン φ が誤認識される確率 $\mathrm{Prob}\{y\geqq 0\}$

$$=\int_{-\infty}^{+\infty} dy\ h(y)\cdot f(y) \leqq g(b) \text{ for any } b>0 \tag{5.61}$$

（ii）式(5.57)の Chernoff bound g(b)の自然対数関数

$$G(b)\equiv\log_e g(b) \tag{5.62}$$

について、

$$\forall b\in R,\ dG(b)/db = [dg(b)/db]/g(b) \tag{5.63}$$

$$=\int_{-\infty}^{+\infty} dy\ h(y)\cdot y\cdot\exp[by] \Big/ \int_{-\infty}^{+\infty} dy\ h(y)\cdot\exp[by] \tag{5.64}$$

(iii) $\forall b \in R$, $d^2G(b)/db^2$

$$= \int_{-\infty}^{+\infty} dy\,[y - dG(b)/db]^2 \cdot h(y)\cdot \exp[by] / \int_{-\infty}^{+\infty} dw\, h(w)\cdot \exp[bw] \geqq 0 \tag{5.65}$$

(iv) $dG(b)/db$

$$\begin{cases} > \int_{-\infty}^{+\infty} dy\, h(y)\cdot y \text{ if } b > 0 \\ = \int_{-\infty}^{+\infty} dy\, h(y)\cdot y \text{ if } b = 0 \\ < \int_{-\infty}^{+\infty} dy\, h(y)\cdot y \text{ if } b < 0 \end{cases} \tag{5.66}$$

（証明）（ i ）式(5.55)から明らかである。

（ii）微分公式

$$d\log_e x/dx = 1/x \text{ for any } x > 0 \tag{5.67}$$

と、（ロ），式(5.57)とより明らか。

（iii）先ず、

$$d^2g(b)/db^2 = \int_{-\infty}^{+\infty} dy\, h(y)\cdot y^2 \cdot \exp[by] \tag{5.68}$$

に注意する。

$$\forall b \in R,\ d^2G(b)/db^2$$

$$= d[\{dg(b)/db\}/g(b)]/db \quad \because \text{ 式(5.63)} \tag{5.69}$$

$$= \{d^2g(b)/db^2\}/g(b) - [\{dg(b)/db\}/g(b)]^2 \tag{5.70}$$

$$= \int_{-\infty}^{+\infty} dy\, y^2 \cdot 【h(y)\cdot \exp[by] / \int_{-\infty}^{+\infty} dw\, h(w)\cdot \exp[bw]】 - 【\int_{-\infty}^{+\infty} dy\, y\{h(y)\cdot \exp[by] / \int_{-\infty}^{+\infty} dz\, h(z)\cdot \exp[bz]\}】^2$$
$$\because \text{ 3式(5.68), (5.57), (5.64)} \tag{5.71}$$

$$= \int_{-\infty}^{+\infty} dy\,[y - dG(b)/db]^2 \cdot h(y)\cdot \exp[by] / \int_{-\infty}^{+\infty} dw\, h(w)\cdot \exp[bw] \quad \because \text{ 式(5.64)} \tag{5.72}$$

$$\geqq 0 \tag{5.73}$$

を得、証明が終わる。

（iv）上述の（iii）より、

$$dG(b)/db \begin{cases} > dG(b)/db \mid b = 0 \text{ if } b > 0 \\ = dG(b)/db \mid b = 0 \text{ if } b = 0 \\ < dG(b)/db \mid b = 0 \text{ if } b < 0 \end{cases} \tag{5.74}$$

であるが、これに、

$dG(b)/db\mid_{b=0}$

$=[dg(b)/db]/g(b)\mid_{b=0}$ (5.75)

$=[dg(b)/db]\mid_{b=0}/g(0)$

$=\int_{-\infty}^{+\infty} dy\ h(y)\cdot y/1$

$\because$ (ロ), (イ) (5.76)

$=\int_{-\infty}^{+\infty} dy\ h(y)\cdot y$ (5.57)

を考慮したものである。 □

5.4 平均伝送歪み ADM が指定されたときの、平均相互情報量 AMI の最小値を与える通信路行列の算出

5.4.1 $p(T\varphi/\mathfrak{C}_n[j])$を変えた場合の、2カテゴリ相互情報量 $AMI(\underline{\mathfrak{C}}[j];\Phi[t_1,t_2])$の条件付き極小値

これまでは、3.2節で、2条件式(3.11), (3.12)を満たす“パターンモデル $T\varphi$ の出現確率” $q(T\varphi)$と、2条件式(3.13), (3.14)を満たす“カテゴリ $\mathfrak{C}_n[j]$ の条件付き出現確率” $p(\mathfrak{C}_n[j]/T\varphi)$とを与え、式(3.17)の2カテゴリ平均相互情報量 $AMI(\underline{\mathfrak{C}}[j];\Phi[t_1,t_2])$を定義し、論を組み立てて来た。

“カテゴリ $\mathfrak{C}_n[j]$ の出現確率” $p(\mathfrak{C}_n[j])$が式(3.15)の如く定義され、式(3.16)を満たしているように出来た。“$\mathfrak{C}_n[j]$ が出現したときの、$T\varphi$ の条件付き出現確率” $p(T\varphi/\mathfrak{C}_n[j])$は

$p(T\varphi/\mathfrak{C}_n[j])$

$\equiv q(T\varphi)\cdot p(\mathfrak{C}_n[j]/T\varphi)/p(\mathfrak{C}_n[j])$ (5.58)

$=q(T\varphi)\cdot p(\mathfrak{C}_n[j]/T\varphi)$

$/\Sigma_{\varphi\in\Phi[t_1,t_2]}\ q(T\varphi)\cdot p(\mathfrak{C}_n[j]/T\varphi)$

$\because$ 式(3.15) (5.59)

と定義でき、確率条件

$\forall n\in\{1,2\},\ \Sigma_{\varphi\in\Phi[t_1,t_2]}\ p(T\varphi/\mathfrak{C}_n[j])=1$ (5.60)

を満たすことが直ちにわかる。

逆に、入力確率分布

$p(\mathfrak{C}_n[j]),\ n\in\{1,2\}$ (5.61)

が、2条件

$[\forall n\in\{1,2\},\ 0\leqq p(\mathfrak{C}_n[j])\leqq 1]$ (5.62)

$\wedge[\sum_{n=1}^{2} p(\mathfrak{C}_n[j])=1]$ (5.63)

を満たすように与えられ、かつ、入力 $\mathfrak{C}_n[j]$ が与えられた時、出力 $T\varphi$ が出現する条件付き確率分布

$p(T\varphi/\mathfrak{C}_n[j]),\ \varphi\in\Phi[t_1,t_2]$ (5.64)

が、2条件

$[\forall\varphi\in\Phi[t_1,t_2],\ 0\leqq p(T\varphi/\Phi[t_1,t_2])\leqq 1]$ (5.65)

$\wedge[\Sigma_{\varphi\in\Phi[t_1,t_2]}\ p(T\varphi/\mathfrak{C}_n[j])=1]$ (5.66)

を満たすように与えられた場合、“パターンモデル $T\varphi$ の出現確率” $q(T\varphi)$を

$q(T\varphi)$

$\equiv\sum_{n=1}^{2} p(\mathfrak{C}_n[j])\cdot p(T\varphi/\mathfrak{C}_n[j]),$

$\varphi \in \Phi[t_1, t_2]$ (5.67)

とおくと、確率条件

$$\Sigma_{\varphi \in \Phi[t_1, t_2]}\ q(T\varphi) = 1$$

$\because$ 2式(5.66), (5.63) (5.68)

が成り立つ。また、

$$p(\mathfrak{C}_n[j]/T\varphi)$$
$$\equiv p(\mathfrak{C}_n[j]) \cdot p(T\varphi/\mathfrak{C}_n[j])/q(T\varphi) \quad (5.69)$$
$$= p(\mathfrak{C}_n[j]) \cdot p(T\varphi/\mathfrak{C}_n[j])$$
$$/\sum_{n=1}^{2} p(\mathfrak{C}_n[j]) \cdot p(T\varphi/\mathfrak{C}_n[j])$$

$\because$ 式(5.67) (5.70)

と定義でき、確率条件

$$\forall\varphi \in \Phi[t_1, t_2],\ \sum_{n=1}^{2} p(\mathfrak{C}_n[j]/T\varphi) = 1 \quad (5.71)$$

を満たすことが直ちにわかる。

以上では、2つの条件付き確率 $p(\mathfrak{C}_n[j]/T\varphi)$, $p(T\varphi/\mathfrak{C}_n[j])$について、$\mathfrak{C}_n[j]$, $T\varphi$ が同時に出現する結合確率 $p(\mathfrak{C}_n[j], T\varphi)$の2表現

$$\forall\varphi \in \Phi[t_1, t_2],\ \forall n \in \{1, 2\},$$
$$p(\mathfrak{C}_n[j], T\varphi)$$
$$\equiv p(\mathfrak{C}_n[j]/T\varphi) \cdot q(T\varphi) \quad (5.72)$$
$$p(\mathfrak{C}_n[j], T\varphi)$$
$$\equiv p(\mathfrak{C}_n[j]) \cdot p(T\varphi/\mathfrak{C}_n[j]) \quad (5.73)$$

を利用し、等式

$$p(\mathfrak{C}_n[j]/T\varphi) \cdot q(T\varphi)$$
$$= p(\mathfrak{C}_n[j]) \cdot p(T\varphi/\mathfrak{C}_n[j])$$

$\because$ 2式(5.69), (5.58) (5.74)

の成立を前提としているがわかる。

$p(T\varphi/\mathfrak{C}_n[j])$を変えると、式(3.17)の2カテゴリ相互情報量 AMI($\underline{\mathfrak{C}}[j]$; $\Phi[t_1, t_2]$)は変動するが、出力 $T\varphi$ が入力としてのカテゴリ $\mathfrak{C}_n[j]$ に帰属するということが出来るだけ保存されるとい条件の下で、この平均認識歪み AMI の極小値が付録Fで求められている。

5.4.2 シャノンの2つの情報理論

符号化された信号が送信され、受信･復号化されるとき“受信信号の条件つき出現確率確率分布”としてモデル化される通信路(channel)が主体となって論が展開されるシャノンの情報理論(通信の確率モデルに関する数学的理論)には、2つの流れ「情報伝送の理論(transmission theory)」,「情報縮約の理論(rate-distortion theory)」があり、各々に「符号化理論(coding theory)、復号化の理論(decoding theory)」が主として関係している。前者は与えられた確率特性の通信路を固定して、**最大**伝送速度(maximum transmission rate ; 単位時間に入力を伝送できる情報量の最大値)を達成するのに伝送される入力情報を如何に符号化するのかを考え、後者は可能な限り原信号を忠実に再生するかの観点から、与えられた入力アルファベットの確率分布を固定して、**最小**伝送歪み(minimum transmission distortion ; 受信する出力情報が入力情報に比べどの程度歪んでいるかに関する最小値)を達成するのに如何なる確率特性の通信路を使用すべきかを研究している。

各入力(カテゴリ)$\mathfrak{C}_n[j]$の、式(3.15)の出現確率 $p(\mathfrak{C}_n[j])\ (n \in \{1, 2\})$を変えて、与えられた"各入力 $\mathfrak{C}_n[j]$が観測された条件の下で、各出力(パターンモデル)$T\varphi$の出現に関する条件付き確率(**通信路行列**)" $p(T\varphi/\mathfrak{C}_n[j])\ (\varphi \in \Phi[t_1, t_2])$が与えられた場合、式(3.17)の平均相互情報量

$$\mathrm{AMI}(\underline{\mathfrak{C}}[j]\ ;\ \Phi[t_1, t_2])$$
$$= \sum_{n=1}^{2} \sum_{\varphi \in \Phi[t_1, t_2]}$$
$$p(\mathfrak{C}_n[j]) \cdot p(T\varphi/\mathfrak{C}_n[j]) \cdot$$
$$\log_e[p(T\varphi/\mathfrak{C}_n[j])/q(T\varphi)]$$
$$\because\ 3式(5.72) \sim (5.74) \qquad (5.75)$$

の最大値を求め、この最大値としての**通信容量**(channel capacity)を求めるのが、従来の**シャノンの最大伝送速度に関する情報理論**であり、この結果、通信容量をもたらす最適な入力確率分布

$$p(\mathfrak{C}_n[j])\ (n \in \{1, 2\}) \qquad (5.76)$$

が判明する。この最適入力確率分布に従って、各入力を出現させ、最大伝送速度を達成することになる。

情報縮約の方に目を向けよう。

歪み測度(distortion measure)、忠実度規準(fideity criterion)$\mathrm{dtr}(T\varphi/\underline{\mathfrak{C}}[j])$を先ず、説明しよう。任意の $T\varphi \in T \cdot \Phi[t_1, t_2]$の情報圧縮に伴う歪み(distortion)$\mathrm{dtr}(T\varphi/\mathfrak{C}_n[j])$は、一般には、axiom 2 を満たす式(2.12)の類似関数 SM を使って、

$$\mathrm{dtr}(T\varphi/\mathfrak{C}_n[j]) =$$
$$\begin{cases} 1 - \mathrm{SM}(T\varphi, T\omega_j) \cdots n = 1 \text{ のとき} \\ 1 - \max_{i \in J - \{j\}} \mathrm{SM}(T\varphi, T\omega_i) \cdots n = 2 \text{ のとき} \end{cases} \qquad (5.77)$$

と表されると想定されてよい。簡単には、

$$\mathrm{dtr}(T\varphi/\mathfrak{C}_n[j]) =$$
$$\begin{cases} 0 \\ \quad \cdots n = 1 \wedge T\varphi = T\omega_j \text{ のとき} \\ \alpha \text{ (正なる実定数)} \\ \quad \cdots n = 2 \vee T\varphi = T\omega_j \text{ のとき} \end{cases} \qquad (5.78)$$

と設定されてよい。以後、歪み $\mathrm{dtr}(T\varphi/\mathfrak{C}_n[j])$を $\rho(\mathfrak{C}_n[j], T\varphi)$で表そう。そうすると、2式(5.77), (5.78)の他に、

$$\rho(\mathfrak{C}_n[j],\ T\varphi) =$$
$$\begin{cases} \| T\varphi - T\omega_j \|^2 \qquad \cdots \varphi \text{ が第}j \in J\text{番目の} \\ \quad \text{カテゴリ } \mathfrak{C}_j \text{ に帰属するように生成されたとき} \\ \max_{i \in J - \{j\}} \| T\varphi - T\omega_i \|^2 \cdots \varphi \text{ が第}i \in J - \{j\} \\ \quad \text{番目のカテゴリ } \mathfrak{C}_i \text{ に帰属するように生成されたとき} \end{cases} \qquad (5.79)$$

などのように、

$$\rho(\mathfrak{C}_n[j],\ T\varphi) \geqq 0：\mathfrak{C}_n[j]\text{を表すように、}T\varphi\text{を生成するときの歪み(distortion)} \qquad (5.80)$$

を与え、各入力 $\mathfrak{C}_n[j]\ (n \in \{1, 2\})$が対応する出力 $T\varphi\ (\varphi \in \Phi[t_1, t_2])$として可能な限り忠実に生成されるときの平均ひずみ測度

$$\sum_{n=1}^{2} p(\mathfrak{C}_n[j]) \cdot$$
$$\sum_{\varphi \in \Phi[t_1, t_2]} p(T\varphi/\mathfrak{C}_n[j]) \cdot \rho(\mathfrak{C}_n[j], T\varphi) \qquad (5.81)$$

を固定して、平均相互情報量 AMI($\underline{\mathfrak{C}}$ [j]；Φ[t_1, t_2])を最小にする各条件つき確率分布

$$p(\mathfrak{C}_n[j]/T\varphi)(n \in \{1, 2\}) \tag{5.82}$$

をもとめるのが、従来のシャノンの rate-distortion theory (**情報伝送・再生の歪み理論**)である。

5.4.3 一般の場合の伝送歪みに関する情報理論

$$\text{入力アルファベット}\quad a_i,\ i=1\sim n \tag{5.82}$$

$$\text{出力アルファベット}\quad b_j,\ j=1\sim m \tag{5.83}$$

を考えるが、以下の適用を考えるには、対応

$$a_1, a_2 \longleftrightarrow \mathfrak{C}_1[j],\ \mathfrak{C}_2[j] \tag{5.84}$$

$$b_1, b_2, \cdots, bm \longleftrightarrow \varphi_{t1},\ \varphi_{t1}+1, \cdots,\ \varphi_{t2} \tag{5.85}$$

と考えればよい。また、以下では、prob{…} は確率事象(probabilistic event)… の出現確率の意である。

$$p_i \equiv \text{prob}\{a_i\},\ i=1\sim n \tag{5.86}$$

$$Q_{ji} \equiv \text{prob}\{b_j/a_i\},\ \ i=1\sim n,\ j=1\sim m \tag{5.87}$$

が与えられたとしているから、

$$q_j \equiv \text{prob}\{b_j\} \tag{5.88}$$

$$= \sum_{i=1}^{n} \text{prob}\{a_i\} \cdot \text{prob}\{b_j/a_i\} \tag{5.89}$$

$$= \sum_{i=1}^{n} p_i \cdot Q_{ji},\ j=1\sim m \tag{5.90}$$

と各 q_j が求まり、更に、

$$\text{prob}\{a_i/b_j\}$$

$$= \text{prob}\{b_j/a_i\} \cdot \text{prob}\{a_i\}/\text{prob}\{b_j\} \tag{5.91}$$

$$= \text{prob}\{b_j/a_i\} \cdot \text{prob}\{a_i\} \Big/ \sum_{i=1}^{N} \text{prob}\{a_i\} \cdot \text{prob}\{b_j/a_i\} \tag{5.92}$$

$$= Q_{ji} \cdot p_i \Big/ \sum_{k=1}^{n} Q_{jk} \cdot p_k,\quad i=1\sim n,\ j=1\sim m \tag{5.93}$$

と各 prob{a_i/b_j} が求まる。

ここで、a_i が b_j として再生されるときの費用(cost)、つまり、歪み測度(distortion measure)

$$\rho(i, j) \geqq 0 \tag{5.94}$$

を導入し、次の最小化問題を解決することを考えよう。

【平均歪み ADM が一定値 d であると指定されたときの、平均相互情報量 AMI を最小にする通信路行列 Q＝(Q_{ji})を求める最小化問題】

非正定数 s が与えられたとする。平均相互情報量

$$\text{AMI} \equiv \sum_{i=1}^{n} \sum_{j=1}^{m} p_i \cdot Q_{ji} \cdot \log_e [Q_{ji}/q_j] \tag{5.95}$$

を、3条件

$$\forall i \in \{1, 2, \cdots, n\},\ \sum_{j=1}^{m} Q_{ji} = 1\ (\text{条件付き確率}) \tag{5.96}$$

$$\text{ADM} \equiv \sum_{i=1}^{n} \sum_{j=1}^{m} p_i \cdot Q_{ji} \cdot \rho(i, j) = d$$

$$(\text{平均歪み；average distortion measure}) \tag{5.97}$$

の下で、最小化する通信路行列 Q＝(Q_{ji})を求めよ。但し、出力確率の表現

$$q_j = \sum_{k=1}^{n} Q_{jk} \cdot p_k,\ j=1\sim m \tag{5.98}$$

が成立しており、入力確率の正条件

$$\forall i \in \{1, 2, \cdots, n\},\ \ p_i > 0 \tag{5.99}$$

を設定しているものとする。

その最小化問題を解は次のように述べられる。

【上記の最小化問題の解 Q＝(Q) と、AMI の最小値】

条件付き確率 Q_{ji} が

$$Q_{ji} = q_j \cdot \exp[s \cdot \rho(i, j)] / \sum_{\ell=1}^{m} q_\ell \cdot \exp[s \cdot \rho(i, \ell)],\quad i = 1 \sim n,\ j = 1 \sim m \tag{5.100}$$

であるとき、式(5.95)の AMI は最小値になり、各 λ_i を、

$$\lambda_i = 1 / \sum_{\ell=1}^{m} q_\ell \cdot \exp[s \cdot \rho(i, \ell)],\quad i = 1 \sim n \tag{5.101}$$

とおくと、AMI の最小値は

$$\min \mathrm{AMI} = s \cdot d + \sum_{i=1}^{n} p_i \cdot \log_e \lambda_i \tag{5.102}$$

と求められ、式(5.97)の d は

$$d = \sum_{i=1}^{n} \sum_{j=1}^{m} \lambda_i \cdot p_i \cdot q_j \cdot \exp[s \cdot \rho(i, j)] \cdot \rho(i, j) \tag{5.103}$$

と表現される。更に、

$$\begin{aligned} &\mathrm{prob}\{a_i/b_j\} \\ &= \mathrm{prob}\{b_j/a_i\} \cdot \mathrm{prob}\{a_i\} / \mathrm{prob}\{b_j\} \\ &= \mathrm{prob}\{b_j/a_i\} \cdot \mathrm{prob}\{a_i\} / \sum_{k=1}^{n} \mathrm{prob}\{a_k\} \cdot \mathrm{prob}\{b_j/a_k\} \\ &= Q_{ji} \cdot p_i / \sum_{k=1}^{n} Q_{jk} \cdot p_k \\ &= p_i \cdot \exp[s \cdot \rho(i, j)] / \sum_{k=1}^{n} p_k \cdot \exp[s \cdot \rho(k, j)] \\ &,\ i = 1 \sim n,\ j = 1 \sim m \end{aligned} \tag{5.104}$$

と各 $\mathrm{prob}\{a_i/b_j\}$ が求まり、また、

$$\begin{aligned} q_j &\equiv \mathrm{prob}\{b_j\} = \sum_{i=1}^{n} \mathrm{prob}\{a_i\} \cdot \mathrm{prob}\{b_j/a_i\} \\ &= \sum_{i=1}^{n} p_i \cdot Q_{ji} \\ &= \sum_{i=1}^{n} p_i \cdot 【 q_j \cdot \exp[s \cdot \rho(i, j)] / \sum_{\ell=1}^{m} q_\ell \cdot \exp[s \cdot \rho(i, \ell)] 】, \\ & j = 1 \sim m \end{aligned} \tag{5.105}$$

と各 q_j が求まる。

上記の解 Q を導出する計算は次の通りである。

【平均歪み ADM が一定値 d であると指定されたときの、平均相互情報量 AMI を最小にする通信路行列　Q＝(Q_{ji}) の導出と、AMI の最小値の計算】

シャノン情報理論で解決されており [A4]、少し、詳細に説明する。

ラグランジュの未定定数法を適用する。

確率条件式(5.96)にラグランジュの未定定数 $\mu_i (i=1\sim m)$ を対応させ、また、平均歪み式(5.97)にラグランジュの未定定数 s を対応させて、関数

$$\begin{aligned}
&F(Q)\\
&\equiv AMI-\sum_{i=1}^{n}\mu_i\cdot\left[\sum_{j=1}^{m}Q_{ji}-1\right]\\
&-s\cdot[ADM-d]\\
&=\sum_{i=1}^{n}\sum_{j=1}^{m}p_i\cdot Q_{ji}\cdot\log_e[Q_{ji}/q_j]\\
&-\sum_{i=1}^{n}\mu_i\cdot\left[\sum_{j=1}^{m}Q_{ji}-1\right]\\
&-s\cdot\left[\sum_{i=1}^{n}\sum_{j=1}^{m}p_i\cdot Q_{ji}\cdot\rho(i,j)-d\right]
\end{aligned}\tag{5.106}$$

を設ける。このとき、関数 F(Q)が最小値をとるものとすれば、

$$0=\partial F(Q)/\partial Q_{ji},\ i=1\sim n,\ j=1\sim m$$

が成り立つ。

ここで、

$$\begin{aligned}
&①\partial\sum_{i=1}^{n}\sum_{j=1}^{m}p_i\cdot Q_{ji}\cdot\log_e[Q_{ji}/q_j]\\
&/\partial Q_{ji}\\
&=\partial\sum_{i=1}^{n}\sum_{j=1}^{m}p_i\cdot Q_{ji}\cdot\log_e(1/q_j)/\partial Q_{ji}\\
&+\partial\sum_{i=1}^{n}\sum_{j=1}^{m}p_i\cdot Q_{ji}\cdot\log_e Q_{ji}/\partial Q_{ji}
\end{aligned}$$

であるが、右辺の第1項，第2項を各々、計算すれば、

①－1

右辺の第1項

$$\begin{aligned}
&=\partial\sum_{i=1}^{n}\sum_{j=1}^{m}p_i\cdot Q_{ji}\cdot\log_e(1/q_j)/\partial Q_{ji}\\
&=-\partial\sum_{i=1}^{n}\sum_{j=1}^{m}p_i\cdot Q_{ji}\cdot\\
&\quad\log_e\sum_{k=1}^{n}Q_{jk}\cdot p_k/\partial Q_{ji}\quad\because\ \text{式}(5.98)\\
&=-p_i\cdot\log_e\sum_{k=1}^{n}Q_{jk}\cdot p_k\\
&\quad-\sum_{k=1}^{n}\sum_{\ell=1}^{m}p_k\cdot Q_{\ell k}\cdot\partial\log_e\sum_{k=1}^{n}Q_{jk}\cdot p_k/\partial Q_{ji}\\
&=p_i\cdot\log_e(1/q_j)\\
&-\sum_{k=1}^{n}\sum_{\ell=1}^{m}p_k\cdot Q_{\ell k}\cdot\left[\sum_{k=1}^{n}Q_{jk}\cdot p_k\right]^{-1}\cdot p_i\cdot\delta_{\ell j}\\
&=p_i\cdot\log_e(1/q_j)\\
&-\sum_{k=1}^{n}p_k\cdot Q_{jk}\cdot\left[\sum_{k=1}^{n}Q_{jk}\cdot p_k\right]^{-1}\cdot p_i\\
&=p_i\cdot\log_e(1/q_j)-p_i
\end{aligned}$$

①－2

右辺の第2項

$$\begin{aligned}
&=\partial\sum_{i=1}^{n}\sum_{j=1}^{m}p_i\cdot Q_{ji}\cdot\log_e Q_{ji}/\partial Q_{ji}\\
&=\partial[p_i\cdot Q_{ji}\cdot\log_e Q_{ji}]/\partial Q_{ji}\\
&=p_i\cdot\log_e Q_{ji}+p_i
\end{aligned}$$

$$\begin{aligned}
&②\partial\sum_{i=1}^{n}\mu_i\cdot\left[\sum_{j=1}^{m}Q_{ji}-1\right]/\partial Q_{ji}\\
&=\mu_i
\end{aligned}$$

③$\partial[\sum_{i=1}^{n}\sum_{j=1}^{m} p_i \cdot Q_{ji} \cdot \rho(i, j) - d]$
$/\partial Q_{ji}$
$= p_i \cdot \rho(i, j)$

であるから、結局、

$\forall i (=1 \sim n), \ \forall j (=1 \sim m),$
$0 = \partial F(Q) / \partial Q_{ji}$
$= p_i \cdot \log_e(1/q_j) - p_i$
$+ p_i \cdot \log_e Q_{ji} + p_i$
$- \mu_i - s \cdot p_i \cdot \rho(i, j)$
$= p_i \cdot \log_e(Q_{ji}/q_j) - \mu_i - s \cdot p_i \cdot \rho(i, j)$
$= p_i \cdot [\log_e(Q_{ji}/q_j) - \mu_i/p_i - s \cdot \rho(i, j)]$
$\because$ 式(5.99)
$= p_i \cdot [\log_e\{Q_{ji}/(\lambda_i \cdot q_j)\} - s \cdot \rho(i, j)]$
$\because \lambda_i \equiv \exp[\mu_i/p_i]$

を得、式(5.99)を考慮すれば、

$\forall i (=1 \sim n), \ \forall j (=1 \sim m),$
$\log_e\{Q_{ji}/(\lambda_i \cdot q_j)\} - s \cdot \rho(i, j) = 0$
$\therefore Q_{ji} = \lambda_i \cdot q_j \cdot \exp[s \cdot \rho(i, j)]$ (5.107)

と表現される。式(5.96)を考慮すれば、

$\forall i \in \{1, 2, \cdots, n\},$
$1 = \sum_{j=1}^{m} Q_{ji}$
$= \lambda_i \cdot \sum_{j=1}^{m} q_j \cdot \exp[s \cdot \rho(i, j)]$
$\because$ 式(5.107)

を得、各 λ_i は、式(5.101)の如く求まる。式(5.101)を式(5.107)に代入すれば、式(5.100)が得られる。

式(5.100)を式(5.95)に代入すれば、

min AMI
$= \min \sum_{i=1}^{n}\sum_{j=1}^{m} p_i \cdot Q_{ji} \cdot \log_e[Q_{ji}/q_j]$
$= \sum_{i=1}^{n}\sum_{j=1}^{m} p_i \cdot Q_{ji} \cdot$
$\log_e【\lambda_i \cdot \exp[s \cdot \rho(i, j)]】$ $\because$ 式(5.107)
$= \sum_{i=1}^{n}\sum_{j=1}^{m} p_i \cdot Q_{ji} \cdot s \cdot \rho(i, j)$
$+ \sum_{i=1}^{n}\sum_{j=1}^{m} p_i \cdot Q_{ji} \cdot \log e \lambda_i$
$= s \cdot d + \sum_{i=1}^{n} p_i \cdot \log_e \lambda_i$
$\because$ 2式(5.97), (5.96)

と、式(5.102)が求まった。

式(5.103)は式(5.97)に式(5.107)を代入したものである。 □

次の定理5.5は、式(5.95)の平均相互情報量 AMI が式(5.97)の平均歪み ADM の1次式として表されることを明らかにしており、平均相互情報量 AMI が平均歪み ADM の、不等式(5.109)の右辺

の1次関数より小さくできないことを指摘している。

［定理5.5］(**平均伝送歪みとしての平均相互情報量 AMI の下界定理**)

式(5.95)の平均相互情報量 AMI と、式(5.97)の平均歪み ADM について考えよう。

パラメータ s(≦0)と、1 より大きくない非負条件

$$\sum_{i=1}^{n} \lambda_i \cdot p_i \cdot \exp[s \cdot \rho(i, j)] \leqq 1 \tag{5.108}$$

を満足する各 $\lambda_i (1 \leqq i \leqq n)$に対し、不等式

$$\mathrm{AMI} \geqq s \cdot \mathrm{ADM} + \sum_{i=1}^{n} p_i \cdot \log_e \lambda_i \tag{5.109}$$

が成立する。

(証明)

$$\begin{aligned}
&\mathrm{AMI} - s \cdot \mathrm{ADM} - \sum_{i=1}^{n} p_i \cdot \mathrm{loge}\, \lambda_i \\
&= \mathrm{AMI} - s \cdot \mathrm{ADM} - \sum_{i=1}^{n} \sum_{j=1}^{m} p_i \cdot Q_{ji} \cdot \mathrm{loge}\, \lambda_i \\
&\qquad \because \text{ 式}(5.96) \\
&= \sum_{i=1}^{n} \sum_{j=1}^{m} - p_i \cdot Q_{ji} \cdot \\
&\qquad \log_e \left[q_j \cdot \lambda_i \cdot \exp[s \cdot \rho(i, j)] / Q_{ji} \right] \\
&\geqq \sum_{i=1}^{n} \sum_{j=1}^{m} p_i \cdot Q_{ji} \cdot \\
&\qquad [1 - q_j \cdot \lambda_i \cdot \exp[s \cdot \rho(i, j)] / Q_{ji}] \\
&\qquad \because \text{ 式}(3.103) \\
&= 1 - \sum_{i=1}^{n} \sum_{j=1}^{m} p_i \cdot q_j \cdot \lambda_i \cdot \\
&\quad \exp[s \cdot \rho(i, j)] \\
&\qquad \because \sum_{i=1}^{n} \sum_{j=1}^{m} p_i \cdot Q_{ji} = 1 \\
&= 1 - \sum_{j=1}^{m} q_j \cdot \sum_{i=1}^{n} p_i \cdot \lambda_i \cdot \\
&\quad \exp[s \cdot \rho(i, j)] \\
&\geqq 1 - 1 \quad \because \sum_{j=1}^{m} q_j = 1 \wedge \text{式}(5.108) \\
&= 0.
\end{aligned}$$

□

6. 大分類関数 BSC の構成への応用

認識システム RECOGNITRON を構成するのには、axiom 1を満たすパターン集合 Φ と、モデル構成作用素 T との対［Φ, T］を選定し、axiom 2, 3を各々、満たす類似度関数 SM, 大分類関数 BSC を選定すればよい［B4］。そうすれば、文献［B4］の定理A4.1を適用しカテゴリ選択関数 CSF の構造形式が定まり、文献［B4］の付録5での設定より、構造受精変換 TA(μ)T が確定し、構造受精変換の不動点を探索する“パターン $\varphi \in \Phi$ の**カテゴリ帰属知識**(categorical membership-knowledge)を処理する多段階パターン認識の働き”が設定されることになる。

本章では、この種の多段階認識の困難度を表している2カテゴリ分類困難度 DOC(φ, $\underline{\mathfrak{C}}$[j]; SM)の構造形式からヒントを得、axiom 3を満たす大分類関数 BSC を構成する。

6.1 axiom 3を満たす大分類関数 BSC

大分類関数(rough classifier, binary-state classifier)と呼ばれる**2値関数**

$$\mathrm{BSC}: \Phi \times J \to \{0, 1\} \tag{6.1}$$

を、次のaxiom 3 を満たすものとして導入し、解釈

パターン $\varphi\in\Phi$の帰属するカテゴリ候補の1つが第$j\in J$番目のカテゴリ$\mathfrak{C}_j$であるならば、BSC$(\varphi, j)=1$ であることが望ましい (6.2)

を採用しょう。この際、注意すべきは、文献［B4］の付録4，axiom 4のカテゴリ選択関数CSF(φ, γ)の定義での(iii)の場合からわかるように、

BSC$(\varphi, j)=0$ であっても、パターン$\varphi\in\Phi$の帰属するカテゴリ候補の1つは、第$j\in J$番目のカテゴリ$\mathfrak{C}_j$でないとは限らない (6.3)

としていることである。また、axiom 3 の(i)からわかるように、**カテゴリ間の相互排除性**(the mutual exclusion of the one category from the other categories)

$$\forall j\in J,\ \forall i\in J-\{j\},\ \mathrm{BSC}(\omega_i, j)=0 \tag{6.4}$$

を公理として要請していない事実に注意しておこう。

<u>Axiom 3</u>（**大分類関数 BSC の満たすべき公理**）

(i)(**カテゴリ抽出能力**；category separability)

$$\forall j\in J,\ \mathrm{BSC}(\omega_j, j)=1.$$

(ii)(写像 T の下での不変性；invariance under mapping T)

$$\forall\varphi\in\Phi,\ \forall j\in J,\ \mathrm{BSC}(T\varphi, j)=\mathrm{BSC}(\varphi, j).$$ □

6.2 類似度関数SMを用いた大分類関数BSCの構成

axiom 3を満たす式(6.1)の大分類関数 BSC を構成しよう。

$$\mathrm{psn}(u)=1 \text{ if } u\geqq 0,\ =0 \text{ if } u<0 \tag{6.5}$$

と定義される1実変数 u の2値関数 psn を導入して、

$$\begin{aligned}
&\mathrm{BSC}(\varphi, j)\\
&\equiv \mathrm{psn}(W(j, 1)\cdot \mathrm{SM}(\varphi, \omega_j)\\
&+W(j, 2)\cdot \sum_{i\in J-\{j\}} \mathrm{SM}(\varphi, \omega_i)\\
&+W(j, 3)\cdot \mathrm{SM}(\varphi, \omega_j)\cdot [\sum_{i\in J-\{j\}} \mathrm{SM}(\varphi, \omega_i)]\\
&-h(j))
\end{aligned} \tag{6.6}$$

$$\begin{aligned}
&=\mathrm{psn}(\mathrm{SM}(\varphi, \omega_j)\cdot [W(j, 1)-h(j)]\\
&+[1-\mathrm{SM}(\varphi, \omega_j)]\cdot[W(j, 2)-h(j)]\\
&+W(j, 3)\cdot \mathrm{SM}(\varphi, \omega_j)\cdot[1-\mathrm{SM}(\varphi, \omega_j)])\\
&\quad\because\ \text{axiom 2の(ii)}
\end{aligned} \tag{6.7}$$

と定義される式(6.1)の関数 BSC について考えよう。

式(3.72)の2カテゴリ分類困難度 DOC$(\varphi, \underline{\mathfrak{C}}[j]; \mathrm{SM})$，式(4.2)の容易度 EOC$(\varphi, \underline{\mathfrak{C}}[j]; \mathrm{SM})$ を用いると、3式(3.78)，(4.5)，(4.7)が成立しているから、式(6.6)の BSC は、

$$\begin{aligned}
&\mathrm{BSC}(\varphi, j)\\
&=\mathrm{psn}(W(j, 1)\cdot[1+\mathrm{DOC}(\varphi, \underline{\mathfrak{C}}[j]; \mathrm{SM})]^{-1}\\
&+W(j, 2)\cdot[1+\mathrm{EOC}(\varphi, \underline{\mathfrak{C}}[j]; \mathrm{SM})]^{-1}\\
&+W(j, 3)\cdot[d\mathrm{SM}(\varphi, \omega_j)/d\mathrm{EOC}(\varphi, \underline{\mathfrak{C}}[j]; \mathrm{SM})]\\
&-h(j))
\end{aligned} \tag{6.8}$$

と再表現される。

このとき、次の定理6.1が成立し、axiom 3を満たす1つの大分類関数 BSC が構成された。

［定理6.1］（**大分類関数 BSC の、SM による構成定理**）

4つの実定数 $W(j,\ \ell)\ (\ell=1, 2, 3)$, $h(j)\ (j \in J)$ が、2条件

$$\forall j \in J,\ W(j, 1) \geqq h(j) \tag{6.9}$$

$$\forall j \in J,\ W(j, 2) < h(j) \tag{6.10}$$

を満たせば、式(6.6)のように定義された関数 BSC はaxiom 3を満たし、然も、式(6.4)のカテゴリ間の相互排除性を満たす。

（証明） 先ず、axiom 3の(i)が

$$\begin{aligned}
&\forall j \in J,\ BSC(\omega_j, j)\\
&= psn(W(j, 1) \cdot SM(\omega_j,\ \omega_j)\\
&+ W(j, 2) \cdot \sum_{i \in J-\{j\}} SM(\omega_j,\ \omega_i)\\
&+ W(j, 3) \cdot SM(\omega_j,\ \omega_j) \cdot [\sum_{i \in J-\{j\}} SM(\omega_j,\ \omega_i)]\\
&- h(j)) \quad \because \ \text{式}(6.6)\\
&= psn(W(j, 1) \cdot 1 + W(j, 2) \cdot 0 + W(j, 3) \cdot 1 \cdot 0\\
&- h(j)) \quad \because \ \text{axiom 2, (i)}\\
&= psn(W(j, 1) - h(j))\\
&= 1 \quad \because \ \text{式}(6.9)
\end{aligned} \tag{6.11}$$

と得られる。

次に、axiom 3の(ii)がaxiom 2,（iii)を考慮すれば成立していることがわかる。

最後に、式(6.4)のカテゴリ間の相互排除性については、

$$\begin{aligned}
&\forall j \in J,\ \forall i \in J - \{j\},\ BSC(\omega_i, j)\\
&= psn(W(j, 1) \cdot SM(\omega_i,\ \omega_j)\\
&+ W(j, 2) \cdot \ [1 - SM(\omega_i,\ \omega_j)]\\
&+ W(j, 3) \cdot SM(\omega_i,\ \omega_j) \cdot [1 - SM(\omega_i,\ \omega_j)] - h(j))\\
&\qquad \because \ \text{式}(6.6),\ \text{axiom 2の(ii)}\\
&= psn(W(j, 1) \cdot 0 + W(j, 2) \cdot 1 + W(j, 3) \cdot 0 \cdot 1\\
&- h(j)) \qquad \because \ \text{axiom 2の(i)}\\
&= psn(W(j, 2) - h(j))\\
&= 0 \quad \because \ \text{式}(6.10)
\end{aligned} \tag{6.12}$$

と示され、証明が終わる。 □

6.3 2次ニューラルネットによる大分類関数 BSC の構成

上述の定理6.1を一般化して、次の定理6.2が得られる。式(6.15)の psn 内の成分は2次ニューラルネットの形式［B3］である。

［定理6.2］（**大分類関数 BSC の、SM による構成定理**）

2条件

$$\forall j \in J,\ W(j, j) + W(j, j, j) \geqq h(j) \tag{6.13}$$

$$\begin{aligned}&\forall j \in J,\ \forall k \in J - \{j\},\ \ W(j, k)\\ &\qquad\qquad + W(j, k, k) < h(j)\end{aligned} \tag{6.14}$$

の下で、

$$\mathrm{BSC}(\varphi, j)$$
$$\equiv \mathrm{psn}(\sum_{i\in J} W(j, i)\cdot \mathrm{SM}(\varphi, \omega_i)$$
$$+\sum_{i_1\in J}\sum_{i_2\in J} W(j, i_1, i_2)\cdot \mathrm{SM}(\varphi, \omega_{i_1})\cdot$$
$$\mathrm{SM}(\varphi, \omega_{i_2}) - h(j)) \tag{6.15}$$

と定義された式(6.1)の関数 BSC はaxiom 3を満たし、然も、式(6.4)のカテゴリ間の相互排除性を満たす。

（証明） 先ず、axiom 3の(i)が

$$\forall j\in J, \mathrm{BSC}(\omega_j, j)$$
$$= \mathrm{psn}(W(j, j)\cdot \mathrm{SM}(\omega_j, \omega_j)$$
$$+ W(j, j, j)\cdot \mathrm{SM}(\omega_j, \omega_j)$$
$$\cdot \mathrm{SM}(\omega_j, \omega_j) - h(j))$$

∵ 式(6.15), axiom 2の(i)

$$= \mathrm{psn}(W(j, j)$$
$$+ W(j, j, j) - h(j))$$

∵ axiom 2, (i)

$$= 1 \quad \because \text{式}(6.13) \tag{6.11}$$

と得られる。

次に、axiom 3の(ii)がaxiom 2, (iii)を考慮すれば成立していることがわかる。

最後に、式(6.4)のカテゴリ間の相互排除性については、

$$\forall j\in J, \forall k\in J-\{j\}, \mathrm{BSC}(\omega_k, j)$$
$$= \mathrm{psn}(\sum_{i\in J} W(j, i)\cdot \mathrm{SM}(\omega_k, \omega_i)$$
$$+\sum_{i_1\in J}\sum_{i_2\in J} W(j, i_1, i_2)\cdot \mathrm{SM}(\omega_k, \omega_{i_1})\cdot$$
$$\mathrm{SM}(\omega_k, \omega_{i_2}) - h(j))$$

∵ 式(6.15), axiom 2の(ii)

$$= \mathrm{psn}(W(j, k) + W(j, k, k) - h(j))$$

∵ axiom 2の(i)

$$= 0 \quad \because \text{式}(6.14) \tag{6.12}$$

と示され、証明が終わる。 □

7. むすび

シグモイド関数

$$1/[1+\exp(x)]\ (-\infty < x < +\infty) \tag{7.1}$$

情報量密度関数

$$\log_e[x/(1-x)]\ (0 < x < 1) \tag{7.2}$$

エントロピー関数

$$-x\cdot\log_e x - (1-x)\cdot\log_e(1-x)\ (0 < x < 1) \tag{7.3}$$

の3者には次の簡単な7つの関係Ⅰ～Ⅶがあり、本研究において有効に利用された：

Ⅰ．シグモイド1和関係

$1/[1+\exp(x)]+1/[1+\exp(-x)]$
$=1(-\infty<x<+\infty).$

Ⅱ．シグモイド・情報量密度の逆関数関係

$y=1/[1+\exp(-x)]$
$\Leftrightarrow$
$x=\log_e[y/(1-y)]\ (-\infty<x<+\infty,\ 0<y<1).$

Ⅲ．シグモイド下界の評価関係

$1/[1+\exp(x)]-[1-\exp[x]]$
$=\exp[x]\cdot 1/[1+\exp(-x)]\geqq 0(-\infty<x<+\infty).$

Ⅳ．シグモイド上界の評価関係

$1-\exp[x]+\exp[2x]-1/[1+\exp(x)]$
$=\exp[2x]\cdot 1/\ [1+\exp(-x)]\geqq 0(-\infty<x<+\infty).$

Ⅴ．シグモイド微分・シグモイド積の関係

$d/dx【1/[1+\exp(-x)]】$
$=【1/[1+\exp(-x)]】\cdot【1/[1+\exp(x)]】(-\infty<x<+\infty).$

Ⅵ．エントロピー微分・情報量密度の関係

$d/dx【-x\cdot\log_e x-(1-x)\cdot\log_e(1-x)】$
$=-\log_e[x/(1-x)](0<x<1).$

Ⅶ．情報量密度微分の評価関係

$d/dx【\log_e[x/(1-x)】=1/x+1/[1-x]\ (0<x<1).$ □

パターン認識の働きはもともと、多数のパターン事例を知覚したという経験から獲得された感受性、感性が反映されているであろう

“処理の対象とする問題のパターン φ に関し、その帰属するカテゴリを探索することを伴う**帰納的推論**(inductive reasoning)”

でなされている。

補助定理3.3などからわかるように、一般に、式(3.46)の曖昧度 $H(\underline{\mathfrak{C}}[j]/T\varphi)$ が大きい値をとるパターン $\varphi\in\Phi$ ほど、多段階想起形認識の働きで、パターン $\varphi\in\Phi$ を第 $j\in J$ 番目のカテゴリ $\mathfrak{C}_j$ のパターンモデル $T\omega_j$ に変換するには数多くの処理を必要とする。よって、式(3.69)のように、$H(\underline{\mathfrak{C}}[j]/T\varphi)$ の密度(情報量密度)として、処理の対象とする問題のパターン $\varphi\in\Phi$ を多段階パターン変換を介しての認識(2カテゴリ分類)の困難度 $DOC(\varphi,\underline{\mathfrak{C}}[j];SM)$ を定義するのは自然である。にもかかわらず、この種の論が全くなされなかっのは、事後確率 $p(\mathfrak{C}_1[j]/T\varphi)$ を式(3.63)のように、類似度 $SM(\varphi,\omega_j)$ と設定することにこれまでのパターン情報処理研究者が気が付かなかったからである。

本研究では、式(3.72)のように、困難度 $DOC(\varphi,\underline{\mathfrak{C}}[j];SM)$ は、

$$DOC(\varphi,\underline{\mathfrak{C}}[j];SM)$$
$$=\log_e[\{1-p(\mathfrak{C}_1[j]/T\varphi)\}/p(\mathfrak{C}_1[j]/T\varphi)] \tag{7.4}$$

と決定された(2式(3.63), (3.64)を参照)。

式(3.52)の平均相互情報量 $AMI(\underline{\mathfrak{C}}[j];\Phi[t_1,t_2])$ からわかるように、各パターン $\varphi\in\Phi$ に付き曖昧度 $H(\underline{\mathfrak{C}}[j]/T\varphi)$ を小さくすれば、獲得される情報量としての解消される不確定さ AMI($\underline{\mathfrak{C}}$

$[j]$；$\Phi[t_1, t_2]$）が大きくなる。SS理論での不動点多段階想起形認識［B3]，[B4］は解消される不確定さ $AMI(\mathfrak{C}[j] ; \Phi[t_1, t_2])$ が大きくなるようなパターン処理法であるとの結論が本研究によって鮮明にされたといえよう。

パターンをある1つのカテゴリに分類する場面と異なり、パターンからその特徴を抽出する場面においては、パターン形状の複雑さをエントロピーとして計量化し、この種の複雑さによりその特徴抽出の困難度を評価することが重要であろう：

処理の対象とする問題のパターン $\varphi \in \Phi \subseteq \mathfrak{H} = L_2(M ; dm)$ に対応するパターンモデル $T\varphi \in \Phi$ を求め、約束 $0/0=0$ の下で、

（i）座標値 $x \in M$ での振幅 $(T\varphi)(x)$ の、下限値 $i(\varphi)$ からの差 $p_s(x)$

$$\equiv [(T\varphi)(x) - i(\varphi)]/[s(\varphi) - i(\varphi)] \geqq 0$$

（ii）上限値 $s(\varphi)$ からの、座標値 $x \in M$ での振幅 $(T\varphi)(x)$ の差

$p_i(x)$

$$\equiv [s(\varphi) - (T\varphi)(x)]/[s(\varphi) - i(\varphi)] \geqq 0$$

where

$$s(\varphi) \equiv \sup_{x \in M} (T\varphi)(x)$$

$$i(\varphi) \equiv \inf_{x \in M} (T\varphi)(x)$$

をすべての座標点 $x \in M$ にわたり求める。

このとき、確率条件

（iii）$\forall x \in M,\ p_s(x) + p_i(x) \in \{0, 1\}$

が成立しているから、約束 $0\log_e 0=0=0$ の下で、密度関数

$I(\varphi ; x)$

$$\equiv -p_s(x) \cdot \log_e p_s(x) - p_i(x) \cdot \log_e p_i(x)$$

を求め、その積分値

$$SI(T\varphi) \equiv \int_M dm(x) I(\varphi ; x) / \int_M dm(x)$$

が特徴抽出の困難度を表している。

この $SI(T\varphi)$ は有界実数値パターンモデル $(T\varphi)(x)$ の持つ形状の複雑さの程度を表す平均情報量であり、S.Suzukiによって提案されたものである［B7］。 □

最大類似度認識法を用いて2式(4.19)，(4.20)のように入力パターン $\varphi \in \Phi$ を認識した場合、単段階パターン変換

$\varphi \rightarrow \omega_j$, 或いは、$T\varphi \rightarrow \omega_j$

を行っていると言える。最大類似度認識法のこのような単段階パターン変換とは異なり、多段階パターン変換過程を式(4.44)のように導入するのは、単段階パターン変換で φ の帰属するカテゴリを1度に決定したら犯すであろう“認識の誤びゅう”を出来るだけ回避するために、φ の帰属するカテゴリを出来るだけ曖昧なまま処理して行き、最終的に正当なカテゴリを獲得する余地を残存させようとする意図がある。この意図は、ファジィ情報処理(fuzzy information processing)［A6］の考えと一脈通ずるものがある。ファジィ情報処理の目的は何らかの決定をする場合最終段階まで曖昧ななまま処理することにより生じる利点を利用することにあるからである。決定に役立つ情報に多義性(polymeanings)，不確実性(uncertainty)，不完全性(incompleteness)がある場合、その多義性を出来るだけ解消する方向に、その不確実性を出来るだけ減少させる方向に、その不完全

性を出来るだけ補う方向に、曖昧なまま情報処理が可能なのが、ファジィ情報処理であるからである。

式(3.69)のように定義され、式(3.71)のように決定された本研究の2カテゴリ分類困難度DOC(φ, $\underline{\mathfrak{C}}$[j]；SM)は最終的に入力パターンの帰属しないカテゴリの候補を排除し、帰属する正当なカテゴリを抽出する目的を持ち、各認識段階での複数の候補カテゴリを絞っていく機能、即ち、ファジィ(曖昧さ)を減少させてゆく機能のある多段階パターン認識法を評価する1つの手段を提供しているものである。

文　献　A

[A 1] 福村晃夫："情報理論(情報工学3)"，コロナ社，June 1970

[A 2] Solomon W.Gelomb："A new derivation of the entropy expressions", IRE Transactions on Information Theory, vol.IT-7, no.3, pp.166-167, July 1961

[A 3] 小野田崇，Gunnar Rätsch, Klaus R.Müller："2値分類問題におけるAdaBoostの漸近特性解析と改善"，人工知能学会誌，vol.15, no.2, pp.287-295, Mar.2000

[A 4] 有本卓："情報理論(共立数学講座22)"，共立出版，Feb.1976

[A 5] 青木利夫，高橋渉："集合・位相空間要論"，培風館，Sept.1979

[A 6] L.Zadeh："Fuzzy Sets", Information and Control, vol.8, pp.338-353, 1965

[A 7] 栗田多喜夫，麻生英樹，大津展之："数量化手法の背後確率による解釈について"，電子情報通信学会論文誌，vol.J70-D, no.4, pp.760-768, Apr.1987

[A 8] Ashok K.Agrawala："Learning with a probabilistic teacher", IEEE Trans. on information theory, vol.IT-16, no.4, pp.373-379, July 1970

[A 9] Vladimir N.Vanik："The nature of statistical learning theory", Springer-Verlag New York, Inc., 1995

[A10] ヨアブ・フロインド，ロバート・シャピリ(訳 安倍直樹)："ブースティング入門"，人工知能学会誌，vol.14, no.5, pp.771-780, Sept.1999

[A11] 末松伸朗,林朗："ブースティング法に発想を得た確率モデル学習アルゴリズム"，人工知能学会誌，vol.15, no.1, pp.129-136, Jan.2000

[A12] Robert E.Schapire, Yoav Freund, Peter Bartlett, Wee Sun Lee："Boosting the margin", Machine Learning:proceedings of the Fourteenth International Conference (ICML), Nashville, Tennessee, July 8-12, 1997/edited by Douglas H.Fisher, Jr., …Morgan Kaufmann.1997, pp.322-330, 1997

[A13] Luc Devroye,Laszlo Gyorfi,Gabor Lugosi："A Probabilistic Theory of Pattern Recognition", Springer-Verlag New York,Inc.,1996

文　献　B

[B 1] 鈴木昇一："認識工学"，柏書房，Feb.1975

[B 2] 鈴木昇一："ニューラルネットの新数理"，近代文芸社，Sept.1996

[B 3] 鈴木昇一：“パターン認識問題の数理的一般解決”，近代文芸社，June 1997
[B 4] 鈴木昇一：“認識知能情報論の新展開”，近代文芸社，Aug.1998
[B 5] 鈴木昇一：“測度的不変量検出形認識系の構成理論”，電子通信学会論文誌(D)，vol.55-D, no.8, pp.531-538, Aug.1972
[B 6] 鈴木昇一：“パターンのエントロピーモデル”，電子情報通信学会論文誌(D-Ⅱ)，vol.J77-D-Ⅱ, no.10, pp.2220-2238, Nov.1994
[B 7] 鈴木昇一：“新しい情報の測度とパターン情報処理”,情報研究(文教大学・情報学部)，no.13, pp.273-358, Dec.1992
[B 8] 鈴木昇一：“手書き漢字の側抑制効果的分解とその計算機シミュレーション”，情報処理学会誌，vol.15, no.12, pp.927-934, Dec.1974
[B 9] 鈴木昇一：“画像情報量とその手書き漢字への応用”，画像電子学会誌，vol.4, no.1, pp.4-12, Apr.1975
[B10] 鈴木昇一：“抽出された特徴による手書き漢字構造の再生”，情報処理学会誌，vol.18, no.11, pp.1115-1122, Nov.1977
[B11] 鈴木昇一，斉藤静昭，奥野治雄，太田芳雄：“画像の復元とその計算機シュミレーション”，工学院大学研究報告，no.39, pp.198-206, Jan.1976
[B12] 鈴木昇一：“回転群と画像の分解・強調・構造化再構成に関する計算機シュミレーション”，情報研究(文教大学情報学部)，vol.4, pp.36-56, Dec.1983
[B13] 鈴木昇一：“連想形記億器 MEMOTRON と日本語母音系列の再生に関する計算機シミュレーション”，情報研究(文教大学情報学部)，vol.7, pp.14-29, Dec.1986
[B14] 鈴木昇一：“収縮写像の構成用空間回路とその計算機シミュレーション”，情報研究(文教大学・情報学部)，vol.9, pp.17-29, Dec.1988
[B15] 鈴木昇一：“多変量解析に基づく大分類関数の決定とその計算機シミュレーション”，情報研究(文教大学・情報学部)，vol.10, pp.35-49, Dec.1989
[B16] 鈴木昇一：“帰属係数法に基づく類似度、帰属関係あいまい度、認識情報量の計算機シミュレーション”，情報研究(文教大学・情報学部)，vol.11, pp.51-68, Dec.1990
[B17] 鈴木昇一：“構造受精法と日本語単独母音の認識”，情報研究(文教大学・情報学部)，vol.18,pp.17-51, Dec.1998
[B18] 鈴木昇一，前田英明：“有声破裂音の代表パターンの学習的決定と、その計算機シミュレーション”，情報研究(文教大学・情報学部)，no.20,pp.77-95, Dec.1998
[B19] 鈴木昇一，前田英明：“変動エントロピーによる有声破裂音の順序付けと、その計算機シミュレーション”，情報研究(文教大学・情報学部)，no.21,pp.51-78,Mar.1999
[B20] 鈴木昇一：“平均顔を用いた顔画像の２値化、並びに、目・鼻・口の抽出と、その計算機シミュレーション”，情報研究(文教大学・情報学部)，no.22,pp.65-150,Dec.1999
[B21] 鈴木昇一：“界面エネルギーの減少に伴うモデル構成作用素の、顔画像処理に関する計算機シミュレーション”，情報研究(文教大学・情報学部)，no.23,pp.109-182, Mar.2000
[B22] 鈴木昇一：“風景画から知識を抽出し、解釈するシステムの、ファジィ推論ニューラルネットによる構成”，情報研究(文教大学・情報学部)，no.23,pp.183-265, Mar.2000
[B23] 鈴木昇一：“各個人の感性を反映した認識システムRECOGNITRON”，情報研究(文教大学・情報学部)，no.24, pp.185-257, Dec.2000

[B24] 鈴木昇一："プロダクション・システムとしてのファジィ・マルチメディア・コンピュータと、空間多重パターンファジィ推論系"，情報研究（文教大学・情報学部），no.24, pp.105-183, Dec.2000

付録A. 他のカテゴリの代表パターンモデルから定まる"axiom 2を満たす類似度関数 SM"の構成論

本付録Aでは、axiom 2を満たす類似度関数 SM が他のカテゴリに帰属するパターンのモデル集合から定まるように構成される。

A.1 類似度関数 SM の構成における諸前提条件

全カテゴリ集合

$$\underline{\mathfrak{C}} \equiv \{\mathfrak{C}_j \mid j \in J\} \tag{A1.1}$$

についての全代表パターン集合

$$\Omega \equiv \{\omega_j \mid j \in J\} \tag{A1.2}$$

は1次独立であり、同時に、

$$T \cdot \Omega \equiv \{T\omega_j \mid j \in J\} \tag{A1.3}$$

も1次独立であることを要請する。第$j \in J$番目のカテゴリ$\mathfrak{C}_j$の持つ諸性質を典型的に代表するパターンがω_jであり、そのモデルが$T\omega_j$である。

確率条件

$$\forall j \in J, 0 < p(\mathfrak{C}_j) < 1 \tag{A1.4}$$

$$\wedge \sum_{j \in J} p(\mathfrak{C}_j) = 1 \tag{A1.5}$$

を満たす"パターンω_jをその代表パターンとする第$j \in J$番目のカテゴリ$\mathfrak{C}_j$の生起確率$p(\mathfrak{C}_j)$"を導入しておく。

式(A1.2)のΩ内の代表パターン間の分離条件

$$\forall j \in J, \forall i \in J - \{j\}, \ \| \omega_i - \omega_j \| > 0 \tag{A1.6}$$

の下で（Ωの1次独立性）、2条件

$$\forall j \in J, \forall i \in J - \{j\}, \ \Psi_i \cap \Psi_j = \phi \ \text{(パターン集合間の非交差条件)} \tag{A1.7}$$

$$\forall j \in J, \ \omega_j \in \Psi_j \ \text{(代表パターン包含条件)} \tag{A1.8}$$

を満たすようにパターン集合Ψ_jの系

$$\Psi_j (\subset \Phi), \ j \in J \tag{A1.9}$$

を選定しておく。このとき、

$$\forall j \in J, \forall i \in J - \{j\}, \ \omega_i \overline{\in} \Psi_j \ \text{(他代表パターンの非包含条件)} \tag{A1.10}$$

が成立していることに注意しておく。

処理の対象とする問題のパターン$\varphi \in \Phi$のモデル$T\varphi$の集合

$$T \cdot \Phi \equiv \{T\varphi \mid \varphi \in \Phi\} \tag{A1.11}$$

と、パターン集合Ψ_jのモデル集合

$$T \cdot \Psi_j \equiv \{T\varphi \mid \varphi \in \Psi_j\} \tag{A1.12}$$

とを導入しておく。このとき、2式(A2.2)，(A2.3)が成立することを可能ならしめるため、式(A1.3)の$T \cdot \Omega$内の代表パターンモデル間の分離条件

$\forall j \in J,\ \forall i \in J-\{j\},\ \ \| T\omega_i - T\omega_j \| > 0$ (A1.13)

の下で($T\cdot\Omega$ の1次独立性)、2条件

$\forall j \in J,\ \forall i \in J-\{j\},\ \ T\cdot\Psi_i \cap T\cdot\Psi_j = \phi$ (モデル集合間の非交差条件) (A1.14)

$\forall j \in J,\ T\cdot\omega_j \in T\cdot\Psi_j$ (代表パターンモデル包含条件) (A1.15)

をも要請する。当然、

$\forall j \in J,\ \forall i \in J-\{j\},\ \ T\cdot\omega_i \overline{\in} T\cdot\Psi_j$ (他代表パターンモデルの非包含条件) (A1.16)

が成立している。

A.2 類似度関数 SM の構成、再帰的構成

本節では、処理の対象とする問題のパターン $\varphi \in \Phi$ について、他のカテゴリに帰属するパターンのモデルから相違していればいるほど、問題のカテゴリに帰属するパターンのモデルとの類似性の程度が大きくなるように、axiom 2を満たす類似度関数

$$SM: \Phi \times \Omega \rightarrow \{s \mid 0 \leqq s \leqq 1\} \tag{A2.1}$$

を得る構成手法が説明される。

先ず、次の定理A1は、他のカテゴリに帰属するパターンモデル集合 $\Psi_{J-\{j\}}$ から、第 $j \in J$ 番目の類似度関数 $SM(\cdot,\ \omega_j)$ が構成され得ることを指摘している。

[定理A1] **(他カテゴリからの SM 構成基本定理)**

2条件

$$\text{(a)}\ \forall i \in J,\ \min_{\eta \in \Psi_i} dsm_i(T\omega_i,\ T\eta) = 0 \tag{A2.2}$$

$$\text{(b)}\ \forall j \in J,\ \forall i \in J-\{j\},\ \min_{\eta \in \Psi_i} dsm_i(T\omega_j,\ T\eta) > 0 \tag{A2.3}$$

を満たす相違度関数(dissimilarity-measure function) dsm_j の系

$$dsm_j: T\cdot\Phi \times T\cdot\Psi_j \rightarrow R^+\ (\text{非負実数全体の集合}),\quad j \in J \tag{A2.4}$$

が構成されたとしよう。このとき、

$$SM(\varphi,\ \omega_j) = \begin{cases} \prod_{i \in J-\{j\}} \min_{\eta \in \Psi_i} dsm_i(T\varphi,\ T\eta) \\ \quad / \sum_{k \in J} \prod_{i \in J-\{k\}} \min_{\eta \in \Psi_i} dsm_i(T\varphi,\ T\eta) \\ \quad \text{if } \sum_{k \in J} \prod_{i \in J-\{k\}} \min_{\eta \in \Psi_i} dsm_i(T\varphi,\ T\eta) > 0 \\ p(\mathfrak{C}_j) \\ \quad \text{if } \sum_{k \in J} \prod_{i \in J-\{k\}} \min_{\eta \in \Psi_i} dsm_i(T\varphi,\ T\eta) = 0 \end{cases} \tag{A2.5}$$

と定義された式(Y2.1)の関数 SM はaxiom 2を満たす。

(証明) axiom 2, (i)(直交性)が成立することを示そう。

2式(A2.2), (A2.3)に注意すれば、

①$\varphi = \omega_j$ のとき

$$\begin{aligned} SM(\varphi,\ \omega_j) &= \prod_{i \in J-\{j\}} \min_{\eta \in \Psi_i} dsm_i(T\omega_j,\ T\eta) \\ &\quad / \sum_{k \in J} \prod_{i \in J-\{k\}} \min_{\eta \in \Psi_i} dsm_i(T\omega_j,\ T\eta) \\ &= \prod_{i \in J-\{j\}} \min_{\eta \in \Psi_i} dsm_i(T\omega_j,\ T\eta) \\ &\quad / \prod_{i \in J-\{j\}} \min_{\eta \in \Psi_i} dsm_i(T\omega_j,\ T\eta) \end{aligned}$$

$\because$ 2式(A2.3), (A2.2)

$= 1 \quad \because$ 式(A2.3) (A2.6)

であるし、

②$\varphi = \omega_\ell \ (\ell \neq j)$のとき

$$
\begin{aligned}
&\mathrm{SM}(\varphi, \omega_j) \\
&= \prod_{i \in J - \{j\}} \min_{\eta \in \Psi_i} \mathrm{dsm}_i(T\omega_\ell, T\eta) \\
&\quad / \sum_{k \in J} \prod_{i \in J - \{k\}} \min_{\eta \in \Psi_i} \mathrm{dsm}_i(T\omega_\ell, T\eta) \\
&= 0 / \prod_{i \in J - \{\ell\}} \min_{\eta \in \Psi_i} \mathrm{dsm}_i(T\omega_\ell, T\eta) \\
&\quad \because \text{2式(A2.2), (A2.3)} \\
&= 0 \quad \because \text{式(A2.3)} \qquad \text{(A2.7)}
\end{aligned}
$$

あることがわかる。

次に、axiom 2, (ii)(規格化条件)

$$\forall \varphi \in \Phi, \sum_{j \in J} \mathrm{SM}(\varphi, \omega_j) = 1 \qquad \text{(A2.8)}$$

が成立することは、SM を定義している式(A2.5)と、2式(A1.5), (A1.6)とから明らか。

最後に、axiom 2, (iii)(T-不変性)

$$\forall \varphi \in \Phi, \forall j \in J, \mathrm{SM}(T\varphi, \omega_j) = \mathrm{SM}(\varphi, \omega_j) \qquad \text{(A2.9)}$$

が成立することは、axiom 1, (iii)の後半から明らか。 □

次の2定理A2, A3は、axiom 2を満たす類似度関数 SM´からaxiom 2を満たす今1つの類似度関数 SM を構成可能な"再帰性"を指摘するものである。

[定理A2] (**他カテゴリからの SM 再帰構成定理1**)

特に、各 Ψ_j を

$$\forall j \in J, \Psi_j = \{\omega_j\} \qquad \text{(A2.10)}$$

と設定しているとき、axiom 2を満たし、且つ、T-不変性

$$\forall \varphi \in \Phi, \forall i \in J, \mathrm{SM}'(T\varphi, T\omega_i) = \mathrm{SM}'(T\varphi, \omega_i) \qquad \text{(A2.11)}$$

が成り立つような類似度関数

$$\mathrm{SM}' : \Phi \times \Omega \rightarrow \{s \mid 0 \leqq s \leqq 1\} \qquad \text{(A2.12)}$$

を利用して、

$$
\begin{aligned}
&\min_{\eta \in \Psi_i} \mathrm{dsm}_i(T\varphi, T\eta) \\
&= \mathrm{dsm}_i(T\varphi, T\omega_i) \quad \because \text{式(A2.10)} && \text{(A2.13)} \\
&\equiv 1 - \mathrm{SM}'(T\varphi, T\omega_i) && \text{(A2.14)} \\
&= 1 - \mathrm{SM}'(T\varphi, \omega_i) \quad \because \text{式(A2.11)} \\
&= 1 - \mathrm{SM}'(\varphi, \omega_i) \quad \because \text{axiom 2, (iii)} && \text{(A2.15)}
\end{aligned}
$$

とおくことができ、定理A1が適用できる。

(証明) 式(A2.2)の成立は、

$$
\begin{aligned}
&\forall i \in J, \min_{\eta \in \Psi_i} \mathrm{dsm}_i(T\omega_i, T\eta) \\
&= 1 - \mathrm{SM}'(\omega_i, \omega_i) \quad \because \text{式(A2.15)} \\
&= 1 - 1 \quad \because \text{axiom 2, (i)} \\
&= 0 && \text{(A2.16)}
\end{aligned}
$$

と確かめられ、式(A2.3)の成立は、

$\forall j \in J,\ \forall i \in J - \{j\},$

$$
\begin{aligned}
&\min_{\eta \in \Psi_i} \mathrm{dsm}_i(T\omega_j, T\eta) \\
&= 1 - SM'(\omega_j, \omega_i) \quad \because \ \text{式}(A2.15) \\
&= 1 - 0 \quad \because \ \text{axiom 2, (i)} \\
&= 1 > 0 \qquad (A2.17)
\end{aligned}
$$

と確かめられた。 □

［定理A3］(**他カテゴリからの SM 再帰構成定理2**)

式(A2.10)を設定しているとき、

$$\forall \varphi \in \Phi,\ \forall i \in J,\ SM'_i(T\varphi, T\omega_i) = SM'_i(T\varphi, \omega_i) \qquad (A2.18)$$

が成立するようなaxiom 2を満たす類似度関数 SM ′ i, i∈J の系

$$SM'_i : \Phi \times \Omega \rightarrow \{s \mid 0 \leqq s \leqq 1\},\ i \in J \qquad (A2.19)$$

を利用して、

$$
\begin{aligned}
&\min_{\eta \in \Psi_i} \mathrm{dsm}_i(T\varphi, T\eta) \\
&= \mathrm{dsm}_i(T\varphi, T\omega_i) \quad \because \ \text{式}(A2.10) \qquad (A2.20) \\
&\equiv 1 - SM'_i(T\varphi, T\omega_i),\ i \in J \\
&= 1 - SM'_i(T\varphi, \omega_i),\ i \in J \quad \because \ \text{式}(A2.18) \\
&= 1 - SM'_i(\varphi, \omega_i) \quad \because \ \text{axiom 2, (iii)} \qquad (A2.21)
\end{aligned}
$$

とおくことができ、定理A1が適用できる。

(証明) 式(A2.2)の成立は、

$$
\begin{aligned}
&\forall i \in J,\ \min_{\eta \in \Psi_i} \mathrm{dsm}_i(T\omega_i, T\eta) \\
&= 1 - SM'_i(\omega_i, \omega_i) \quad \because \ \text{式}(A2.21) \\
&= 1 - 1 \quad \because \ \text{axiom 2, (i)} \\
&= 0 \qquad (A2.22)
\end{aligned}
$$

と確かめられ、式(A2.3)の成立は、

$\forall j \in J,\ \forall i \in J - \{j\},$

$$
\begin{aligned}
&\min_{\eta \in \Psi_i} \mathrm{dsm}_i(T\omega_j, T\eta) \\
&= 1 - SM'_i(\omega_j, \omega_i) \quad \because \ \text{式}(A2.21) \\
&= 1 - 0 \quad \because \ \text{axiom 2, (i)} \\
&= 1 > 0 \qquad (A2.24)
\end{aligned}
$$

と確かめられた。 □

尚、axiom 2を満たす式(A2.12)の類似度関数 SM′について、式(A2.11)が成立するように構成することは容易である。これまで構成られている SM′はすべて、式(A2.11)を満たしている［B3］,［B4］。

A.3 2条件式(A2.2), (A2.3)を満たす相違度関数 dsm_i の構成に必要な性質と、SMの簡単化

定理A1で構成された式(A2.5)の類似度関数 SM を簡単化したり、次節の構成諸例を簡単に理解したりするのに役立つ3定理A4～A6が説明される。

［定理A4］(**2条件式 (A2.2), (A2.3) の満足定理**)

(c) (dsm_i の零条件)

$$【\forall i \in J,$$
$$[\| T\varphi - T\eta \| = 0 \Rightarrow dsm_i(T\varphi, T\eta) = 0]】 \tag{A3.1}$$
$$\Rightarrow$$
$$【\forall i \in J,$$
$$\| T\omega_i - T\eta \| = 0 \Rightarrow dsm_i(T\omega_i, T\eta) = 0】 \tag{A3.2}$$
(d) (dsm_i の正条件)
$$【\forall i \in J,$$
$$[\| T\varphi - T\eta \| > 0 \Rightarrow dsm_i(T\varphi, T\eta) > 0]】 \tag{A3.3}$$
$$\Rightarrow$$
$$【\forall j \in J, \forall i \in J - \{j\},$$
$$[\| T\omega_j - T\eta \| > 0 \Rightarrow dsm_i(T\omega_j, T\eta) > 0]】 \tag{A3.4}$$
の両者を満たす式(A2.4)の関数系 dsmi, $i \in J$ は2条件式(A2.2), (A2.3)を満たす。

(証明)　明らかに、式(A3.1)から式(A3.2)が従う。

式(A3.2)の成立から、条件式(A2.2)の成立を示そう。

明らかに、式(A1.15)より、
$$\forall i \in J,\ \eta = \omega_i \in \Psi_i \tag{A3.5}$$
と選ぶことができ、このとき、
$$\forall i \in J, \exists \eta \in \Psi_i,\ \| T\omega_i - T\eta \| = 0 \tag{A3.6}$$
が成り立ち、よって、
$$\forall i \in J, \exists \eta \in \Psi_i,\ dsmi(T\omega_i, T\eta) = 0 \quad \because \text{式(A3.2)} \tag{A3.7}$$
が成り立ち、よって、条件式(A2.2)が成立することがわかる。

また、明らかに、式(A3.3)から式(A3.4)が従う。

式(A3.4)の成立から、条件式(A2.3)の成立を示そう。

明らかに、式(A1.16)より、
$$\forall j \in J, \forall i \in J - \{j\},\ \eta = \omega_i (\in \Psi_i) \overline{\in} \Psi_j \tag{A3.8}$$
と選ぶことができ、このとき、
$$\forall j \in J, \forall i \in J - \{j\},$$
$$\exists \eta \in \Psi_i,\ \| T\omega_j - T\eta \| > 0 \quad \because \text{式(A1.13)} \tag{A3.9}$$
が成り立ち、よって、
$$\forall j \in J, \forall i \in J - \{j\},$$
$$\exists \eta \in \Psi_i,\ dsm_i(T\omega_j, T\eta) > 0 \quad \because \text{式(A3.4)} \tag{A3.10}$$
が成り立ち、よって、条件式(A2.3)が成立することがわかる。　□

尚、式(A1. 9)からわかるように、
$$\forall j \in J,\ \Psi_j \supset \{\omega_j\} \tag{A3.11}$$
であり、特に、
$$\exists i \in J,\ \Psi_i = \{\omega_i\} \tag{A3.12}$$
であってもかまわないし、この場合、次の定理A5が成り立つ。

［定理A5］（**単一元から成る Ψ_i についての相違度 dsm_i を用いたときの、類似度関数表現定理**）

処理の対象とする問題のパターン集合Φから抜き取られたサンプルパターン集合 Ψ_i の構成が式(A3.13)のように単一の元 ω_i のように設定されていれば、

$$【\Psi_i = \{\omega_i\} \tag{A3.13}$$

$$\Rightarrow$$

$$\min_{\eta\in\Psi_i} dsm_i(T\varphi, T\eta)$$

$$= dsm_i(T\varphi, T\omega_i)】 \tag{A3.14}$$

が成り立ち、定理Y1の式(A2.5)の SM は、簡単に、

$$\forall\varphi\in\Phi, SM(\varphi, \omega_j) =$$

$$\begin{cases} \prod_{i\in J-\{j\}} dsm_i(T\varphi, T\omega_i) \\ \quad / \sum_{k\in J} \prod_{i\in J-\{k\}} dsm_i(T\varphi, T\omega_i) \\ \quad \text{if } \sum_{k\in J} \prod_{i\in J-\{k\}} dsm_i(T\varphi, T\omega_i) > 0 \\ p(\mathfrak{C}_j) \\ \quad \text{if } \sum_{k\in J} \prod_{i\in J-\{k\}} dsm_i(T\varphi, T\omega_i) = 0 \end{cases} \tag{A3.15}$$

と表現される。

（証明）　明らか。 □

次の定理 Y6は、定理A1の式(A2.5)の SM が極めて簡単になる形式(A3.18)を指摘している。

［定理A6］（**類似度関数 SM の簡単化定理**）

正条件

$$\sum_{k\in J} \prod_{i\in J-\{k\}} \min_{\eta\in\Psi_i} dsm_i(T\varphi, T\eta) > 0 \tag{A3.16}$$

が満たされている処理対象とするパターン $\varphi\in\Phi$ については、定理A1の式(A2.5)の SM は、簡単に、

$$SM(\varphi, \omega_j) =$$

$$\prod_{i\in J-\{j\}} \min_{\eta\in\Psi_i} dsm_i(T\varphi, T\eta)$$

$$/ \sum_{k\in J} \prod_{i\in J-\{k\}} \min_{\eta\in\Psi_i} dsm_i(T\varphi, T\eta) \tag{A3.17}$$

と表現され、更に、処理の対象とする問題のパターン集合 Φ から抜き取られたサンプルパターン集合 Ψ_i の構成条件式(A3.13)が成立していれば、式(A3.14)が成り立ち、正条件式(A3.16)は

$$\sum_{k\in J} \prod_{i\in J-\{k\}} dsm_i(T\varphi, T\omega_i) > 0 \tag{A3.18}$$

となり、式(A3.18)が満たされている処理対象とするパターン $\varphi\in\Phi$ については、

$$SM(\varphi, \omega_j) =$$

$$\prod_{i\in J-\{j\}} dsm_i(T\varphi, T\omega_i)$$

$$/ \sum_{k\in J} \prod_{i\in J-\{k\}} dsm_i(T\varphi, T\omega_i) \tag{A3.19}$$

と表現される。

（証明）　前半の式(A3.17)は式(A2. 5)から明らかであり、後半の2式(A3.18)，(A3.19)は定理A5から明らかである。 □

ここで、2式(A3.16)，(A3.18)を次の定理A7で書き直しておこう。

［定理A7］（**正条件同値定理**）

式(A3.16)は、

$$\exists k\in J, \forall i\in J-\{k\},$$

$$\forall \eta \in \Psi_i,\ dsm_i(T\varphi, T\eta) > 0 \tag{A3.20}$$

と同値である。

また、式(A3.18)は、

$$\exists k \in J,\ \forall i \in J-\{k\},$$
$$dsmi(T\varphi, T\omega_i) > 0 \tag{A3.21}$$

と同値である。

(証明)　前半は

$$\sum_{k \in J} \prod_{i \in J-\{k\}} \min_{\eta \in \Psi_i} dsm_i(T\varphi, T\eta) > 0$$
$$\Leftrightarrow \exists k \in J,\ \prod_{i \in J-\{k\}} \min_{\eta \in \Psi_i} dsm_i(T\varphi, T\eta) > 0 \tag{A3.22}$$
$$\Leftrightarrow \exists k \in J,\ \forall i \in J-\{k\},$$
$$\min_{\eta \in \Psi_i} dsm_i(T\varphi, T\eta) > 0 \tag{A3.23}$$
$$\Leftrightarrow \exists k \in J,\ \forall i \in J-\{k\},$$
$$\forall \eta \in \Psi_i,\ dsm_i(T\varphi, T\eta) > 0$$

と、示された。後半は、2式(A3.13), (A3.14)から明らかである。　□

A.4　パターンモデル間のノルム距離, 内積(相関値)を利用した関数 dsm_j の構成諸例

本節では、以上の6定理A1～A6を適用可能にする式(A2.4)の相違度関数の系 dsmi, $i \in J$ の構成諸例が示される。具体的には、上述の定理A1の2条件式(A2. 2), (A2. 3)を満たす式(A2. 4)の関数 dsm_j の諸例がパターンモデル間のノルム距離、内積(相関値)を利用して、構成される。

[構成例1]

$$dsmi(T\varphi, T\eta)$$
$$= \| T\varphi - T\eta \| \ (\text{パターンモデル間のノルム距離}) \tag{A4.1}$$

を採用すると、2条件式(A3.1), (A3.3)を満たし、よって、定理A4より、2条件式(Y2.2), (Y2.3)を満たす。

式(A3.13)が成り立つように設定しておけば、

$$\min_{\eta \in \Psi_i} dsm_i(T\varphi, T\eta)$$
$$= dsm_i(T\varphi, T\omega_i)$$
$$= \| T\varphi - T\omega_i \| \tag{A4.2}$$

であり、この場合、式(A2.5)の SM は

$$SM(\varphi, \omega_j) =$$
$$\begin{cases} \prod_{i \in J-\{j\}} \| T\varphi - T\omega_i \| \\ \quad / \sum_{k \in J} \prod_{i \in J-\{k\}} \| T\varphi - T\omega_i \| \\ \quad \text{if } \sum_{k \in J} \prod_{i \in J-\{k\}} \| T\varphi - T\omega_i \| > 0 \\ p(\mathfrak{C}_j) \\ \quad \text{if } \sum_{k \in J} \prod_{i \in J-\{k\}} \| T\varphi - T\omega_i \| = 0 \end{cases} \tag{A4.3}$$

と表されるが、パターンモデル間の分離条件式(A1.13)を考慮すれば、

$$\sum_{k \in J} \prod_{i \in J-\{k\}} \| T\varphi - T\omega_i \| > 0 \tag{A4.4}$$
$$\Rightarrow$$
$$SM(\varphi, \omega_j) =$$

$$\prod_{i \in J-\{j\}} \| T\varphi - T\omega_i \| \Big/ \sum_{k \in J} \prod_{i \in J-\{k\}} \| T\varphi - T\omega_i \| \qquad \text{(A4.5)}$$

と、簡単に表現される。 □

[構成例2]

正条件

$$\forall i \in J,\ a_i > 0 \qquad \text{(A4.6)}$$

を満たす係数 a_i の組 $\{a_i\}_{i \in J}$ を導入して、

$$dsm_i(T\varphi, T\eta) = 1 - \exp[-a_i^{-1} \cdot \| T\varphi - T\eta \|] \qquad \text{(A4.7)}$$

を採用すると、2条件式(A3.1), (A3.3)を満たし、よって、定理A4より、2条件式(A2.2), (A2.3)を満たす。 □

[構成例3]

正条件式(A4.6)を満たす係数 a_i の組 $\{a_i\}_{i \in J}$ を導入して、

$$dsm_i(T\varphi, T\eta) = \log_e[1 + a_i^{-1} \cdot \| T\varphi - T\eta \|] \qquad \text{(A4.8)}$$

を採用すると、2条件式(A3.1), (A3.3)を満たし、よって、定理A4より、2条件式(A2.2), (A2.3)を満たす。 □

[構成例4]

正条件式(A4.6)を満たす係数 a_i の組 $\{a_i\}_{i \in J}$ を導入して、

$$dsm_i(T\varphi, T\eta) = a_i^{-1} \cdot \| T\varphi - T\eta \| / [1 + a_i^{-1} \cdot \| T\varphi - T\eta \|] \qquad \text{(A4.9)}$$

を採用すると、2条件式(A3.1), (A3.3)を満たし、よって、定理A4より、2条件式(A2.2), (A2.3)を満たす。 □

[構成例5]

2条件

$$f_i(0) = 0 \qquad \text{(A4.10)}$$

$$\forall x > 0,\ f_i(x) > 0 \qquad \text{(A4.11)}$$

を満たす関数 f_i の系

$$f_i : R^+ (\text{非負実数全体の集合}) \to R^+,\ i \in J \qquad \text{(A4.12)}$$

を導入する。また、正条件式(A4.6)を満たす係数 a_i の組 $\{a_i\}_{i \in J}$ を導入して、

$$dsm_i(T\varphi, T\eta) = f_i(a_i^{-1} \cdot \| T\varphi - T\eta \|) \qquad \text{(A4.13)}$$

を採用すると、2条件式(A3.1), (A3.3)を満たし、よって、定理A4より、2条件式(A2.2), (A2.3)を満たす。 □

[構成例6]

式(A3.13)が成立しているとする。

2条件式(A4.10), (A4.11)を満たす関数の式(A4.12)の系 f_i, $i \in J$ を導入する。また、正条件式(A4.6)を満たす係数 a_i の組 $\{a_i\}_{i \in J}$ を導入して、

$$dsm_i(T\varphi, T\omega_i) = f(a_i^{-1} \cdot \| T\varphi - T\omega_i \| \Big/ \sum_{\ell \in J} a_\ell^{-1} \cdot \| T\varphi - T\omega_\ell \|) \qquad \text{(A4.14)}$$

を採用すると、2条件式(A3.1)，(A3.3)を満たし、よって、定理A4より、2条件式(A2.2)，(A2.3)を満たす。 □

［構成例7］

式(A3.13)が成立しているとする。

2条件式(A4.10)，(A4.11)を満たす関数の式(A4.12)の系 f_i, $i \in J$ を導入する。また、正条件式(A4.6)を満たす係数 a_i の組 $\{a_i\}_{i \in J}$ を導入して、

$$\begin{aligned} &dsm_i(T\varphi, T\omega_i) \\ &= f(a_\ell^{-1} \cdot \| T\varphi - T\omega_i \| \\ &\quad / \max_{\ell \in J} a_\ell^{-1} \cdot \| T\varphi - T\omega_\ell \|) \end{aligned} \tag{A4.15}$$

を採用すると、2条件式(A3.1)，(A3.3)を満たし、よって、定理A4より、2条件式(A2.2)，(A2.3)を満たす。 □

次の定理A8は、パターンモデル $T\varphi$, $T\eta$ 間のノルム距離 $\| T\varphi - T\eta \|$ のみならず、内積(相関値) $(T\varphi, T\eta)$ を利用して、式(A2.4)の相違度関数の系 dsm_i, $j \in J$ を構成できることを指摘している。

［定理A8］**(パターンモデル間の内積を利用しての、SM の構成例定理)**

以上の［構成例1～7］において、$\| T\varphi - T\eta \|$, $\| T\varphi - T\omega_i \|$ の代りに、次の諸例を考えても、式(A2.4)の相違度関数の系 dsm_i, $j \in J$ は2条件式(A3.1)，(A3.3)を満たし、よって、定理A4より、2条件式(A2.2)，(A2.3)を満たす：

$$[1 - |(T\varphi \| T\varphi \|^{-1}, T\eta \| T\eta \|^{-1})|^2]^{1/2} \tag{A4.16}$$

$$[1 - |(T\varphi \| T\varphi \|^{-1}, T\omega_i \| T\omega_i \|^{-1})|^2]^{1/2} \tag{A4.17}$$

$$[\| T\varphi \|^2 \cdot \| T\eta \|^2 - |(T\varphi, T\eta)|^2] \tag{A4.18}$$

$$[\| T\varphi \|^2 \cdot \| T\omega_i \|^2 - |(T\varphi, T\omega_i)|^2] \tag{A4.19}$$

(証明)　Schwarzの不等式

$$\forall \varphi \in \mathfrak{H}, \forall \eta \in \mathfrak{H}, |(\varphi, \eta)| \leqq \| \varphi \| \cdot \| \eta \| \tag{A4.20}$$

ここに、

$$|(\varphi, \eta)| = \| \varphi \| \cdot \| \eta \| \tag{A4.21}$$

$$\begin{aligned} &\Leftrightarrow \\ &\text{【}[\| \varphi \| = 0 \vee \| \eta \| = 0] \\ &\vee\ [\exists a (\neq 0) \in Z(\text{複素数全体の集合}), \\ &\quad \varphi = a \cdot \eta \neq 0]\text{】} \end{aligned} \tag{A4.22}$$

から、式(A2.4)の相違度関数の系 dsm_i, $j \in J$ は2条件式(A3.1)，(A3.3)を満たすことが知れる。

□

A.5　1次従属係数に基づいた dsm_i の構成諸例

本節では、1次従属係数 $d_\eta(\varphi)$ $(\eta \in \Psi_i, i \in J)$ を利用して、2条件式(A2.2)，(A2.3)を満たす式(A2.4)の相違度関数の系 dsm_i, $i \in J$ の諸例を構成する。

A.5.1　1次従属係数 $d_\eta(\varphi)$，$\eta \in \Psi_i (i \in J)$ の定義

先ず、処理の対象とする問題のパターン $\varphi \in \Phi$ の第 $\eta \in \Psi_i (i \in J)$ 番目の1次従属係数と呼ばれる $d_\eta(\varphi)$，$\eta \in \Psi_i (i \in J)$ を定義しよう。

任意にカテゴリ番号 $j \in J$ を選び、固定する。パターンモデル $T\varphi$ を各 $T\eta$ $(\eta \in \Psi_j)$ の1次結合

$$\sum_{\eta \in \Psi_j} d_\eta \cdot T\eta \tag{A5.1}$$

で近似するときの、誤差

$$T\varphi - \sum_{\eta \in \Psi_j} d_\eta \cdot T\eta \tag{A5.2}$$

の平均自乗ノルム

$$\| T\varphi - \sum_{\eta \in \Psi_j} d_\eta \cdot T\eta \|^2 \tag{A5.3}$$

を極小ならしめ、$T \cdot \Psi_j$ への1次従属性を表す1次結合係数 d_η は、最小自乗法を適用して、連立1次方程式

$$\sum_{\eta \in \Psi_j} (T\eta, T\eta') \cdot d_\eta(\varphi) = (T\varphi, T\eta'),$$
$$\eta' \in \Psi_j \tag{A5.4}$$

を解いて得られる。得られた各 d_η は $d_\eta(\varphi)$ と表現されている。

明らかに、3性質

(イ)(規格化直交性)

$$d_\eta(\varphi) = \begin{cases} 1 \text{ if } \| \varphi - \eta \| = 0 \\ 0 \text{ if } \| \varphi - \eta \| > 0 \end{cases} \tag{A5.5}$$

(ロ)(T-不変性1)

$$\forall \varphi \in \Phi, \forall \eta \in \Psi_j, d_{T\eta}(\varphi) = d_\eta(\varphi)$$
$$\because \text{ axiom 1, (iii)の後半 } T \cdot T = T \tag{A5.6}$$

(ロ)(T-不変性2)

$$\forall \varphi \in \Phi, \forall \eta \in \Psi_j, d_{T\eta}(\varphi) = d_\eta(\varphi)$$
$$\because \text{ axiom 1, (iii)の後半 } T \cdot T = T \tag{A5.7}$$

が成り立つ。

このとき、$T\varphi$ の、$T \cdot \Psi_j$ による1次結合式

$$\forall j \in J, \exists (T\varphi)_{j,\perp} \in \mathfrak{H},$$
$$T\varphi = \sum_{\eta \in \Psi_j} d_\eta(\varphi) \cdot T\eta + (T\varphi)_{j,\perp} \tag{A5.8}$$
$$\wedge [\forall \eta \in \Psi_j, ((T\varphi)_{j,\perp}, T\eta) = 0] \tag{A5.9}$$

が成り立つ。第 $\eta \in \Psi_j$ 番目の1次結合係数 $d_\eta(\varphi)$ の絶対値の自乗

$$| d_\eta(\varphi) |^2 \tag{A5.10}$$

は、パターンモデル $T\varphi \in T \cdot \Phi$ が今1つのパターンモデル $T\eta \in T \cdot \Psi_j$ と1次従属の関係にある強さの程度を反映したものである。

ここで、2条件

(一)(一致条件)$\forall i \in J$,

$$\exists \eta \in \Psi_i,$$
$$\forall \eta' \in \Psi_i, | d_{\eta'}(T\omega_i) |^2 = | d_{\eta'}(T\eta) |^2 \tag{A5.11}$$

(二)(非一致条件)$\forall j \in J, \forall i \in J - \{j\}$,

$$\exists \eta \in \Psi_i,$$
$$\exists \eta' \in \Psi_i,$$
$$| d_{\eta'}(T\omega_j) |^2 \neq | d_{\eta'}(T\eta) |^2 \tag{A5.12}$$

を要請しておく。

3条件式(A1.14)～(A1.16)を要請していることに注意しよう。一致条件式(A5.11)は、条件式(A1.15)を設定していることより、$\eta=\omega_i$ と選ぶことができ、自動的に満たされている。更に、条件式(A1.16)を設定していることより、$\eta=\eta'=\omega_i$ と選ぶことができ、非一致条件式(A5.12)を満たすような式(A1.9)のパターン集合の系 Ψ_j, $j\in J$ が選定されている。

A.5.2 相違度関数の系 dsmi, i∈J の構成諸例

2条件式(A5.11), (A5.12)を満たす式(A1.9)のパターン集合の系 Ψ^j, $j\in J$ を選定しているとして、式(A2.4)の相違度関数の系 dsm_i, $i\in J$ の諸例を構成する。

[構成例1]

正条件式

$$\forall i\in J,\ \forall \eta'\in\Psi_i,\ a_i(\eta')>0 \tag{A5.13}$$

を満たす係数 $a_i(\eta')$ の組 $\{a_i(\eta')\}_{\eta'\in\Psi_i}$ を導入して、

$$dsm_i(T\varphi,T\eta)=\Big[\sum_{\eta'\in\Psi_i}a_i(\eta')\cdot|d_{\eta'}(T\varphi)-d_{\eta'}(T\eta)|^2\Big]^{1/2} \tag{A5.14}$$

$$dsm_i(T\varphi,T\eta)=\max_{\eta'\in\Psi_i}a_i(\eta')\cdot|d_{\eta'}(T\varphi)-d_{\eta'}(T\eta)| \tag{A5.15}$$

$$dsm_i(T\varphi,T\eta)=\sum_{\eta'\in\Psi_i}a_i(\eta')\cdot|d_{\eta'}(T\varphi)-d_{\eta'}(T\eta)| \tag{A5.16}$$

などを採用すると、2条件式(A5.11), (A5.12)を満たす式(Y1.9)のパターン集合の系 Ψ_j, $j\in J$ を選定していれば、2条件式(A2.2), (A2.3)が満たされ、定理A1が適用できる。 □

[構成例2]

2条件式(A4.10), (A4.11)を満たす関数の式(A4.12)の系 f_i, $i\in J$ を導入する。例えば、正条件

$$\forall i\in J,\ w_i>0 \tag{A5.17}$$

を満たす係数 w_i の組 $\{w_i\}$ $i\in J$ を導入して、

$$f_i(x)=w_i\cdot x \tag{A5.18}$$

$$f_i(x)=1-\exp[-w_i\cdot x] \tag{A5.19}$$

$$f_i(x)=\log_e[1+w_i\cdot x] \tag{A5.20}$$

$$f_i(x)=w_i\cdot x/[1+w_i\cdot x] \tag{A5.21}$$

がそうである。また、正条件式(A5.13)を満たす係数 $a_i(\eta)$ の組 $\{a_i(\eta)\}_{\eta\in\Psi_i}$ を導入して、

$$dsm_i(T\varphi,T\eta)=f_i\Big(\Big[\sum_{\eta'\in\Psi_i}a_i(\eta')\cdot|d_{\eta'}(T\varphi)-d_{\eta'}(T\eta)|^2\Big]^{1/2}\Big) \tag{A5.22}$$

$$dsmi(T\varphi,T\eta)=f_i\Big(\Big[\max_{\eta'\in\Psi_i}a_i(\eta')\cdot|d_{\eta'}(T\varphi)-d_{\eta'}T\eta)|^2\Big]^{1/2}\Big) \tag{A5.23}$$

$$dsmi(T\varphi,T\eta)=f_i\Big(\Big[\sum_{\eta'\in\Psi_i}a_i(\eta')\cdot|d_{\eta'}(T\varphi)-d_{\eta'}(T\eta)|\Big]\Big) \tag{A5.24}$$

などを採用すると、2条件式(A5.11), (A5.12)を満たす式(A1.9)のパターン集合の系 Ψ_j, $j\in J$ を選

定していれば、2条件式(A2.2), (A2.3)が満たされ、定理A1が適用できる。 □

3式(A5.5)～(A5.7)を考慮すれば、次の構成例3の成立が直ちにわかる。

[構成例3]

$$\mathrm{dsm_i}(T\varphi, T\eta) = \max_{\eta' \in \Psi_i} |d_{T\eta'}(T\varphi)|^2 - |d_{T\eta}(T\varphi)|^2 \tag{A5.25}$$

$$\mathrm{dsm_i}(T\varphi, T\eta) = 1 - |d_{T\eta}(T\varphi)|^2 / \max_{\eta' \in \Psi_i} |d_{T\eta'}(T\varphi)|^2 \tag{A5.26}$$

$$\mathrm{dsm_i}(T\varphi, T\eta) = 1 - |d_{T\eta}(T\varphi)|^2 / \sum_{\eta' \in \Psi_i} |d_{T\eta}(T\varphi)|^2 \tag{A5.27}$$

などを採用すると、2条件式(A5.11), (A5.12)を満たす式(A1.9)のパターン集合の系 $\Psi_j, j\in J$ を選定していれば、2条件式(A2.2), (A2.3)が満たされ、定理A1が適用できる。 □

[構成例4]

2条件式(A4.10), (A4.11)を満たす関数の式(A4.12)の系 $f_i, i\in J$ を導入する。例えば、正条件(A5.16)を満たす係数 w_i の組 $\{w_i\}_{i\in J}$ を導入して、4式(A5.17)～(A5.20)がそうである。また、正条件式(A5.21)を満たす係数 $a_i(\eta)$ の組 $\{a_i(\eta)\}_{\eta\in\Psi_i}$ を導入して、

$$\mathrm{dsm_i}(T\varphi, T\eta) = f_i\left(\max_{\eta' \in \Psi_i} |d_{T\eta'}(T\varphi)|^2 - |d_{T\eta}(T\varphi)|^2\right) \tag{A5.28}$$

$$\mathrm{dsm_i}(T\varphi, T\eta) = f_i\left(1 - |d_{T\eta}(T\varphi)|^2 / \max_{\eta' \in \Psi_i} |d_{T\eta'}(T\varphi)|^2\right) \tag{A5.29}$$

$$\mathrm{dsmi}(T\varphi, T\eta) = f_i\left(1 - |d_{T\eta}(T\varphi)|^2 / \sum_{\eta' \in \Psi_i} |d_{T\eta}(T\varphi)|^2\right) \tag{A5.30}$$

などを採用すると、2条件式(A5.11), (A5.12)を満たす式(A1.9)のパターン集合の系 $\Psi_j, j\in J$ を選定していれば、2条件式(A2.2), (A2.3)が満たされ、定理A1が適用できる。 □

A.6 特徴抽出写像に基づいた $\mathrm{dsm_i}$ の構成諸例

本節では、特徴抽出写像

$$u : \Phi \times L \rightarrow Z\text{(複素数全体の集合)} \tag{A6.1}$$

を使って、式(A2.4)の相違度関数の系 $\mathrm{dsm_i}, i\in J$ の諸例を構成する。ここに、$u(\varphi, \ell)\in Z$ はパターン $\varphi\in\Phi$ から抽出された第 $\ell\in L$ 番目の特徴量である。

2条件

(一)(一致条件) $\forall i\in J$,

$$\exists \eta \in \Psi_i,\ \forall \ell \in L,\ u(T\omega_i, \ell) = u(T\eta, \ell) \tag{A6.2}$$

(二)(非一致条件) $\forall j\in J,\ \forall i\in J-\{j\}$,

$\exists \eta \in \Psi_i$,

$$\exists \ell \in L,\ u(T\omega_j, \ell) \neq u(T\eta, \ell) \tag{A6.3}$$

を要請しておく。

3条件式(A1.14)～(A1.16)を要請していることに注意しよう。一致条件式(A6. 2)は、条件式(A1.15)を設定していることより、$\eta=\omega_i$ と選ぶことができ、自動的に満たされている。更に、条件式(A1.16)を設定していることより、$\eta=\omega_i$ と選ぶことができるような式(A1.2)の Ω を選定す

るなど、非一致条件式(A6. 3)を満たすような式(A6.1)の特徴抽出写像 u を選定することは可能である。

[構成例1]

正条件式

$$\forall i \in J,\ \forall \ell \in L,\ a_i(\ell) > 0 \tag{A6.4}$$

を満たす係数 $a_i(\ell)$の組 $\{a_i(\ell)\}\ \ell \in L$ を導入して、

$$dsm_i(T\varphi, T\eta) = [\sum_{\ell \in L} a_i(\ell) \cdot |u(T\varphi, \ell) - u(T\eta, \ell)|^2]^{1/2} \tag{A6.5}$$

$$dsm_i(T\varphi, T\eta) = \max_{\ell \in J} a_i(\ell) \cdot |u(T\varphi, \ell) - u(T\eta, \ell)| \tag{A6.6}$$

$$dsm_i(T\varphi, T\eta) = \sum_{\ell \in L} a_i(\ell) \cdot |u(T\varphi, \ell) - u(T\eta, \ell)| \tag{A6.7}$$

などを採用すると、2条件式(A6.2), (A6.3)を満たす式(Y6.1)の特徴抽出写像 u を採用していれば、2条件式(A2.2), (A2.3)が満たされ、定理A1が適用できる。 □

[構成例2]

2条件式(A4.10), (A4.11)を満たす関数の式(A4.12)の系 f_i, $i \in J$ を導入する。例えば、正条件式(A5.17)を満たす係数 a_i の組 $\{a_i\}\ i \in J$ を導入して、4式(A5.18)～(A5.21)のように選定できる。

このとき、正条件式(A6.4)を満たす係数 $a_i(\ell)$の組 $\{a_i(\ell)\}_{\ell \in L}$ を導入して、

$$dsm_i(T\varphi, T\eta) = f_i([\sum_{\ell \in L} a_i(\ell) \cdot |u(T\varphi, \ell) - u(T\eta, \ell)|^2]^{1/2}) \tag{A6.8}$$

$$dsm_i(T\varphi, T\eta) = f_i(\max_{\ell \in L} a_i(\ell) \cdot |u(T\varphi, \ell) - u(T\eta, \ell)|) \tag{A6.9}$$

$$dsm_i(T\varphi, T\eta) = f_i(\sum_{\ell \in L} a_i(\ell) \cdot |u(T\varphi, \ell) - u(T\eta, \ell)|) \tag{A6.10}$$

などを採用すると、2条件式(A6.2), (A6.3)を満たす式(A6.1)の特徴抽出写像 u を採用していれば、2条件式(A2.2), (A2.3)が満たされ、定理A1が適用できる。 □

付録B. 類似度関数 SM を再帰的に構成する基本的な手法

本付録 B では、類似度関数 SM を再帰的に構成する基本的な手法が研究される。

axiom 2を満たす類似度関数 SM には、逆ノルム距離 $\| T\varphi - T\omega_j \|^{-1}$, 内積$(T\varphi, T\omega_j)$, 線形1次結合 $\sum_{j \in J} d_j(\varphi) \cdot T\omega_j$ に基づく基本的な3種類のものがある(文献 [B20] の付録Iを参照)。本付録B では、この事実を利用して、パターンモデル間の分離があからさまに改良されるように、axiom 2 を満たす類似度関数 SM′を SM に再帰的に構成し直す手法が説明される。

B.1 SM の再帰的構成における諸条件

以下では、axiom 2を満たす類似度関数

$$SM' : \Phi \times \Omega \rightarrow \{s \mid 0 \leqq s \leqq 1\} \tag{B1.1}$$

を構成し直して、新しい類似度関数

$$SM : \Phi \times \Omega \rightarrow \{s \mid 0 \leqq s \leqq 1\} \tag{B1.2}$$

を得る再帰的構成手法が説明される。

全カテゴリ集合

$$\underline{\mathfrak{C}} \equiv \{\mathfrak{C}_j \mid j \in J\} \tag{B1.3}$$

についての全代表パターン集合

$$\Omega \equiv \{\omega_j \mid j \in J\} \tag{B1.4}$$

は1次独立であり、同時に、

$$T \cdot \Omega \equiv \{T\omega_j \mid j \in J\} \tag{B1.5}$$

も1次独立であることを要請し、$T \cdot \Omega$ 間の任意の2要素 $T\omega_i, T\omega_j$ 間の分離性

$$\forall j \in J, \forall i \in J - \{j\}, \| T\omega_i - T\omega_j \| > 0 \tag{B1.6}$$

が成立していると仮定しておく。

確率条件

$$\forall j \in J, 0 < p(\mathfrak{C}_j) < 1 \tag{B1.7}$$

$$\wedge \sum_{j \in J} p(\mathfrak{C}_j) = 1 \tag{B1.8}$$

を満たす“パターン ω_j をその代表パターンとする第$j \in J$番目のカテゴリ $\mathfrak{C}_j$ の生起確率 $p(\mathfrak{C}_j)$”を導入しておく。

式(B1.4)の Ω 内の代表パターン間の分離条件

$$\forall j \in J, \forall i \in J - \{j\}, \quad \| \omega_i - \omega_j \| > 0 \tag{B1.9}$$

の下で、2条件

$$\forall j \in J, \forall i \in J - \{j\}, \quad \Psi_i \cap \Psi_j = \phi \tag{B1.10}$$

$$\forall j \in J, \quad \omega_j \in \Psi_j \tag{B1.11}$$

を満たすようにパターン集合 Ψ_j の系

$$\Psi_j (\subset \Phi), \quad j \in J \tag{B1.12}$$

を選定しておく。このとき、

$$\forall j \in J, \forall i \in J - \{j\}, \quad \omega_i \in \Psi_j \tag{B1.13}$$

が成立していることに注意しておく。

B.2 類似度関数 SM の一般的再帰的構成

次の定理B.1は、axiom 2を満たす類似度関数 SM を、次節以降で再帰的に構成するとき利用される。本定理B.1内の条件式(B2.1)では、f_j の値が $+\infty$ になってもよいことを許していることに注意しておく。

［定理B.1］ **(類似度関数 SM の一般的再帰的構成定理)**

2条件

$$\forall j \in J, +\infty \geqq \max_{\eta \in \Psi_j} f_j(T\omega_j, T\eta) > 0 \tag{B2.1}$$

$$\forall j \in J, \forall i \in J - \{j\},$$

$$\max_{\eta \in \Psi_j} f_j(T\omega_j, T\eta) \geqq 0 \tag{B2.2}$$

を満たす関数

$f_j : T\cdot\Phi \times T\cdot\Psi_j \rightarrow R^+$ (非負実数全体の集合) (B2.3)

と、axiom 2を満たす式(B1.1)の類似度関数 SM´との両者を用いて、

$$SM(\varphi, \omega_j)=$$
$$\begin{cases} SM'(\varphi, \omega_j)\cdot \max_{\eta\in\Psi_j} f_j(T\varphi, T\eta) \\ \quad / \sum_{i\in J} SM'(\varphi, \omega_i)\cdot \max_{\eta\in\Psi_i} f_i(T\varphi, T\eta) \\ \quad \text{if } \sum_{i\in J} SM'(\varphi, \omega_i)\cdot \max_{\eta\in\Psi_i} f_i(T\varphi, T\eta)>0 & \text{(B2.4)} \\ p(\mathfrak{C}_j) \\ \quad \text{if } \sum_{i\in J} SM'(\varphi, \omega_i)\cdot \max_{\eta\in\Psi_i} f_i(T\varphi, T\eta)=0 & \text{(B2.5)} \end{cases}$$

と定義される式(B1.2)の関数 SM はaxiom 2を満たす。

(証明) axiom 2, (i)の成立:

(a) $\varphi=\omega_j$のとき

$$SM(\varphi, \omega_j)$$
$$=[1+\sum_{i\in J-\{j\}} \{SM'(\varphi, \omega_i)\cdot \max_{\eta\in\Psi_i} f_i(T\varphi, T\eta)\}/\{SM'(\varphi, \omega_j)\cdot \max_{\eta\in\Psi_j} f_j(T\varphi, T\eta)\}]^{-1} \quad \text{(B2.6)}$$
$$=[1+\sum_{i\in J-\{j\}} \{SM'(\omega_j, \omega_i)\cdot \max_{\eta\in\Psi_i} f_i(T\omega_j, T\eta)\}/\{SM'(\omega_j, \omega_j)\cdot \max_{\eta\in\Psi_j} i_j(T\omega_j, T\eta)\}]^{-1} \quad \text{(B2.7)}$$
$$=[1+\sum_{i\in J-\{j\}} \{0\cdot \max_{\eta\in\Psi_i} f_i(T\omega_j, T\eta)\}/\{1\cdot \max_{\eta\in\Psi_j} f_j(T\omega_j, T\eta)\}]^{-1}$$

∵ SM´のaxiom 2の(i)の直交性

$=1/[1+0]$ ∵ 2式(B2.1), (B2.2)

$=1$ ∵ 式(B2.1) (B2.8)

(b) $\varphi=\omega_k (k\neq j)$のとき、

$$SM(\varphi, \omega_j)=$$
$$=SM'(\omega_k, \omega_j)\cdot \max_{\eta\in\Psi_j} f_j(T\omega_k, T\eta) / \sum_{i\in J} SM'(\omega_k, \omega_i)\cdot \max_{\eta\in\Psi_i} f_i(T\omega_k, T\eta) \quad \text{(B2.9)}$$
$$=0\cdot \max_{\eta\in\Psi_j} f_j(T\omega_k, T\eta) / 1\cdot \max_{\eta\in\Psi_k} f_k(T\omega_k, T\eta)$$

∵ SM´のaxiom 2の(i)の直交性

$=0$ ∵ 2式(B2.1), (B2.2) (B2.10)

axiom 2, (ii)の成立: 2定義式(B2.4), (B2.5)から明らかである。

axiom 2, (iii)の成立: SM´がaxiom 2, (iii)を満たすことと、axiom 1, (iii)の後半である $T\cdot T=T$ から明らかである。 □

B.3 内積形再帰的類似度(タイプ1のSS再帰的類似度)SM

$$T\varphi \,\|T\varphi\|^{-1}=0 \text{ if and only if } \|T\varphi\|=0 \quad \text{(B3.1)}$$

と約束する。

Schwarzの不等式から、不等式

$$\forall\varphi\in\Phi,\ \forall\eta\in\Phi,$$
$$|(T\varphi \,\|T\varphi\|^{-1}, T\eta \,\|T\eta\|^{-1})| \leqq 1 \quad \text{(B3.2)}$$

が成り立っている。更に、

$$(T\varphi\,\|T\varphi\|^{-1},\,T\eta\,\|T\eta\|^{-1})=\pm 1$$
$$\Leftrightarrow \exists c\in R^{++}(\text{正実数全体}),\ T\varphi=c\cdot T\eta\neq 0 \tag{B3.3}$$

も成り立っている。また、式(B3.1)の約束から、

$$(T\varphi\,\|T\varphi\|^{-1},\,T\eta\,\|T\eta\|^{-1})=0$$
$$\Leftrightarrow (T\varphi,\,T\eta)=0 \tag{B3.4}$$

であることも明らかである。

さて、処理の対象とする任意の問題のパターン $\varphi\in\Phi$ について、

$$(T\varphi,\,T\eta),\quad \eta\in\Psi_j,\,j\in J\ \text{は実数値の系である} \tag{B3.5}$$

とする。

式(B2.3)の関数 f_j として、

$$f_j(T\varphi,\,T\eta)$$
$$\equiv[1+(T\varphi\,\|T\varphi\|^{-1},\,T\eta\,\|T\eta\|^{-1})] \tag{B3.6}$$
$$\therefore$$
$$\max_{\eta\in\Psi_j} f_j(T\varphi,\,T\eta)$$
$$=[1+\max_{\eta\in\Psi_j}(T\varphi\,\|T\varphi\|^{-1},\,T\eta\,\|T\eta\|^{-1})] \tag{B3.7}$$

を採用すれば、

$$\text{(a)}\ \forall j\in J,\ \max_{\eta\in\Psi_j} f_j(T\omega_j,\,T\eta)=2>0$$
$$\because\ 2\text{式}(B1.11),\ (B3.3) \tag{B3.8}$$
$$\text{(b)}\ \forall j\in J,\ \forall i\in J-\{j\},$$
$$\max_{\eta\in\Psi_i} f_j(T\omega_j,\,T\eta)\geqq 0$$
$$\because\ 2\text{式}(B3.2)\,(B1.13) \tag{B3.9}$$

を得、2式(B2.1), (B2.2)を満たしており、定理B.1を適用できる。

式(B3.6)の関数 f_j 内の

$$(T\varphi\,\|T\varphi\|^{-1},\,T\eta\,\|T\eta\|^{-1}) \tag{B3.10}$$

は、2つのパターンモデル $T\varphi,T\eta$ 間の線形相関の程度を表している"$T\varphi$ が $T\eta$ に直交していない程度を表す規格化値"であることに注意しておこう。

更に、パターン

$$\psi\equiv[(T\varphi,\,T\eta)/(T\eta,\,T\eta)]\cdot T\eta \tag{B3.11}$$
$$=(T\varphi,\,T\eta\,\|T\eta\|^{-1})\cdot T\eta\,\|T\eta\|^{-1} \tag{B3.12}$$

に含まれているパターンモデル $T\varphi$ の程度を**情報量**(amount of information)$I(T\varphi,\ \psi)$として計量化した $I(T\varphi,\ \psi)$は、

$$I(T\varphi,\ \psi)=-(1/2)\cdot\log_e[1-\|\psi\|^2/\|T\varphi\|^2] \tag{B3.13}$$

と定義できる(文献［B20］の付録F, 文献［B4］の定理A 2.3を参照)、再表現

$$I(T\varphi,\ \psi)$$
$$=-(1/2)\cdot\log_e[1-|(T\varphi,\,T\eta\,\|T\eta\|^{-1})|^2$$
$$/\|T\varphi\|^2] \tag{B3.14}$$
$$=-(1/2)\cdot\log_e[1-$$
$$|(T\varphi\cdot\|T\varphi\|^{-1},\,T\eta\cdot\|T\eta\|^{-1})|^2 \tag{B3.15}$$

が得られることにも注意しておこう。

B.4　内積形再帰的類似度（タイプ2のSS再帰的類似度）SM

式(B2.3)の関数 f_j として、

$$f_j(T\varphi, T\eta) \equiv |(T\varphi \| T\varphi \|^{-1}, T\eta \| T\eta \|^{-1})|^2 \quad \text{(B4.1)}$$

を採用すれば、

(a) $\forall j \in J, \max_{\eta \in \Psi_j} f_j(T\omega_j, T\eta) = 1 > 0$

$\because$　2式(B1.11)，(B3.3)　(B4.2)

(b) $\forall j \in J, \forall i \in J - \{j\},$

$\max_{\eta \in \Psi_i} f_j(T\omega_j, T\eta) \geqq 0$

$\because$　2式(B3.2)，(B1.13)　(B4.3)

を得、2式(B2.1)，(B2.2)を満たしており、定理B.1を適用できる。

B.5　内積形再帰的類似度（タイプ3のSS再帰的類似度）SM

式(B2.3)の関数 f_j として、

$$f_j(T\varphi, T\eta) \equiv |(T\varphi, T\eta)|^2 \quad \text{(B5.1)}$$

$\therefore$

$\max_{\eta \in \Psi_j} f_j(T\varphi, T\eta)$

$\leqq \| T\varphi \|^2 \cdot \max_{\eta \in \Psi_j} \| T\eta \|^2$

$\because$　式(B3.2)　(B5.2)

を採用すれば、

(a) $\forall j \in J,$

$\max_{\eta \in \Psi_j} f_j(T\omega_j, T\eta) = \| T\omega_j \|^4 > 0$

$\because$　2式(B1.11)，(B3.3)　(B5.3)

(b) $\forall j \in J, \forall i \in J - \{j\},$

$\max_{\eta \in \Psi_i} f_j(T\omega_j, T\eta) \geqq 0$

$\because$　2式(B3.2)，(B1.13)　(B5.4)

を得、2式(B2.1)，(B2.2)を満たしており、定理B.1を適用できる。

B.6　逆自乗ノルム距離形再帰的類似度（タイプ4のSS再帰的類似度）SM

式(B2.3)の関数 fj として、

$$f_j(T\varphi, T\eta) \equiv \| T\varphi - T\eta \|^{-2} / \sum_{\eta \in \Psi_j} \| T\varphi - T\eta \|^{-2} \quad \text{(B6.1)}$$

を採用すれば、

(a) $\forall j \in J, \max_{\eta \in \Psi_j} f_j(T\omega_j, T\eta) = 1 > 0$

$\because$　式(B1.11)　(B6.2)

(b) $\forall j \in J, \forall i \in J - \{j\},$

$\max_{\eta \in \Psi_i} f_j(T\omega_j, T\eta) \geqq 0$

$\because$　式(B1.13)　(B6.3)

を得、2式(B2.1)，(B2.2)を満たしており、定理B.1を適用できる。

尚、処理の対象とする問題のパターンφの集合 Φ、並びに、サンプルパターン集合 Ψ_j が有限集合であると仮定して得られる汎関数

$$F(\{a_\eta \mid \eta \in \Psi_j, \varphi \in \Phi\} ; T\cdot\Phi, T\cdot\Psi_j)$$
$$= \sum_{\varphi \in \Phi} \sum_{\eta \in \Psi_j} a_\eta^{2} \cdot \| T\varphi - T\eta \|^{2} \tag{B6.4}$$
$$\text{on condition that } \sum_{j \in J} a_\eta = 1 \tag{B6.5}$$
$$\text{for any } \varphi \in \Phi \tag{B6.6}$$

は各$T\eta$の回りに$T\varphi$が集中している程度を計量化したもの(自乗ノルム基準)であり、この $F(\{a_\eta \mid \eta \in \Psi_j, \varphi \in \Phi\} ; T\cdot\Phi, T\cdot\Psi_j)$を極小にする各変数 a_η の値 $a_\eta(\varphi)$ は、fuzzy $|J|$-means algorithm によれば、式(B6.1)の $f_j(T\varphi, T\eta)$であって、

$$a_\eta(\varphi)$$
$$= \| T\varphi - T\eta \|^{-2} / \sum_{\eta \in \Psi j} \| T\varphi - T\eta \|^{-2}$$
$$\text{for any } \eta \in \Psi_j \text{ and } \varphi \in \Phi \tag{B6.7}$$

と与えられる。

B.7　逆自乗ノルム距離形再帰的類似度(タイプ5のSS再帰的類似度)SM

式(B2.3)の関数 f_j として、

$$f_j(T\varphi, T\eta)$$
$$\equiv \| T\varphi - T\eta \|^{-2} / \max_{\eta \in \Psi_j} \| T\varphi - T\eta \|^{-2} \tag{B7.1}$$

を採用すれば、

$$\text{(a) } \forall j \in J, \max_{\eta \in \Psi_j} f_j(T\omega_j, T\eta) = 1 > 0$$
$$\because \text{ 式(B1.11)} \tag{B7.2}$$
$$\text{(b) } \forall j \in J, \forall i \in J - \{j\},$$
$$\max_{\eta \in \Psi_i} f_j(T\omega_j, T\eta) \geqq 0$$
$$\because \text{ 式(B1.13)} \tag{B7.3}$$

を得、2式(B2.1), (B2.2)を満たしており、定理B.1を適用できる。

B.8　逆自乗ノルム距離形再帰的類似度(タイプ6のSS再帰的類似度)SM

式(B2.3)の関数 f_j として、

$$f_j(T\varphi, T\eta)$$
$$\equiv \| T\varphi - T\eta \|^{-2} \tag{B8.1}$$

を採用すれば、

$$\text{(a) } \forall j \in J, \max_{\eta \in \Psi_j} f_j(T\omega_j, T\eta) = +\infty$$
$$\because \text{ 式(B1.11)} \tag{B8.2}$$
$$\text{(b) } \forall j \in J, \forall i \in J - \{j\},$$
$$\max_{\eta \in \Psi_i} f_j(T\omega_j, T\eta) \geqq 0$$
$$\because \text{ 式(B1.13)} \tag{B8.3}$$

を得、2式(B2.1), (B2.2)を満たしており、定理B.1を適用できる。

B.9　自乗ノルム距離形再帰的類似度(タイプ7のSS再帰的類似度)SM

正条件

$\forall \eta \in \Psi_j, a_\eta > 0$ (B9.1)

を満たす係数 a_η の組 $\{a_\eta \mid \eta \in \Psi_j\}$ を導入して、式(B2.3)の関数 f_j として、

$$f_j(T\varphi, T\eta) \equiv \exp[-a_\eta^{-1} \cdot \| T\varphi - T\eta \|^2] / \sum_{\eta \in \Psi_j} \exp[-a_\eta^{-1} \cdot \| T\varphi - T\eta \|^2] \quad (B9.2)$$

を採用すれば、

(a) $\forall j \in J, \max_{\eta \in \Psi_j} f_j(T\omega_j, T\eta) > 0$

$\because$ 式(B1.11) (B9.3)

(b) $\forall j \in J, \forall i \in J - \{j\},$

$\max_{\eta \in \Psi_i} f_j(T\omega_j, T\eta) > 0$

$\because$ 式(B1.13) (B9.4)

を得、2式(B2.1), (B2.2)を満たしており、定理B.1を適用できる。

B.10 自乗ノルム距離形再帰的類似度(タイプ8のSS再帰的類似度)SM

正条件式(B9.1)を満たす係数 a_η の組 $\{a_\eta \mid \eta \in \Psi_j\}$ を導入して、式(B2.3)の関数 f_j として、

$$f_j(T\varphi, T\eta) \equiv \exp[-a_\eta^{-1} \cdot \| T\varphi - T\eta \|^2] / \max_{\eta \in \Psi_j} \exp[-a_\eta^{-1} \cdot \| T\varphi - T\eta \|^2] \quad (B10.1)$$

を採用すれば、

(a) $\forall j \in J, \max_{\eta \in \Psi_j} f_j(T\omega_j, T\eta) = 1 > 0$ (B10.2)

(b) $\forall j \in J, \forall i \in J - \{j\},$

$\max_{\eta \in \Psi_i} f_j(T\omega_j, T\eta) > 0$ (B10.3)

を得、2式(B2.1), (B2.2)を満たしており、定理B.1を適用できる。

B.11 自乗ノルム距離形再帰的類似度(タイプ9のSS再帰的類似度)SM

正条件式(B9.1)を満たす係数 a_η の組 $\{a_\eta \mid \eta \in \Psi_j\}$ を導入して、式(B2.3)の関数 f_j として、

$$f_j(T\varphi, T\eta) \equiv \exp[-a_\eta^{-1} \cdot \| T\varphi - T\eta \|^2] \quad (B11.1)$$

を採用すれば、

(a) $\forall j \in J, \max_{\eta \in \Psi_j} f_j(T\omega_j, T\eta) = 1 > 0$

$\because$ 式(B1.11) (B11.2)

(b) $\forall j \in J, \forall i \in J - \{j\},$

$\max_{\eta \in \Psi_i} f_j(T\omega_j, T\eta) > 0$ (B11.3)

を得、2式(B2.1), (B2.2)を満たしており、定理B.1を適用できる。

尚、以上の6節B.6～11の $f_j(T\varphi, T\eta)$ の定義式(B6.1), (B7.1), (B8.1), (B9.2), (B10.1), (B11.1)内での、$T\varphi, T\eta$ 間のノルム距離 $\| T\varphi - T\eta \|$ の代りに、2性質

$dsm(T\varphi, T\eta) > 0$ if $\| T\varphi - T\eta \| > 0$ (B11.4)

$dsm(T\varphi, T\eta) = 0$ if $\| T\varphi - T\eta \| = 0$ (B11.5)

を満たす2変数関数(相違度関数；dissimilarity-measure function)

$$\mathrm{dsm}: \mathrm{T}\cdot\Phi \times \mathrm{T}\cdot\Phi \to \mathbf{R}^{+} \tag{B11.6}$$

で定義される $\mathrm{dsm}(\mathrm{T}\varphi, \mathrm{T}\eta)$ を用いても、全く同様に、axiom 2を満たす式(B1.2)の類似度関数 SM が得られることが容易に理解できよう。

例えば、2性質(B11.4)，(B11.5)を満たす dsm として、

$$\mathrm{dsm}(\mathrm{T}\varphi, \mathrm{T}\eta) = [1 - |(\mathrm{T}\varphi \| \mathrm{T}\varphi \|^{-1}, \mathrm{T}\eta \| \mathrm{T}\eta \|^{-1})|^{2}]^{1/2} \tag{B11.7}$$

があるが、以上を定理B.2にまとめておこう。

[定理B.2] (**相違度関数を用いた類似度関数 SM の再帰構成定理**)

6節B.6～11の $\mathrm{f_j}(\mathrm{T}\varphi, \mathrm{T}\eta)$ の定義式(B6.1)，(B7.1)，(B8.1)，(B9.2)，(B10.1)，(B11.1)において、$\mathrm{T}\varphi$, $\mathrm{T}\eta$ 間のノルム距離 $\| \mathrm{T}\varphi - \mathrm{T}\eta \|$ の代りに、2性質(B11.4)，(B11.5)を満たす式(B11.6)の関数 dsm を用いて $\mathrm{dsm}(\mathrm{T}\varphi, \mathrm{T}\eta)$ を採用しても、axiom 2を満たす式(B1.2)の類似度関数 SM が得られる。 □

B.12 1次従属形再帰的類似度(タイプ10のS再帰的S類似度)SM

パターンモデル $\mathrm{T}\varphi$ を各 $\mathrm{T}\eta$ $(\eta \in \Psi_{\mathrm{j}})$ の1次結合

$$\sum_{\eta \in \Psi_{\mathrm{j}}} \mathrm{d}_{\eta} \cdot \mathrm{T}\eta \tag{B12.1}$$

で近似するときの、誤差

$$\mathrm{T}\varphi - \sum_{\eta \in \Psi_{\mathrm{j}}} \mathrm{d}_{\eta} \cdot \mathrm{T}\eta \tag{B12.2}$$

の平均自乗ノルム

$$\| \mathrm{T}\varphi - \sum_{\eta \in \Psi_{\mathrm{j}}} \mathrm{d}_{\eta} \cdot \mathrm{T}\eta \|^{2} \tag{B12.3}$$

を極小ならしめ、$\mathrm{T}\cdot\Psi_{\mathrm{j}}$ への1次従属性を表す1次結合係数 d_{η} は、連立1次方程式

$$\sum_{\eta \in \Psi_{\mathrm{j}}} (\mathrm{T}\eta, \mathrm{T}\eta') \cdot \mathrm{d}_{\eta}(\varphi) = (\mathrm{T}\varphi, \mathrm{T}\eta'), \quad \eta' \in \Psi_{\mathrm{j}} \tag{B12.4}$$

を解いて得られる。得られた各 $\mathrm{d_j}$ は $\mathrm{d_j}(\varphi)$ と表現されている。

明らかに、2性質

(イ)(規格化直交性)

$$\mathrm{d}_{\eta}(\mathrm{T}\eta') = \begin{cases} 1 \text{ if } \| \eta - \eta' \| = 0 \\ 0 \text{ if } \| \eta - \eta' \| > 0 \end{cases} \tag{B12.5}$$

(ロ)(T-不変性)

$$\forall \varphi \in \Phi, \forall \eta \in \Psi_{\mathrm{j}}, \mathrm{d}_{\eta}(\mathrm{T}\varphi) = \mathrm{d}_{\eta}(\varphi) \quad \because \text{ axiom 1, (iii)の後半 } \mathrm{T}\cdot\mathrm{T} = \mathrm{T} \tag{B12.6}$$

が成り立つ。

このとき、$\mathrm{T}\varphi$ の、$\mathrm{T}\cdot\Psi_{\mathrm{j}}$ による1次結合式

$$\exists (\mathrm{T}\varphi)_{\perp} \in \mathfrak{H},$$

$$\mathrm{T}\varphi = \sum_{\eta \in \Psi_{\mathrm{j}}} \mathrm{d}_{\eta}(\varphi) \cdot \mathrm{T}\eta + (\mathrm{T}\varphi)_{\perp} \tag{B12.7}$$

$$\wedge [\forall \eta \in \Psi_{\mathrm{j}}, ((\mathrm{T}\varphi)_{\perp}, \mathrm{T}\eta) = 0] \tag{B12.8}$$

が成り立つ。第 $\eta \in \Psi_{\mathrm{j}}$ 番目の1次結合係数 $\mathrm{d}_{\eta}(\varphi)$ の絶対値の自乗

$$|d_\eta(\varphi)|^2 \tag{B12.9}$$

は、$T\varphi$ が $T\eta$ と1次従属の関係にある強さの程度を反映したものである。

式(B2.3)の関数 f_j として、

$$f_j(T\varphi, T\eta) \equiv |d_\eta(\varphi)|^2 / \sum_{\eta\in\Psi_j} |d_\eta(\varphi)|^2 \tag{B12.10}$$

を採用すれば、

$$\begin{aligned}&(a)\ \forall j\in J,\ \max_{\eta\in\Psi_j} f_j(T\omega_j, T\eta)=1>0\\ &\qquad \because\ 2式(B1.11),\ (B12.5)\end{aligned} \tag{B12.11}$$

$$\begin{aligned}&(b)\ \forall j\in J,\ \forall i\in J-\{j\},\\ &\qquad \max_{\eta\in\Psi_i} f_j(T\omega_j, T\eta)>0\\ &\qquad \because\ 3式(B1.13),\ (B12.7),\ (B12.8)\end{aligned} \tag{B12.12}$$

を得、2式(B2.1),(B2.2)を満たしており、定理B.1を適用でき、$T\varphi$ が各 $T\omega_j$ と1次従属の関係にある程度を計量化した1次従属形類似度と称される式(B1.2)の SM が得られたことになる。

尚、例えば、2性質(B11.4),(B11.5)を満たす式(11.6)の dsm として、

$$dsm(T\varphi,\ T\eta) = [1-|d_\eta(\varphi)|^2 / \sum_{\eta\in\Psi_j} |d_\eta(\varphi)|^2]^{1/2} \tag{B12.13}$$

があり、この場合も定理B.2を適用できる。

B.13 1次従属形再帰的類似度(タイプ11のS再帰的S類似度)SM

式(B2.3)の関数 f_j として、

$$f_j(T\varphi, T\eta) \equiv |d_\eta(\varphi)|^2 / \max_{\eta\in\Psi_j} |d_\eta(\varphi)|^2 \tag{B13.1}$$

を採用すれば、

$$(a)\ \forall j\in J,\ \max_{\eta\in\Psi_j} f_j(T\omega_j, T\eta)=1>0 \tag{B13.2}$$

$$\begin{aligned}&(b)\ \forall j\in J,\ \forall i\in J-\{j\},\\ &\qquad \max_{\eta\in\Psi_i} f_j(T\omega_j, T\eta)>0\\ &\qquad \because\ 2式(B1.11),\ (B1.13)\end{aligned} \tag{B13.3}$$

を得、2式(B2.1),(B2.2)を満たしており、定理B.1を適用でき、$T\varphi$ が各 $T\omega_j$ と1次従属の関係にある程度を計量化した1次従属形類似度と称される式(B1.2)の SM が得られたことになる。

尚、例えば、2性質(B11.4),(B11.5)を満たす式(11.6)の dsm として、

$$dsm(T\varphi, T\eta) = [1-|d_\eta(\varphi)|^2 / \max_{\eta\in\Psi_j} |d_\eta(\varphi)|^2]^{1/2} \tag{B13.4}$$

があり、この場合も定理B.2を適用できる。

B.14 1次従属形再帰的類似度(タイプ12のS再帰的S類似度)SM

式(B2.3)の関数 f_j として、

$$f_j(T\varphi, T\eta) \equiv |d_\eta(\varphi)|^2 \tag{B14.1}$$

を採用すれば、

(a) $\forall j \in J, \max_{\eta \in \Psi_j} f_j(T\omega_j, T\eta) > 0$

$\because$ 式(B1.11) (B14.2)

(b) $\forall j \in J, \forall i \in J - \{j\}$,

$\max_{\eta \in \Psi_i} f_j(T\omega_j, T\eta) > 0$

$\because$ 式(B1.13) (B14.3)

を得、2式(B2.1), (B2.2)を満たしており、定理B.1を適用でき、$T\varphi$ が各 $T\omega_j$ と1次従属の関係にある程度を計量化した1次従属形類似度と称される式(B1.2)の SM が得られたことになる。

付録C. 各種認識方式

本付録Cでは、処理の対象とする問題のパターン $\varphi \in \Phi$ に対し、帰属するかも知れないカテゴリに関する不確かさが 0 に解消されるようなカテゴリを選択する多段階認識法として、不動点探索形構造受精多段階認識法があることが説明される、最大類似度認識法を初めとする各種単段階認識法も説明される。

[パターン φ, パターンモデル $T\varphi$ に関する知覚仮定]

パターンモデル $T\varphi$ を見たり聞いたりしたならば、

原パターン $\varphi \in \Phi$ と同じように知覚される (C.0)

が以下では、採用されている。

C1. 諸前提

C1.1 全カテゴリ集合 $\underline{\mathfrak{C}}$ と、その生起確率$p(\mathfrak{C}_j)$

全カテゴリ集合

$$\underline{\mathfrak{C}} = \{\mathfrak{C}_j \mid j \in J\} \tag{C.1}$$

における第 $j \in J$ 番目のカテゴリ $\mathfrak{C}_j$ の先見的な生起確率 $p(\mathfrak{C}_j)$は、確率条件

$$[\forall j \in J, 0 \leqq p(\mathfrak{C}_j) \leqq 1] \tag{C.2}$$

$$\wedge \sum_{j \in J} p(\mathfrak{C}_j) = 1 \tag{C.3}$$

を満たしている。

処理の対象とする問題のパターン φ はある可分なヒルベルト空間 $\underline{\mathfrak{H}}$ の元であり、正常なパターン φ は、全カテゴリ集合 $\underline{\mathfrak{C}}$ 内の、唯1つのカテゴリ、例えば、第$j \in J$番目のカテゴリ $\mathfrak{C}_j$に帰属している。

C1.2 axiom 1を満たすパターン集合 Φ と、モデル構成作用素 T との対 $[\Phi, T]$

axiom 1を満たすパターン集合$\Phi(\subset \mathfrak{H})$と、モデル構成作用素

$$T : \Phi \to \Phi \tag{C.4}$$

との対 $[\Phi, T]$ を導入しなければならない。

処理の対象とする問題のパターン φ の集合 Φ の構成的性質について説明しておこう。

パターンと判明している φ の集合としての、基本領域$(0 \in)\Phi_B(\subset \mathfrak{H})$を持ち出す。

パターンモデル $T\varphi$ の集合

$$T \cdot \Phi \equiv \{T\varphi \mid \varphi \in \Phi\} \tag{C.5}$$

と、パターン φ の集合 Φ の任意の正定数倍の集合

$$R^{++}\cdot\Phi\equiv\{T\varphi \mid r^{++}\in R^{++},\ \varphi\in\Phi\} \tag{C.6}$$

ここに、R^{++} は正実数の集合

とを導入して、パターンの帰納的定義から定まる再帰領域方程式

$$\Phi=\Phi_B\cup T\cdot\Phi\cup R^{++}\cdot\Phi \tag{C.7}$$

の解

$$\Phi=R^{++}\cdot(\Phi_B\cup T\cdot\Phi_B) \tag{C.8}$$

に注目する。2性質

$$[R^{++}\cdot\Phi=\Phi]\wedge[T\cdot\Phi=T\cdot\Phi_B\subset\Phi] \tag{C.9}$$

が成立している。

処理の対象とする問題のパターン φ の集合は式(C.8)のように表現され得る“$\mathfrak{H}$ の部分集合としての構成的集合” Φ である。

C1.3 代表パターン ω_j と、そのパターンモデル $T\omega_j$ の2つの集合 Ω, $T\cdot\Omega$ の1次独立性

第$j\in J$番目のカテゴリ $\mathfrak{C}_j$ の代表パターン ω_j の集合

$$\Omega\equiv\{\omega_j \mid j\in J\}\subset\Phi \tag{C.10}$$

と、パターンモデル $T\omega_j$ の集合

$$T\cdot\Omega\equiv\{T\omega_j \mid j\in J\} \tag{C.11}$$

とは1次独立であるとする。

C2. 最大類似度認識法

第$t(=0, 1, 2, \cdots)$パターン変換段階で得られている“axiom 2を満たす類似度関数”

$$SM(\cdot, \cdot\ ;t):\Phi\times\Omega\rightarrow\{s \mid 0\leqq s\leqq 1\} \tag{C.12}$$

を考える。

不等式

$$SM(\varphi,\ \omega_j\ ;t)\geqq\max_{i\in J-\{j\}} SM(\varphi,\ \omega_i\ ;t) \tag{C.13}$$

を満たす最も若いカテゴリ番号 $j\in J$ を見つけて、つまり、

$$\mathrm{argmax}_{k\in J}\, SM(\varphi,\ \omega_k\ ;t)=j\in J \tag{C.14}$$

であれば、

第 t パターン変換段階では、処理の対象とする問題
の入力パターン $\varphi\in\Phi$ は第$j\in J$番目のカテゴリ
$\mathfrak{C}_j$ に帰属する　(C.15)

と、認識する。

C3. 最大相互情報量認識法

C3.1 パターンモデル $T\varphi$ の出現確率 $p_1(T\varphi)$

2条件

$$\forall i\in J,\ \forall j\in J-\{i\},\ \ \Phi_i\cap\Phi_j=\phi \tag{C.16}$$

$$\forall j\in J,\ \omega_j\in\Phi_j \tag{C.17}$$

を満たすように、パターンの有限集合 Φ_j の系

$$\Phi_j(\subset\Phi),\ \ j\in J \tag{C.18}$$

を選定する。このとき、

$$\forall j \in J,\ \forall i \in J-\{j\},\quad \omega_i \overline{\in} \Phi_j \tag{C.19}$$

も成立していることに注意しておく。

確率条件

$$[\forall \varphi \in \Phi,\ 0 \leqq p_1(T\varphi) \leqq 1] \wedge \sum_{\varphi \in \Phi} p_1(T\varphi)=1 \tag{C.20}$$

を満たすパターンモデル $T\varphi$ の出現確率 $p_1(T\varphi)$ を、以下の4種類などの $q(T\varphi)$ を用い、

$$p_1(T\varphi) \equiv \begin{cases} q(T\varphi) / \sum_{\psi \in \Phi} q(T\psi) \\ \quad \cdots \sum_{\psi \in \Phi} q(T\psi) > 0 \text{ のとき} \\ 1/|\Phi| \quad \cdots \sum_{\psi \in \Phi} q(T\psi) = 0 \text{ のとき} \end{cases} \tag{C.21}$$

と与えよう。ここに、$(\ ,\ )$, $\|\cdot\| \equiv \sqrt{(\cdot,\cdot)}$ は各々、$\mathfrak{H}$ の内積とノルムである：

例えば、非一致条件

$$\forall j \in J,\ \forall i \in J-\{j\},\ \| T\omega_i - T\omega_j \| > 0 \tag{C.21}$$

の下で、考えよう。

①$q(T\varphi)$

$$\equiv \prod_{k \in J} \max_{\eta \in \Phi_k} \exp\left[-\| T\varphi - T\eta \|^2 / (2\sigma_k{}^2)\right]$$

$$,\ \sigma_k > 0. \tag{C.22}$$

②$q(T\varphi)$

$$\equiv \prod_{k \in J} \max_{\eta \in \Phi_k} \exp\left[-\{\| T\varphi \| - \| T\eta \|\}^2\right.$$

$$\left. / (2\sigma_k{}^2)\right],\quad \sigma_k > 0. \tag{C.23}$$

③$q(T\varphi)$

$$\equiv \prod_{k \in J} \max_{\eta \in \Phi_k} \left[1 - \exp\{-|(T\varphi, T\eta)\right.$$

$$\left. / \{\| T\varphi \| \cdot \| T\eta \|\}|^2 / (2\sigma_k{}^2)\}\right],\quad \sigma_k > 0. \tag{C.24}$$

④実数値条件

$$\forall \varphi \in \Phi,\ \forall i \in J,\ (T\varphi, T\omega_i) \in R \text{(実数全体の集合)}$$

を設け、非一致条件式(C.21)の下で、

$$\text{分散}\ \ \sigma_\eta(\varphi) \equiv \| T\omega_j \| \cdot \sqrt{1 - |\rho_\eta(\varphi)|^2} \tag{C.25}$$

規格化相関値 $\rho_\eta(\varphi) \equiv$

$$\begin{cases} (T\varphi, T\eta) / [\| T\varphi \| \cdot \| T\eta \|] \\ \quad \cdots \| T\varphi \| \cdot \| T\eta \| > 0 \text{ のとき} \\ 0 \cdots \| T\varphi \| \cdot \| T\eta \| = 0 \text{ のとき} \end{cases} \tag{C.26}$$

を導入し、$T\varphi$, $T\eta\ (\eta \in \Phi_k)$ 間の規格化相関値 $\rho_\eta(\varphi)$ が大きくなればなるほど大きい値をとる非負量

$$q_\eta(\varphi)$$

$$\equiv$$

$$\begin{cases} \exp\left[-\| (T\eta - \| T\eta \|)\right. \\ \quad - \sigma_\eta(\varphi) \cdot \{\| T\eta \| / \| T\varphi \|\} \cdot (T\varphi - \| T\varphi \|) \|^2 \\ \quad \left. / \{2\sigma_\eta(\varphi)^2\}\right] \quad \cdots |\rho_\eta(\varphi)| < 1 \text{ のとき} \\ 1 \quad \cdots |\rho_\eta(\varphi)| = 1 \text{ のとき} \end{cases} \tag{C.27}$$

を定義して、Gauss確率密度関数からhintを得た量

$$q(T\varphi)\equiv\prod_{k\in J}\max_{\eta\in\Phi_k}q_\eta(\varphi). \tag{C.28}$$

④については、

文献［B18］の1.4節の例1.1を参照 □

C3.2 TΦ, Ω 間の期待相互情報量 EMI(TΦ, Ω)

以後、式(C.12)の SM(φ, ω_j ; t)での助変数(段階番号)t を裏に隠して、SM(φ, ω_j)と表す。

次に、Tφ, ω_jの同時出現確率 p(Tφ, ω_j)を

$$p(T\varphi,\ \omega_j)\equiv SM(T\varphi,\ \omega_j)\cdot p_1(T\varphi) \tag{C.29}$$

と定義する。このとき、Tφが出現したとき、ω_jの条件付き出現確率 $p_{2/1}(\omega_j/T\varphi)$は、

$$p_{2/1}(\omega_j/T\varphi)\equiv p(T\varphi,\ \omega_j)/p_1(T\varphi) \tag{C.30}$$

$$=SM(T\varphi,\ \omega_j) \tag{C.31}$$

ということになる。そして、ω_j の出現確率 $p_2(\omega_j)$は、

$$p_2(\omega_j)\equiv\sum_{\varphi\in\Phi}p(T\varphi,\ \omega_j) \tag{C.32}$$

$$=\sum_{\varphi\in\Phi}SM(T\varphi,\ \omega_j)\cdot p_1(T\varphi) \tag{C.33}$$

と表される。因みに、ω_jが出現したとき、パターンモデルTφの条件付き出現確率 $p_{1/2}(T\varphi/\omega_j)$は

$$p_{1/2}(T\varphi/\omega_j)\equiv p(T\varphi,\ \omega_j)/p_2(\omega_j) \tag{C.34}$$

と表される。

$$\begin{aligned}&EMI(T\Phi,\ \Omega)\\&\equiv\sum_{\varphi\in\Phi}\sum_{j\in J}p(T\varphi,\ \omega_j)\cdot\log_e\left[p(T\varphi,\ \omega_j)/\right.\\&\quad\left.\{p_1(T\varphi)\cdot p_2(\omega_j)\}\right]\end{aligned} \tag{C.35}$$

は、TΦ, Ω 間の期待相互情報量(expected mutual information)といわれる。

C3.3 認識システムが認識処理した後に、Φ の各元 φ のカテゴリ帰属に関し、獲得された平均的情報量としての期待相互情報量 EMI(TΦ, Ω)

非負量であるShannon平均的情報量

$$H_2(\Omega)\equiv-\sum_{j\in J}p_2(\omega_j)\cdot\log_e p_2(\omega_j) \tag{C.36}$$

は

パターン集合 Φ を認識処理する前に、各$\varphi\in\Phi$ が
式(C.10)のΩ を代表パターン集合とする式(A.1)の
全カテゴリ集合 $\underline{\mathfrak{C}}$ の内のどの1つに帰属するかに関した
平均的不確定さ (C.37)

を表し、また、非負量であるShannon平均的情報量

$$\begin{aligned}&H_{1/2}(\Omega/T\varphi)\\&\equiv-\sum_{j\in J}p_{2/1}(\omega_j/T\varphi)\cdot\log_e p_{2/1}(\omega_j/T\varphi)\end{aligned} \tag{C.38}$$

は、

特定のパターン $\varphi\in\Phi$ を認識処理した後に、この
特定の $\varphi\in\Phi$ が式(A .10)のΩ を代表パターン集
合とする式(C.1)の全カテゴリ集合 $\underline{\mathfrak{C}}$ の内どの
1つに帰属するかに関した不確定さ (C.39)

を表す。最後に、この不確定さをパターンモデル Tφ の、式(C.21)の出現確率 $p_1(T\varphi)$で平均し

た非負量であるShannon平均的情報量

$$H_{1/2}(\Omega/T\cdot\Phi)$$

$$\equiv \sum_{\varphi\in\Phi} p^1(T\varphi)\cdot H_{1/2}(\Omega/T\varphi) \qquad (C.40)$$

$$= -\sum_{\varphi\in\Phi} p_1(T\varphi)$$

$$\cdot\sum_{j\in J} p_{2/1}(\omega_j/T\varphi)\cdot\log_e p_{2/1}(\omega_j/T\varphi)$$

$$\because \text{ 式(C.38)} \qquad (C.41)$$

$$= -\sum_{\varphi\in\Phi}\sum_{j\in J} p(T\varphi,\ \omega_j)$$

$$\cdot\log_e p_{2/1}(\omega_j/T\varphi)$$

$$\because \text{ 式(C.30)} \qquad (C.42)$$

は、

パターン集合 Φ を認識処理した後、各 $\varphi\in\Phi$ が
式(C.10)の Ω を代表パターン集合とする式(C.1)
の全カテゴリ集合 $\underline{\mathfrak{C}}$ の内のどの1つに帰属するか
に関した平均的不確定さ (C.43)

を表している。

2式(C.32), (C.30)を適用して、

$$\mathrm{EMI}(T\Phi,\ \Omega) = H_2(\Omega) - H_{1/2}(\Omega/T\cdot\Phi) \qquad (C.44)$$

と表されるから、

パターンモデル $T\varphi$ を見たり聞いたりしたならば、
原パターン $\varphi\in\Phi$ と同じに見えたり聞こえたりする (C.45)

という式(C.0)の［パターン φ, パターンモデル $T\varphi$ に関する知覚仮定］の立場からは、

$\mathrm{EMI}(T\Phi,\ \Omega)$
＝［認識システムが処理する前に、Φ の各元 φ の
カテゴリ帰属に関し、認識システムが持っていた平
均的不確定さ］
－［認識システムが処理した後に、Φ の各元 φ の
カテゴリ帰属に関し、いまだ認識システムが持って
いるであろう平均的不確定さ］ (C.46)
＝認識システムが処理した後に、Φ の各元 φ の
カテゴリ帰属に関し、解消され取り去られた平均的
不確定さ (C.47)
＝認識システムが処理した後に、Φ の各元 φ の
カテゴリ帰属に関し、獲得された平均的情報量 (C.48)

ということになる。

C3.4 取り去られた最大の不確定さに基づいて、個々のパターン φ を認識する方法

非負量

$$\mathrm{MI}(T\varphi,\ \omega_j)$$

$$\equiv -p_2(\omega_j)\cdot\log_e p_2(\omega_j)$$

$$-\{-p_{2/1}(\omega_j/T\varphi)\cdot\log_e p_{2/1}(\omega_j/T\varphi)\} \qquad (C.49)$$

$$=-\left[\sum_{\varphi\in\Phi}SM(T\varphi,\ \omega_j)\cdot p_1(T\varphi)\right]$$
$$\cdot\log_e\left[\sum_{\varphi\in\Phi}SM(T\varphi,\ \omega_j)\cdot p_1(T\varphi)\right]$$
$$-\ \{-SM(T\varphi,\ \omega_j)\cdot\log_e SM(T\varphi,\ \omega_j)\}$$
$$\because\ \ 2式(C.33),\ (C.31) \qquad (C.50)$$

を定義すれば、期待相互情報量 $EMI(T\Phi,\ \Omega)$は、

$$EMI(T\Phi,\ \Omega)$$
$$=H_2(\Omega)-H_{1/2}(\Omega/T\cdot\Phi)\quad\because\ 式(C.43)$$
$$=\sum_{\varphi\in\Phi}p_1(T\varphi)\cdot H_2(\Omega)-H_{1/2}(\Omega/T\cdot\Phi)$$
$$\because\ 式(C.20) \qquad (C.51)$$
$$=\sum_{\varphi\in\Phi}p_1(T\varphi)\cdot\left[-\sum_{j\in J}p_2(\omega_j)\right.$$
$$\left.\cdot\log_e p_2(\omega_j)\right]$$
$$-\left[-\sum_{\varphi\in\Phi}p_1(T\varphi)\right.$$
$$\left.\cdot\sum_{j\in J}p_{2/1}(\omega_j/T\varphi)\cdot\log_e p_{2/1}(\omega_j/T\varphi)\right]$$
$$\because\ 式(C.36),\ (C.41)$$
$$=\sum_{\varphi\in\Phi}p_1(T\varphi)\cdot\sum_{j\in J}\left[-p_2(\omega_j)\right.$$
$$\cdot\log_e p_2(\omega_j)$$
$$\left.-\{-p_{2/1}(\omega_j/T\varphi)\cdot\log_e p_{2/1}(\omega_j/T\varphi)\}\right]$$
$$=\sum_{\varphi\in\Phi}p_1(T\varphi)\cdot\sum_{j\in J}MI(T\varphi,\ \omega_j) \qquad (C.52)$$

と変形できるから、知覚仮定式(C.45)の立場からは、式(C.43)の $EMI(T\Phi,\ \Omega)$は、

処理の対象としている問題のパターン集合 Φ の各
元 $\varphi\in\mathfrak{H}$ がどの1つのカテゴリに最も大に従属的
に帰属するかを決定する認識処理について、取り去
られる平均的不確定さ (C.53)

を表しているということを思い起こすと、式(C.49)の $MI(T\varphi,\ \omega_j)$は、

処理の対象としている問題のパターン $\varphi\in\mathfrak{H}$ が
第$j\in J$のカテゴリ $\mathfrak{C}_j$ に帰属するかを決定する認識処理に
ついて、取り去られる不確定さ (C.54)

を表している。

よって、

カテゴリ帰属判断において取り去られる不確定さが最も大なる
カテゴリを、問題の特定のパターン $\varphi\in\Phi$が帰属する
カテゴリであるとする次の最大相互情報量認識法 (C.55)

は、無意味なことではない： 不等式

$$MI(T\varphi,\ \omega_j)\geqq\max_{i\in J-\{j\}}MI(T\varphi,\ \omega_i) \qquad (C.56)$$

を満たす1つのカテゴリ番号 $j\in J$ を見つけ、つまり、

$$\mathrm{argmax}_{i\in J}\ MI(T\varphi,\ \omega_i)=j\in J \qquad (C.57)$$

を見つけ、

φ belongs to the j-th category (C.58)

という具合に決定する。 □

C3.5 2つの適切さ

次の2事項(a), (b)に注意する。

(a)式(C.33)の $p_2(\omega_j)$ の表現式は、axiom 2を満たす類似度関数 SM を助変数として含んでいるから、$p_2(\omega_j)$ を $p_2(\omega_j\,;\,SM)$ と表そう。そうすると、式(C.37)の $H_2(\Omega)$ は $H_2(\Omega\,;\,SM)$ と表さなければならない。

処理の対象とするパターン集合 Φ を固定している場合、2つの“axiom 2を満たす類似度関数 SM_1, SM_2”について、不等式

$$H_2(\Omega\,;\,SM_1) \geqq H_2(\Omega\,;\,SM_2) \tag{C.59}$$

が成り立っていれば、

SM_2 の方が SM_1 より、Φ の各元 φ を正しく認識するにあたり、パターン集合 Φ に関して最大相互情報量認識法が適切に機能すると言えよう。

(b)パターン集合 Φ の持つ平均情報量(個々のパターンモデル $T\varphi \in \Phi$ を指定するときの平均的不確定さ)

$$H_1(T\cdot\Phi) \equiv -\sum_{\varphi\in\Phi} p_1(T\varphi)\cdot\log_e p_1(T\varphi) \tag{C.60}$$

を定義する。いま1つのパターン集合 Ψ を持ち出し、式(A.21)内の q とaxiom 2を満たす類似度関数 SM とを固定し、$H_1(T\cdot\Psi)$ を計算する。その結果、

$$|T\cdot\Phi| = |T\cdot\Psi| \wedge H_1(T\cdot\Phi) \geqq H_1(T\cdot\Psi) \tag{C.61}$$

であれば、

Φ は Ψ に比べ、最大相互情報量による認識が容易でない

と言えよう。

C4. 最大遷移確率に基づく認識法

第 $j\in J$ 番目のカテゴリ $\mathfrak{C}_j$ の代表パターン ω_j の出現確率 $p_2(\omega_j)$ は2式(C.32), (C.33)で求められている。代表パターン ω_j が出現したとき、パターンモデル $T\varphi$ の条件付き出現確率 $p_{1/2}(T\varphi/\omega_j)$ は、

$$\begin{aligned}
&p_{1/2}(T\varphi/\omega_j)\\
&=SM(T\varphi,\ \omega_j)\cdot p_1(T\varphi)\\
&\quad /\sum_{\varphi\in\Phi} SM(T\varphi,\ \omega_j)\cdot p_1(T\varphi)p_2(\omega_j)\\
&\qquad\because\ \text{式(C.34), (C.29), (C.33)}
\end{aligned} \tag{C.62}$$

と表されることに留意して、式(C.10)の代表パターン集合 Ω に基づいて、式(C.5)のパターンモデル集合 $T\cdot\Phi$ 内の2つのパターンモデル $T\varphi$, $T\eta$ の間の関係を要約した統計量 $sr(T\varphi, T\eta)$ として、

$$\begin{aligned}
&sr(T\varphi, T\eta)\\
&\equiv\sum_{j\in J} p_{1/2}(T\varphi/\omega_j)\cdot p_{1/2}(T\eta/\omega_j)\cdot p_2(\omega_j)
\end{aligned} \tag{C.63}$$

を定義する。

式(C.63)は、文献［A7］での数量化手法の背後確率にhintを得たものである。

$$\begin{aligned}
&sr(T\varphi, T\eta)\\
&=\sum_{j\in J} p_{1/2}(T\varphi/\omega_j)\cdot p(T\eta,\ \omega_j)\\
&\qquad\because\ \text{式(C.34)}\\
&=\sum_{j\in J} p(T\varphi,\ \omega_j)\cdot p(T\eta/\omega_j)/p_2(\omega_j)
\end{aligned}$$

$\because$ 式(C.34) (C.64)

$= \sum_{j \in J} SM(T\varphi, \omega_j) \cdot p_1(T\varphi) \cdot SM(T\eta, \omega_j)$

$\cdot p_1(T\eta) / \sum_{\psi \in \Phi} SM(T\psi, \omega_j) \cdot p_1(T\psi)$

$\because$ 2式(C.29), (C.33) (C.65)

とも再表現されることに注意しておこう。

注目すべき等式

$\sum_{\varphi \in \Phi} sr(T\varphi, T\eta)$

$= \sum_{j \in J} p_{1/2}(T\eta / \omega_j) \cdot p_2(\omega_j)$

$\because$ 2式(C.34), (C.32)

$= \sum_{j \in J} p(T\eta, \omega_j)$ $\because$ 式(C.34)

$= p_1(T\eta)$ $\because$ 周辺確率 $p_1(T\eta)$の定義 (C.66)

が成立する。モデル$T\eta$の出現確率$p_1(T\eta)$は式(C.21)で与えられているものである。同様に、等式

$\sum_{\eta \in \Phi} sr(T\varphi, T\eta) = p_1(T\varphi)$ (C.67)

も成り立つ。よって、

$tp(T\varphi / T\eta) \equiv sr(T\varphi, T\eta) / p_1(T\eta)$ (C.68)

について、

$p_1(T\eta) > 0$ であれば、$\sum_{\varphi \in \Phi} tp(T\varphi / T\eta) = 1$

$\because$ 式(C.66) (C.69)

が成り立つ。同様に、

$tp(T\eta / T\varphi) \equiv sr(T\varphi, T\eta) / p_1(T\varphi)$ (C.70)

について、

$p_1(T\varphi) > 0$ であれば、$\sum_{\eta \in \Phi} tp(T\eta / T\varphi) = 1$

$\because$ 式(C.67) (C.71)

が成り立つ。

式(C.71)を考慮すると、式(C.70)の$tp(T\eta / T\varphi)$は、

式(C.10)の Ω に基づいた、式(C.11)の $T \cdot \Phi$

内のパターンモデル $T\varphi$ から $T\eta$ への

遷移確率(transition probability)である (C.72)

という解釈を可能にする。

$\forall \varphi \in \Phi, tp(T\varphi / T\varphi) \gg \max_{\eta \in \Phi - \{\varphi\}} tp(T\eta / T\varphi)$ (C.73)

が成立していこと、特に、

$\forall j \in J, tp(T\omega_j / T\omega_j)$

$\gg \max_{i \in J - \{j\}} tp(T\omega_i / T\omega_j)$ (C.74)

が成立していることが望ましいが、

$\Omega \subset \Phi$ の場合、$T\varphi$ からどの $T\omega_j$ へ最も大なる確率で遷移するかを考えて、不等式

$tp(T\omega_j / T\varphi) \geqq \max_{i \in J - \{j\}} tp(T\omega_i / T\varphi)$ (C.75)

を満たす最も若いカテゴリ番号 $j \in J$ を見つけ、つまり、

$\mathrm{argmax}_{i \in J}\, tp(T\omega_i / T\varphi) = j \in J$ (C.76)

を見つけ、式(C.58)の如く認識する“最大遷移確率認識法”が提案される。

C5. 組み合わせ山登り法に基づいた多段階想起形認識法

文献［B20］の付録Cに解説されている。

C6. ε－最小ステップ認識法

文献［B20］の付録Dに解説されている。

C7. 想起形認識法としての、不動点探索形構造受精多段階認識法

カテゴリ帰属知識とは、パターンと、このパターンが帰属する可能性があると思われる候補カテゴリの番号リストとの順序対である［B3］，［B4］。

本章では、カテゴリ帰属知識を用いないで、SS理論における不動点探索形構造受精多段階想起認識法を簡単に説明しておこう。

C7.1 不動点を連想する形式を採用した不動点探索形構造受精多段階認識法の、記述的な説明

パターン認識の数学的理論(SS理論)［B1］～［B6］では、次のように、認識システムの持っているカテゴリ帰属知識に関する多段階パターン連想法、正確にいえば、(カテゴリ帰属知識の)**不動点**(を)**連想**(する)**形**(式の)**多段階認識法**(不動点探索形構造受精多段階想起認識法)が考えられている：

(1)先ず、処理の対象とする問題の入力パターン φ に対応するパターンモデル $T\varphi$ を求める。

(2)その後、$T\varphi$ から "現在の認識段階で確保されているカテゴリ帰属知識を変換する機能を備えた構造受精変換" の不動点パターンモデルを連想形認識方程式を解く形式で多段階にわたって連想する。

(3)こうして得られた不動点パターンモデルは、極端に変形していた入力パターン φ である場合を除いて、ある1つのカテゴリの代表パターンのモデルに近い構造に変換され得られているから、入力パターン φ の帰属するカテゴリを容易に決定することができる。 □

C7.2 多段階想起認識過程の近似的収束

SM(η, ω_j)は、パターン$\eta \in \Phi$が $\omega_j \in \Omega \subset \Phi$
をその代表パターンとする第$j \in J$番目のカテゴリ
$\mathfrak{C}_j$ を表している程度(類似度)である (C.77)

と解釈することに注意する。

「パターンモデル $T\varphi \in \Phi$ を見たり聞いたりすれば、処理の対象とする問題の
入力パターン(原パターン)$\varphi \in \Phi$と同じように見えたり聞こえたりする」 (C.78)

という**"パターンモデル $T\varphi$ と原パターン φ との同一知覚仮定"**を採用し、

ある非負整数 $t\,(=0, 1, 2, \cdots)$が存在して、

パターンモデル $T\varphi$ を第0認識段階で得られたパターンとして、後続の第 t 認識段階で得られたパターン $\varphi_t \in \Phi$ について

不等式

$$0 \leqq \delta < 2^{-1} \tag{C.79}$$

を満たす非負数 δ について

不等式

$$\exists j \in J, \mathrm{SM}(\varphi_t, \omega_j) \geqq 1 - \delta \tag{C.80}$$

が成立すれば(多段階認識過程が近似的に収束すれば)、

「$T\varphi$, 即ち、φ に対応して知覚的に最終的に
記憶系から想起される知覚的記憶像」 (C.81)

は、

「φ の帰属するカテゴリが第$j\in J$番目のカテゴリ $\mathfrak{C}_j$で
あるならば、ω_jを代表パターンとする第$j\in J$番目の
カテゴリ $\mathfrak{C}_j$のパターンモデル $T\omega_j$ である」 (C.82)

という解釈を可能にする多段階想起認識法が前節の不動点探索形構造受精多段階認識法である。

C8. 逐次観測に基づいた認識法

逐次観測方式による認識の働き(the sequential Bayesian decision scheme or the sequential Bayes classifier)について説明しよう。

C8.1 逐次観測方式による認識法

これまで通り、

$|J|$(カテゴリ総数)≥ 2

とする。

探索力学系(the search dynamics)での探索法は次のように説明される。

処理の対象とする問題のパターン $\varphi\in\Phi$ がカテゴリ集合

$$\mathfrak{C}_j,\ j\in\gamma \tag{C.83}$$

の何れか1つのカテゴリに帰属すると判明していると判明しているとしよう。$\gamma\subset J$ をリストで表現して、パターン φ の候補カテゴリ番号リストと呼ぶ。このパターン$\varphi\in\Phi$に対し、候補カテゴリ番号リストμ_sの列

$$\mu_0,\ \mu_1,\ \mu_2,\ \cdots,\ \mu_s,\ \cdots,\ \mu_t\in 2^J \tag{C.84}$$

を適切にこの順序通りに選んでいって、

$$\varphi_0\equiv T\varphi\in\Phi \tag{C.85}$$

$$\varphi_s\equiv TA(\mu_{s-1}\cap\gamma_{s-1})T\varphi_{s-1}\in\Phi\ (s=1,2,\cdots,t) \tag{C.86}$$

と、

$$\gamma_0\equiv\gamma\subset J \tag{C.87}$$

$$\gamma_s\equiv CSF(\varphi_{s-1},\ \mu_{s-1}\cap\gamma_{s-1})\in 2^J(s=1,2,\cdots,t) \tag{C.88}$$

とを、求める(T, $A(\mu)$ $(\mu\subset J)$, CSFは各々、モデル構成作用素、構造受精作用素、カテゴリ選択関数である [B3], [B4])。言い替えれば、カテゴリ帰属知識の列

$$\langle\varphi_0,\ \gamma_0\rangle=_\triangle\langle T\varphi,\ \gamma\rangle\in\langle\Phi,\ 2^J\rangle \tag{C.89}$$

$$\langle\varphi_s,\ \gamma_s\rangle=_\triangle TA(\mu_{s-1})T\cdot\langle\varphi_{s-1},\ \gamma_{s-1}\rangle\in\langle\Phi,\ 2^J\rangle \quad (s=1,2,\cdots,t) \tag{C.90}$$

を求める(探索力学;the search dynamics)。

得られたパターンの系列

$$\varphi_0,\ \varphi_1,\ \varphi_2,\ \cdots,\ \varphi_s,\ \cdots,\ \varphi_t\in\Phi \tag{C.91}$$

を、この並べた順序を保存して、Φ_0^tと書くことにしよう。

このとき、

$\exists j\in J$, $p(\mathfrak{C}_j/\Phi_0^t)$ (Φ_0^tが与えられたとき、第$j\in J$ 番目の
カテゴリ $\mathfrak{C}_j$の条件付き生起確率)

$\geqq 1-\delta$ (C.92)

ここに、$0 \leqq \delta < 2^{-1}$ (C.93)

を満たす第 t 認識段階が存在すれば、

The pattern φ in question is assigned to one,
i.e., the j-th category $\mathfrak{C}_j$ of the $|J|$
possible categories $\underline{\mathfrak{C}} = \{\mathfrak{C}_j \mid j \in J\}$ (C.94)

と、入力パターン $\varphi \in \Phi$ を認識する。 □

条件付き誤認識確率(the conditional probability of misclassification) $p(\mathrm{error}/\Phi_0^t)$ は、

$$p(\mathrm{error}/\Phi_0^t) \equiv 1 - p(\mathfrak{C}_j/\Phi_0^t) \tag{C.95}$$

と与えられる。

C8.2 $p(\mathfrak{C}_j/\Phi_0^t)$ の求め方

Bayes' theorem を適用すると、事前生起確率分布(a priori probability distribution of occurrences)

$$p(\mathfrak{C}_j),\ j \in J \tag{C.96}$$

から、事後生起確率分布(a posteriori probability distribution of occurrences)

$$p(\mathfrak{C}_j/\Phi_0^t),\ j \in J \tag{C.97}$$

への計算公式

$$p(\mathfrak{C}_j/\Phi_0^t) = [p(\Phi_0^t/\mathfrak{C}_j)/p(\Phi_0^t)] \cdot p(\mathfrak{C}_j),\ j \in J \tag{C.98}$$

を逐次的に分解しよう。それには、

an iterative application of Bayes' theorem を適用して、次の2式(C.101), (C.102)の the recursive computation を行う。

初期条件

$$p(\mathfrak{C}_j/\Phi_0^0) = p(\mathfrak{C}_j/\varphi_0) \tag{C.99}$$

$$\begin{aligned} p(\varphi_1/(\mathfrak{C}_i,\ \Phi_0^0)) &= p(\varphi_1/(\mathfrak{C}_i,\ \varphi_0)) \\ &= p(\mathfrak{C}_i,\ \varphi_0,\ \varphi_1)/p(\mathfrak{C}_i,\ \varphi_0) \\ &= p(\mathfrak{C}_i,\ \Phi_0^1)/p(\mathfrak{C}_i,\ \varphi_0) \end{aligned} \tag{C.100}$$

を設定し、

$s = 1, 2, \cdots$ について、

$$\begin{aligned} p(\mathfrak{C}_j/\Phi_0^s) &= [p(\varphi_s/(\mathfrak{C}_j,\ \Phi_0^{s-1}))/p(\varphi_s/\Phi_0^{s-1})] \cdot \\ &\quad p(\mathfrak{C}_j/\Phi_0^{s-1}),\ j \in J \end{aligned}$$

(the updating equations for $p(\mathfrak{C}_j/\Phi_0^s)$) (C.101)

$$p(\varphi_s/\Phi_0^{s-1}) = \sum_{j \in J} p(\varphi_s/(\mathfrak{C}_j,\ \Phi_0^{s-1})) \cdot p(\mathfrak{C}_j/\Phi_0^{s-1}) \tag{C.102}$$

と、計算を行えばよい。

2式(C.101), (C.102)の導出については、文献［A8］の2式(2), (3)から、hintを得ているが、因みに、2式(C.101), (C.102)を証明しておこう。

先ず、式(C.101)の成立については、

式(C.101)の右辺

$= [\{p(\mathfrak{C}_j,\ \Phi_0^{s-1},\ \varphi_s)/p(\mathfrak{C}_j,\ \Phi_0^{s-1})\}$

$$/\{p(\Phi_0^{s-1},\ \varphi_s)/p(\Phi_0^{s-1})\}]\cdot$$
$$[p(\mathfrak{C}_j,\ \Phi_0^{s-1})/p(\Phi_0^{s-1})]$$
$$=p(\mathfrak{C}_j,\ \Phi_0^{s-1},\ \varphi_s)/p(\Phi_0^{s-1},\ \varphi s)$$
$$=p(\mathfrak{C}_j,\ \Phi_0^{s})/p(\Phi_0^{s})$$
=式(C.101)の左辺 (C.103)

と示され、次に、式(C.102)の成立については、

式(C.102)の右辺

$$=\sum_{j\in J}[p(\mathfrak{C}_j,\ \Phi_0^{s-1},\ \varphi_s)/p(\mathfrak{C}_j,\ \Phi_0^{s-1})]\cdot$$
$$[p(\mathfrak{C}_j,\ \Phi_0^{s-1})/p(\Phi_0^{s-1})]$$
$$=\sum_{j\in J}p(\mathfrak{C}_j,\ \Phi_0^{s-1},\ \varphi_s)/p(\Phi_0^{s-1})$$
$$=p(\Phi_0^{s-1},\ \varphi_s)/p(\Phi_0^{s-1})$$
=式(C.102)の左辺 (C.104)

と示される。 □

C8.3 式(C.99)の $p(\mathfrak{C}_j/\varphi_0)$，式(C.100)の $p(\varphi_1/(\mathfrak{C}_i,\ \Phi_0^0))$，2式(C.101)，(C.102)内の$p(\varphi_s/(\mathfrak{C}_j,\ \Phi_0^{s-1}))$と、式(C.101)内の $p(\varphi_s/\Phi_0^{s-1})$の、類似度 SM による近似的計算

逐次観測方式による認識システムは、無記憶であると仮定して、然も処理対象とする問題のパターン集合Φは高々可算集合としの基本領域(basic domain) Φ_B に限定し、式(C.99)の $p(\mathfrak{C}_j/\varphi_0)$，式(C.100)の $p(\varphi_1/(\mathfrak{C}_i,\ \Phi_0^0))$，2式(C.101)，(C.102)内の $p(\varphi_s/(\mathfrak{C}_j,\ \Phi_0^{s-1}))$と、式(C.101)内の $p(\varphi_s/\Phi_0^{s-1})$を、類似度 SM を使って表現し、その近似的計算が可能であることを示そう。

先ず、

式(C.99)の $p(\mathfrak{C}_j/\varphi_0)$

$$=p(\mathfrak{C}_j)\cdot[SM(\varphi_0,\ \omega_j)/\sum_{\varphi\in\Phi_B}SM(\varphi_0,\ \omega_j)]/$$
$$\sum_{i\in J}p(\mathfrak{C}_i)\cdot[SM(\varphi_0,\ \omega_i)/\sum_{\varphi\in\Phi_B}SM(\varphi_0,\ \omega_i)] \qquad (C.105)$$

と近似され、その次に、

式(C.100)の $p(\varphi_1/(\mathfrak{C}_i,\ \Phi_0^0))$

$\fallingdotseq p(\varphi_1/\mathfrak{C}_i)$ ∵ 無記憶の仮定

$$=SM(\varphi_1,\ \omega_i)/\sum_{\varphi\in\Phi_B}SM(\varphi_1,\ \omega_i) \qquad (C.106)$$

と近似され、更に、

2式(C.101)，(C.102)内の $p(\varphi_s/(\mathfrak{C}_j,\ \Phi_0^{s-1}))$

$\fallingdotseq p(\varphi_s/\mathfrak{C}_j)$ ∵ 無記憶の仮定

$$=SM(\varphi_s,\ \omega_j)/\sum_{\varphi\in\Phi_B}SM(\varphi_s,\ \omega_j) \qquad (C.107)$$

と近似され、最後に、

式(C.101)内の $p(\varphi_s/\Phi_0^{s-1})$

$\fallingdotseq p(\varphi_s)$ ∵ 無記憶の仮定

$$=\sum_{i\in J}p(\mathfrak{C}_i)\cdot p(\varphi_s/\mathfrak{C}_i)$$
$$=\sum_{i\in J}p(\mathfrak{C}_i)\cdot SM(\varphi_s,\ \omega_i)/\sum_{\varphi\in\Phi_B}SM(\varphi_s,\ \omega_i) \qquad (C.108)$$

と近似される。

C9. certainty factorによる認識法

紙面の都合上、別の機会に説明される。

付録D. 2カテゴリ分類容易度 EOC の、帰納的分類規則としての類似度関数 SM の適応的更新への応用

理系・文系双方に関した人類の持つ「知の財産」を受け継ぎながら、知の体系を情報という観点から構築し、他の学問分科の発展を支え、推進するのが、情報学がその誕生以来背負わなければならない宿命である。この宿命を背負う情報学が他の学問分科と異なるのは現在の知識の持つ不完全性を克服し知識を洗練するために、知識推論する情報システムがそのシステム性能の増大を目指しその構造を自己組織化し外界から学習する手法を積極的に研究しなれればならないことである。

本付録Dでは、典型的な情報システムとしての認識システムの持つ各種帰納的分類規則(inductive classification rule)の学習問題において、最初ランダムな分類規則よりもほんの少しだけ良好な分類規則さえ与えられれば、訓練系列(training sequence)を用い、その分類性能を適応的に終局的に分類誤差 0 に改善するために考えられている“確率的学習モデルに関するブースティングアルゴリズム AdaBoost [A10]”からhintを得、axiom 2を満たす類似度関数 SM の性能を改善する適応的手法を2カテゴリ分類容易度 EOC の使用の下で構成してみよう。

処理の対象とする問題のパターン φ の集合 Φ を

$$\Phi = \cup_{j \in J} \Phi_j \cup \Phi 0$$

$$\text{such that } \Phi_i \cap \Phi_j = \phi \, (i \neq j) \tag{D.1}$$

と有限分割しよう。ここに、Φ_0 はその帰属するカテゴリ(類概念 ; category)が全く存在しないか、或いは複数個存在するようなパターンの集合であり、Φ_j は第 $j \in J$ 番目のカテゴリ $\mathfrak{C}_j$ に一意的に帰属するパターン φ の集合である。

任意の $j \in J$ について

$$f_j(\Phi)_1 \equiv \{\varphi \in \Phi \mid f_j(\varphi) = 1\} = \Phi_j \tag{D.2}$$

であるような帰納的分類規則を表す写像

$$f_j : \Phi \rightarrow \{s \mid 0 \leqq s \leqq 1\} \tag{D.3}$$

を想定すれば、axiom 2を満たす類似度関数

$$SM(\cdot, \omega_j) : \Phi \times \{\omega_j\} \rightarrow \{s \mid 0 \leqq s \leqq 1\} \tag{D.4}$$

について、包含関係

$$SM(\Phi, \omega_j)_1 \equiv \{\varphi \in \Phi \mid SM(\varphi, \omega_j) = 1\} \subseteq f_j(\Phi)_1 \tag{D.5}$$

が成り立ち、

写像 $SM(\cdot, \omega_j)$ の拡張が f_j である (D.6)

ということになる。このように、類概念 $\mathfrak{C}_j$ の内包とは、パターン集合領域 Φ から真理値(truth-value)の集まりとしての非負単位区間 $I\,[0, 1] \equiv \{s \mid 0 \leqq s \leqq 1\}$ への写像 f_j であり、

$1 \in I\,[0, 1]$ は真(truth)を表し、$0 \in I\,[0, 1]$ は偽(falsity)を表し、$(0<)\,s\,(<1)$ は中間的な真理値(the middle truth-value)を表す (D.7)

と解釈可能な“関数 f_j に真理値 1 を与えるパターン φ の集合領域”が類概念 $\mathfrak{C}_j$ の外延である。

時刻変数 $t \in \{0, 1, 2, \cdots\}$ を考え、認識システムRECOGNITRON［B3］,［B4］の構造が時刻 t と共に変容していくものとしよう。但し、ここでは、モデル構成作用素 T, 類似度関数 SM, 大分類関数 BSC の3構造成分の内、axiom 2を満たす類似度関数 SM のみが、

$$SM_t : \Phi \times \Omega \rightarrow \{s \mid 0 \leqq s \leqq 1\} \tag{D.8}$$

と変容するものとする。axiom 2を満たす類似度関数 SM_t を採用している認識システム RECOGNITRON(t)が時刻 t において $\varphi_t \in \Phi$ を認識処理したとしよう。このとき、次の(イ),(ロ),(ハ)の3つの量 RTN_t, β, α を導入する。(ハ)の $\alpha(\varphi_t, j)$ はSS理論では、パターン φ_t の、2カテゴリ分類容易度 $EOC(\varphi_t, \underline{\mathfrak{C}}[j] ; SM)$ と呼ばれている量である(式(4.2)を参照)：

(イ)認識結果の2分類 $RTN_t(\varphi_t, j) =$

$$\begin{cases} +1 \cdots \text{パターン } \varphi_t \in \Phi \text{ が第 } j \in J \text{ 番目のカテゴリ} \\ \qquad \mathfrak{C}_j \text{ に帰属すると、認識システム} \\ \qquad \text{RECOGNITRON(t) によって推断されたとき} \\ -1 \cdots \text{推断されなかったとき} \end{cases} \tag{D.9}$$

(ロ) $\beta(\varphi_t, j) =$

$$\begin{cases} +1 \cdots \text{パターン } \varphi_t \in \Phi \text{ が第 } j \in J \text{ 番目のカテゴリ} \\ \qquad \mathfrak{C}_j \text{ に一意的に帰属するとき、つまり、} \varphi_t \in \Phi_j \text{ のとき} \\ -1 \cdots \text{その他のとき} \end{cases} \tag{D.10}$$

(ハ) $\alpha(\varphi_t, j)$

$$\begin{cases} \equiv \log_e [SM(\varphi_t, \omega_j) / \{1 - SM(\varphi_t, \omega_j)\}] = \\ +\infty \quad \text{if } SM(\varphi_t, \omega_j) = 1 \\ 0 \quad \text{if } SM(\varphi_t, \omega_j) = 1/2 \\ -\infty \quad \text{if } SM(\varphi_t, \omega_j) = 0 \end{cases} \tag{D.11}$$

□

認識システム RECOGNITRON(t)がパターン φ_t を認識処理したことを経験を反映させ、時刻 t の SM_t を SM_{t+1} へと更新する方式は次の①, ②のように述べられる。

先ず、不等式

$$0 < \delta_0(j, t) \leqq 2^{-1}, \; 2^{-1} \leqq \delta_1(j, t) < 1 \tag{D.12}$$

を満たす2つの助変数 δ_0, δ_1 を設けておく。

単調増大性 $\delta_0(j, t) \leqq \delta_0(j, t+1)$

単調減少性 $\delta_1(j, t) \geqq \delta_1(j, t+1)$

for any $t \in \{0, 1, 2, \cdots\}$ (D.13)

が望ましい。

［SMt を SM_{t+1} へと更新する方式］

①$SMt(\varphi_t, \omega_j) \leqq \delta_0(j, t) \vee SMt(\varphi_t, \omega_j) \geqq \delta_1(j, t)$ のとき

SMt を更新しないで、

$SM_{t+1}(\varphi, \omega_j) = SM_t(\varphi, \omega_j)$

for any $j \in J$ and any $\varphi \in \Phi$ (D.14)

とおく。

② $\delta_0(j, t) < SM_t(\varphi_t, \omega_j) < \delta_1(j, t)$ のとき

SM_t を更新し、

$$SM_{t+1}(\varphi,\ \omega_j)=$$
$$SM_t(\varphi,\ \omega_j)\cdot\exp[-\alpha(\varphi_t,j)\cdot\beta(\varphi_t,j)\cdot RTN_t(\varphi_t,j)]$$
$$/\left[\sum_{i\in J} SM_t(\varphi,\ \omega_i)\cdot\exp[-\alpha(\varphi_t,i)\cdot\beta(\varphi_t,i)\cdot RTN_t(\varphi_t,i)]\right]$$
$$\text{for any } j\in J \text{ and any } \varphi\in\Phi \tag{D.15}$$

とおく。 □

更新式(D.15)において、$\alpha(\varphi_t, j)$の定義式(D.11)を考慮すれば、

$$-\alpha(\varphi_t, j)=DOC(\varphi_t, \underline{\mathfrak{C}}[j]\ ;\ SM) \tag{D.16}$$

であることが2カテゴリ分類困難度 DOC の2式(3.71), (4.1)から明らかである。

このとき、次の定理D1が成立し、関数 SM_{t+1} を RECOGNITRON(t+1)内の類似度関数として採用できることがわかる。

[定理D1] **(類似度関数 SM の再帰的構成定理)**

上述の①, ②で得られた関数

$$SM_{t+1}:\Phi\times\Omega\to\{s\mid 0\leqq s\leqq 1\} \tag{D.17}$$

はaxiom 2を満たす。

(証明) ①の場合、SMt はaxiom 2を満たすことから明らか。

②の場合を考えよう。

axiom 2の(i)(正規直交性)の成立は、

$$SMt(\varphi,\ \omega_j)=1 \text{ if } i=j,\ =0 \text{ if } i\neq j \tag{D.18}$$

を考慮すれば、容易に確かめることができる。

axiom 2の(ii)(確率性)の成立は定義式(D.15)より明らかである。

axiom 2の(iii)(T-不変性)の成立は SMt がT-不変性を満たすことから明らかである。 □

上述の更新方式が理にかなっているどうか検討しよう。

$$\alpha(\varphi_t, j)$$
$$\begin{cases} >0 \text{ if } 2^{-1}<SM_t(\varphi_t,\ \omega_j)<1 \\ =0 \text{ if } SM_t(\varphi_t,\ \omega_j)=2^{-1} \\ <0 \text{ if } 0<SM_t(\varphi,\ \omega_j)<2^{-1} \end{cases} \tag{D.19}$$

に注意すれば、次の(一), (二)が結論できる。

(一) $\beta(\varphi_t, j)=RTN_t(\varphi_t, j)$が成立する場合

$$\exp[-\alpha(\varphi_t,j)\cdot\beta(\varphi_t,j)\cdot RTN_t(\varphi_t,j)]$$
$$=\exp[-\alpha(\varphi_t,j)]$$
$$<1 \text{ if } 2^{-1}<SM_t(\varphi_t,\ \omega_j)<1 \tag{D.20}$$
$$=1 \text{ if } SM_t(\varphi_t,\ \omega_j)=2^{-1} \tag{D.21}$$
$$>1 \text{ if } 0<SM_t(\varphi_t,\ \omega_j)<2^{-1} \tag{D.22}$$

が成立する。

式(D.20)の成立は、更新式(D.15)の分子の、$SMt(\varphi,\ \omega_j)$に乗じられる係数 $\exp[-\alpha(\varphi_t,j)\cdot\beta(\varphi_t,j)\cdot RTN_t(\varphi_t,j)]$ は、パターン φ_t がRECOGNITRON(t)で正認識されるならば、$SM_{t+1}(\cdot,\ \omega_j)$を $SM_t(\cdot,\ \omega_j)$より減少させる効果を持っている係数となっているから、式(D.20)の成立は、RECOGNITRON(t)で正認識されているパターン φ_tが RECOGNITRON(t+1)では誤認識される可能性が生じてしまう。

また、式(D.21)の成立は、RECOGNITRON(t)で正認識されているパターン φ_t が

RECOGNITRON(t+1)でも正認識される可能性を保つ。

最後に、式(D.22)の成立は、RECOGNITRON(t)で正認識されているパターンφ_tが一層強くRECOGNITRON(t+1)では正認識される可能性が生じる。

(二) $\beta(\varphi_t, j) \neq RTN_t(\varphi_t, j)$が成立する場合

$$\exp[-\alpha(\varphi_t, j)\cdot\beta(\varphi_t, j)\cdot RTN_t(\varphi_t, j)]$$
$$=\exp[\alpha(\varphi_t, j)]$$
$$\begin{cases} >1 \text{ if } 2^{-1}<SM_t(\varphi_t,\ \omega_j)<1 & \text{(D.23)} \\ =1 \text{ if } SM_t(\varphi_t,\ \omega_j)=2^{-1} & \text{(D.24)} \\ <1 \text{ if } 0<SM_t(\varphi_t,\ \omega_j)<2^{-1} & \text{(D.25)} \end{cases}$$

が成立する。

式(D.23)の成立は、更新式(D.15)の分子の、$SM_t(\varphi,\ \omega_j)$に乗じられる係数$\exp[-\alpha(\varphi_t, j)\cdot\beta(\varphi_t, j)\cdot RTN_t(\varphi_t, j)]$ は、パターンφ_tがRECOGNITRON(t)で誤認識されるならば、$SM_{t+1}(\cdot,\ \omega_j)$を $SM_t(\cdot,\ \omega_j)$より増加させる効果を持っている係数となっているから、式(D.23)の成立は、RECOGNITRON(t)で誤認識されているパターンφ_tが RECOGNITRON(t+1)では正認識される可能性が生じる。

また、式(D.24)の成立は、RECOGNITRON(t)で誤認識されているパターンφ_tがRECOGNITRON(t+1)でも誤認識される可能性を保つ。

最後に、式(D.25)の成立は、RECOGNITRON(t)で誤認識されているパターンφ_tが一層強くRECOGNITRON(t+1)でも誤正認識される可能性が生じる。

以上、2カテゴリ分類容易度 EOC を使って、帰納的分類規則を表す類似度関数 SM を、認識システムRECOGNITRON(t)が認識したパターン φ_tにつき適応的に更新するその都度逐次的学習する方式が研究された。

付録E. Kullback-Leiblerの情報量を平均化したものとしての、平均相互情報量の、下からの評価

本付録Eでは、先ず、Kullback-Leiblerの情報量KLIが非負であり、その次に、その下限を評価し、最後に確率分布の差の絶対値の、上からの評価がこの情報量の関数として与えられることを証明する。結果として、平均相互情報量 AMI の、下からの評価を評価可能なことが示される。最後の章において、シャノン (Shannon)情報理論の基本諸量が解説される。

E1. Kullback-Leiblerの情報量 KLI の3性質

先ず、2つの離散確率分布 $\underline{x}\equiv\{x_k\}_{k\in K}$, $\underline{y}\equiv\{y_k\}_{k\in K}$ の実数値関数

$$KLI(\underline{x}\ ;\ \underline{y})\equiv\sum_{k\in K} x_k\cdot\log_e\ [x_k/y_k]$$
$$=-\sum_{k\in K} x_k\cdot\log_e y_k-\{-\sum_{k\in K} x_k\cdot\log_e x_k\} \tag{E.1}$$

は、自然対数量 $\log_e[x_k/y_k]$, $k\in K$の、確率分布 $\underline{x}$ による平均値として定義されており、**Kullback-Leiblerの情報量**(amount of information suggested by Kullback-Leibler)といわれているものである[B4]。確率分布 $\underline{x}$ が、真の確率分布 $\underline{y}$ からどの程度乖離しているかを測るモノサシとして、使用することの出来る情報量である。事実、以下の式(E.43)では、平均相互情報量 AMI(A, B)が、

Kullback-Leiblerの情報量 $KLI(\{p(a_j/b_k)\}_{j=1\sim n};\{p_1(a_j)\}_{j=1\sim n})$の、各出力記号 b_k の確率分布 $p_2(b_k)$, $k=1\sim m$ を重みとした凸結合として、式(E.42)によって表現されている。

次の補助定理E1を証明しよう。

［補助定理E1］（**Kullback-Leiblerの情報量の3性質**）

2つの実数列 $\underline{x}\equiv\{x_k\}_{k\in K}$, $\underline{y}\equiv\{y_k\}_{k\in K}$ が確率条件

$$[\forall k\in K,\ 0\leqq x_k\leqq 1]\wedge\sum_{k\in K}x_k=1 \tag{E.2}$$

$$[\forall k\in K,\ 0\leqq y_k\leqq 1]\wedge\sum_{k\in K}y_k=1 \tag{E.3}$$

を満たし、

$$\forall k\in K,\ x_k\to 0\Leftrightarrow y_k\to 0 \tag{E.4}$$

も成立するとする。

そうすると、次の3事項(ｉ), (ii), (iii)が成り立つ。

(ｉ)(**非負性**)不等式

$$KLI(x;y)\geqq 0 \tag{E.5}$$

が成り立ち、等号、つまり、$KLI(\underline{x};\underline{y})=0$ は、

$$\forall k\in K,\ x_k=y_k \tag{E.6}$$

の時に限る。

(ii)(**KLI の下からの評価**)

$$KLI(\underline{x};\underline{y}) \geqq \log_e[1-\{2^{-1}\cdot\sum_{k\in K}|x_k-y_k|\}^2]^{-1} \tag{E.7}$$

が成り立ち、等号は、式(E.6)の成立する時に限る。

(iii)(**2つの確率分布 $\underline{x}$, $\underline{y}$ の差の絶対値の、KLI の指数関数による上からの評価**)

$$\sum_{k\in K}|x_k-y_k| \leqq 2\cdot[1-\exp[-KLI(\underline{x};\underline{y})]]^{1/2} \tag{E.8}$$

が成り立ち、等号は、式(E.6)の成立する時に限る。

(証明)　確率密度を持つ連続確率分布の場合は文献［A13］で証明されている。同様な証明法を用いて、以下で少し丁寧に証明する。

$$0\cdot\log_e 0\equiv\lim_{x\to 0}x\cdot\log_e x=0 \tag{E.9}$$

であり、式(E.4)より

$$\lim_{x_k\to 0}x_k\cdot\log_e[x_k/y_k]=0 \tag{E.10}$$

が成り立つから、必要ならば、

$$x_k=0,\ y_\ell=0 \tag{E.11}$$

なる x_k, y_ℓをすべて取り除いておいて、改めて、

$$\forall k\in K'(\subseteq K),\ x_k>0(\Leftrightarrow y_k>0) \tag{E.12}$$

を仮定しておいてよい。以下では、K′ を改めて、K と表現する。

Jesen inequality によれば、$d^2\psi(u)/du^2\leqq 0$ なる1実変数 u の任意実数値関数 $\psi(u)$と、任意の実数値関数 $f(x)\ (0\leqq x\leqq 1)$について、不等式

$$\sum_{k\in K}\psi(f(x_k))\cdot x_k\leqq\psi(\sum_{k\in K}f(x_k)\cdot x_k) \tag{E.13}$$

が成り立つ。

ここで、

$$\psi(u) = \log_e u \tag{E.14}$$

$$(d\psi(u)/du = u^{-1},\ d^2\psi(u)/du^2 = -u^{-2} \leqq 0)$$

$$f(x_k) = y_k/x_k,\ k \in K \tag{E.15}$$

とおけば、

$$\sum_{k \in K} x_k \cdot \log_e [y_k/x_k] \ (= -\mathrm{KLI}(\underline{x}\ ;\ \underline{y}))$$

$$\leqq \log_e [\sum_{k \in K} (y_k/x_k) \cdot x_k] \tag{E.16}$$

$$= \log_e [\sum_{k \in K} y_k]$$

$$= \log_e 1 \quad \because \quad 式(E.3)$$

$$= 0 \tag{E.17}$$

ここで、等号は、式(E.6)の成立する時に限ることは、式(E.2)からわかる。よって、(E.17)の両辺に −1 をかけて、式(E.5)を得、(i)の成立がわかった。

次に、(ii)を証明しよう。

$$K^{+} \equiv \{k \in K \mid y_k/x_k < 1\} \tag{E.18}$$

$$K^{0} \equiv \{k \in K \mid y_k/x_k = 1\} \tag{E.19}$$

$$K^{-} \equiv \{k \in K \mid y_k/x_k > 1\} \tag{E.20}$$

とおくと、

$$K = K^{+} \cup K^{0} \cup K^{-},\ K^{p} \cap K^{q} = \phi\ (p \neq q) \tag{E.21}$$

が成り立ち、

$$\sum_{k \in K} x_k \cdot \log_e [y_k/x_k] \ (= -\mathrm{KLI}(\underline{x}\ ;\ \underline{y}))$$

$$= \sum_{k \in K^0} x_k \cdot \log_e [y_k/x_k]$$

$$+ \sum_{k \in K^-} x_k \cdot \log_e [y_k/x_k]$$

$$+ \sum_{k \in K^+} x_k \cdot \log_e [y_k/x_k] \tag{E.22}$$

と分解すれば、3表現

$$① \sum_{k \in K^0} x_k \cdot \log_e [y_k/x_k]$$

$$= \sum_{k \in K^0} x_k \cdot \log_e 1 = 0 \tag{E.23}$$

$$② \sum_{k \in K^-} x_k \cdot \log_e [y_k/x_k]$$

$$= \sum_{k \in K^-} x_k \cdot \log_e \min\{y_k/x_k,\ 1\}$$

$$= \sum_{k \in K} x_k \cdot \log_e \min\{y_k/x_k,\ 1\} \tag{E.24}$$

$$③ \sum_{k \in K^+} x_k \cdot \log_e [y_k/x_k]$$

$$= \sum_{k \in K^+} x_k \cdot \log_e \max\{y_k/x_k,\ 1\}$$

$$= \sum_{k \in K} x_k \cdot \log_e \max\{y_k/x_k,\ 1\} \tag{E.25}$$

が成り立つから、この①, ②, ③を式(E.22)に代入すれば、

$$-\mathrm{KLI}(\underline{x}\ ;\ \underline{y})$$

$$= \sum_{k \in K} x_k \cdot \log_e \min\{y_k/x_k,\ 1\}$$

$$+ \sum_{k \in K} x_k \cdot \log_e \max\{y_k/x_k,\ 1\} \tag{E.26}$$

が成立する。この式(E.26)にJesen inequalityを適用して、

$$-\mathrm{KLI}(\underline{x}\ ;\ \underline{y})$$

$$\leqq \log_e [\sum_{k \in K} x_k \cdot \min\{y_k/x_k,\ 1\}]$$

$$+\log_e\left[\sum_{k\in K} x_k\cdot\max\{y_k/x_k,\ 1\}\right] \tag{E.27}$$

$$=\log_e\left[\sum_{k\in K}\min\{y_k,\ x_k\}\right]$$

$$+\log_e\left[\sum_{k\in K}\max\{y_k,\ x_k\}\right]$$

$$=\log_e\left\{\left[\sum_{k\in K}\min\{y_k,\ x_k\}\right]\cdot\left[\sum_{k\in K}\max\{y_k,\ x_k\}\right]\right\} \tag{E.28}$$

が得られる。ここで、任意の2実数 a, b について成り立つ2等式

$$\min\{a,\ b\}=2^{-1}\cdot[a+b-|a-b|] \tag{E.29}$$

$$\max\{a,\ b\}=2^{-1}\cdot[a+b+|a-b|] \tag{E.30}$$

を用いて、変形すれば、

$$-\mathrm{KLI}(\underline{x}\ ;\ \underline{y})$$

$$\leqq\log_e\left\{\sum_{k\in K}2^{-1}\cdot[y_k+x_k-|x_k-y_k|]\cdot\sum_{k\in K}2^{-1}\cdot[y_k+x_k+|x_k-y_k|]\right\}$$

$$=\log_e\left\{\left[1-2^{-1}\cdot\sum_{k\in K}|x_k-y_k|\right]\cdot\left[1+2^{-1}\cdot\sum_{k\in K}|x_k-y_k|\right]\right\}$$

$$=\log_e\left[1-\left\{2^{-1}\cdot\sum_{k\in K}|x_k-y_k|\right\}^2\right] \tag{E.31}$$

を得、示された。

式(E.31)の両辺に -1 をかけて、式(E.7)を得、(ii)の成立がわかった。尚、式(E.31)での等号は、式(E.2)を考慮すれば式(E.27)で等号の成立する時に限ることがわかり、よって、式(E.6)の成立する時に限ることになる。

最後に、(iii)を証明しよう。

式(E.7)から、不等式

$$\exp[-\mathrm{KLI}(\underline{x}\ ;\ \underline{y})]$$

$$\leqq 1-\left\{2^{-1}\cdot\sum_{k\in K}|x_k-y_k|\right\}^2 \tag{E.32}$$

を得、この式(E.32)から容易に不等式(E.8)が導かれ、等号の成立は式(E.6)の成立する時に限ることも(ii)の証明からわかる。 □

E2. 平均相互情報量 AMI(A, B)の、下からの評価

上述の補助定理E1を適用すれば、不等式(E.45)で平均相互情報量 AMI(A, B)を下から評価する次の定理 E1 が得られる。

入力アルファベット $A\equiv\{a_j\mid j=1\sim n\}$, 出力アルファベット $B\equiv\{b_k\mid k=1\sim n\}$ を考え、3つの確率分布

$$p_1(a_j),\ j=1\sim n \tag{E.33}$$

$$p(a_j/b_k),\ j=1\sim n\ \text{ for any }\ k\in\{1,2,\cdots,m\} \tag{E.34}$$

$$p_2(b_k),\ k=1\sim m \tag{E.35}$$

を導入する。特に、第 $k\in\{1,2,\cdots,m\}$ 番目の出力 b_k が与えられたときの、第 $j\in\{1,2,\cdots,n\}$ 番目の入力 a_j が出現する条件付き確率 $p(a_j/b_k)$ について

$$\sum_{j=1}^{n}p(a_j/b_k)=1\ \text{ for any }\ k\in\{1,2,\cdots,m\} \tag{E.36}$$

が成立していることに注意しておく。

［定理E1］（**平均相互情報量 AMI (A, B)の下からの評価定理**）

第 $k \in \{1, 2, \cdots, m\}$ 番目の出力 b_k が与えられたときの、入力アルファベット A の条件付き平均情報量(曖昧度；equivocation)

$$H(A/b_k) \equiv \sum_{j=1}^{n} p(a_j/b_k) \cdot \log_e [p(a_j/b_k) \tag{E.37}$$

を含む KLI について、下からの評価

$$KLI(\{p(a_j/b_k)\}_{j=1\sim n}; \{p_1(a_j)\}_{j=1\sim n}) =$$

$$-\sum_{j=1}^{n} p(a_j/b_k) \cdot \log_e p_1(a_j) - H(A/b_k) \tag{E.38}$$

$$= \sum_{j=1}^{n} p(a_j/b_k) \cdot \log_e [p(a_j/b_k)/p_1(a_j)] \tag{E.39}$$

$$\geq \log_e [1 - \{2^{-1} \cdot \sum_{j=1}^{n} | p(a_j/b_k) - p_1(a_j) | \}^2]^{-1} \quad \text{for any } k \in \{1, 2, \cdots, m\} \tag{E.40}$$

が成立し、等号は、

$$\forall j \in J, p(a_j/b_k) = p_1(a_j) \tag{E.41}$$

の時に限る。入力アルファベット A，出力アルファベット B 間の平均相互情報量 AMI(A, B)について、下界の評価

$$AMI(A, B) \equiv \sum_{k=1}^{m} p_2(b_k) \cdot [-\sum_{j=1}^{n} p(a_j/b_k) \cdot \log_e p_1(a_j) - H(A/b_k)] \tag{E.42}$$

$$= \sum_{k=1}^{m} p_2(b_k) \cdot KLI(\{p(a_j/b_k)\}_{j=1\sim n}; \{p_1(a_j)\}_{j=1\sim n}) \tag{E.43}$$

$$= \sum_{k=1}^{m} p_2(b_k) \cdot \sum_{j=1}^{n} p(a_j/b_k) \cdot \log_e [p(a_j/b_k)/p_1(a_j)] \tag{E.44}$$

$$\geq \sum_{k=1}^{m} p_2(b_k) \cdot \log_e [1 - \{2^{-1} \cdot \sum_{j=1}^{n} | p(a_j/b_k) - p_1(a_j) | \}^2]^{-1} \tag{E.45}$$

が成り立ち、等号は、式(E.39)が成立する時に限る。

（証明） 不等式(E.40)、並びに、その等号の成立については、補助定理E1の(ii)を適用したものであり、これから、不等式(E.45)、並びに、その等号の成立については、明らかである。 □

E3. シャノン情報理論の基本諸量

結合確率分布(joint probability distribution)

$$p(a_j, b_k) = p(a_j/b_k) \cdot p_2(b_k) \tag{E.46}$$

を導入して、元来、**平均相互情報量**(average mutual information content) AMI(A, B)は、

$$AMI(A, B) = \sum_{j=1}^{n} \sum_{k=1}^{m} p(a_j, b_k) \cdot \log_e [p(a_j, b_k)/\{p_1(a_j) \cdot p_2(b_k)\}] \tag{E.47}$$

と定義されるものであり、式(E.42)による定義と一致する。その対称性

$$\mathrm{AMI}(A, B) = \mathrm{AMI}(B, A) \tag{E.48}$$

が成立することが直ちにわかる。

結合確率 $p(a_j, b_k)$ については、表現

$$p(a_j, b_k) = p(b_k/a_j) \cdot p_1(a_j) \tag{E.49}$$

も成立し、**周辺確率**(marginal probability)と呼ばれる2つの確率 $p_1(a_j)$, $p_2(b_k)$ は、

$$p_1(a_j) = \sum_{k=1}^{m} p(a_j, b_k),\ \ j = 1 \sim n \tag{E.50}$$

$$p_2(b_k) = \sum_{j=1}^{n} p(a_j, b_k),\ \ k = 1 \sim m \tag{E.51}$$

と表され、

$$p(b_k/a_j)$$

$$= p(a_j, b_k)/p_1(a_j) \tag{E.52}$$

$$= p(a_j, b_k)/\sum_{k=1}^{m} p(a_j, b_k),$$

$$j = 1 \sim n,\ k = 1 \sim m \tag{E.53}$$

とも表現される。

曖昧度(equivocation)と称される非負量

$$H(A/B)$$

$$= -\sum_{j=1}^{n} \sum_{k=1}^{m} p(a_j, b_k) \cdot \log_e p(a_j/b_k) \tag{E.54}$$

を導入し、

1つの出力 $b \in B$ を受信した結果平均何個の入力

$a \in A$ に対応するか

を概算すれば、そのその個数はほぼ、exp $[H(A/B)]$ である。

散布度(dissemination)と称される非負量

$$H(B/A)$$

$$= -\sum_{j=1}^{n} \sum_{k=1}^{m} p(a_j, b_k) \cdot \log_e p(b_k/a_j) \tag{E.55}$$

を導入し、

1つの入力 $a \in A$ を送信した結果平均何個の出力

$b \in B$ に対応するか

を概算すれば、その個数はほぼ、exp $[H(B/A)]$ である。

どの入出力対 $\langle a, b \rangle \in A \times B$ が出現するかに関する平均的な不確定さ(uncertainty)を表す**結合エントロピー**(joint entropy) $H(A, B)$

$$= -\sum_{j=1}^{n} \sum_{k=1}^{m} p(a_j, b_k) \cdot \log_e p(a_j, b_k) \tag{E.56}$$

を導入し、

送受信にあたり、入出力対 $\langle a, b \rangle \in A \times B$ が平均的に出現する個数

を概算すれば、その個数はほぼ、exp $[H(A, B)]$ である。

どの入力 $a \in A$ が出現するかに関する平均的な不確定さを表す入力エントロピー

$$H(A)$$

$$= -\sum_{j=1}^{n} p_1(a_j) \cdot \log_e p_1(a_j) \tag{E.57}$$

を導入し、送信される入力信号の個数を概算すれば、その個数はほぼ、exp $[H(A)]$ である。

どの出力 $b \in B$ が出現するかに関する平均的な不確定さを表す**出力エントロピー**

$$H(B) = -\sum_{k=1}^{m} p_2(b_k) \cdot \log_e p_2(b_k) \tag{E.58}$$

を導入し、受信される出力信号の個数を概算すれば、その個数ははほぼ、exp [H(B)] である。

2つの公式

$$H(A, B) = H(A) + H(B/A) \tag{E.59}$$

$$H(A, B) = H(B) + H(A/B) \tag{E.60}$$

の成立が容易に証明される。

伝送速度(transmission rate) R(A, B) と称される非負量

$$R(A, B) = H(B) - H(B/A) \tag{E.61}$$

は、送信することによって、B について取り去られた平均的不確定さの程度であり、R(A, B) について成り立つ今1つの表現

$$R(A, B) = H(A) - H(A/B) \tag{E.62}$$

を考慮すれば、R(A, B) は受信することによって、A について取り去られた平均的不確定さの程度でもある。

実は、伝送速度 R(A, B) とは平均相互情報量 AMI(A, B) のことであり、公式

$$R(A, B) = AMI(A, B) \tag{E.63}$$

が成り立つ。

Kullback-Leiblerの情報量

$$\begin{aligned} & KLI(\{\{p(b_k/a_j)\}_{k=1\sim m}\},\ \{p_2(b_k)\}_{k=1\sim m}\}) \\ & = \sum_{k=1}^{m} p(b_k/a_j) \cdot \log_e [p(b_k/a_j)/p_2(b_k)] \\ & = -\sum_{k=1}^{m} p(b_k/a_j) \cdot \log_e p_2(b_k) \\ & -[-\sum_{k=1}^{m} p(b_k/a_j) \cdot \log_e [p(b_k/a_j)]] \end{aligned} \tag{E.64}$$

を導入すれば、平均相互情報量 AMI(A, B) は、

$$\begin{aligned} & AMI(A, B) \\ & = \sum_{j=1}^{n} p_1(a_j) \cdot \\ & KLI(\{\{p(b_k/a_j)\}_{k=1\sim m}\},\ \{p_2(b_k)\}_{k=1\sim m}\}) \end{aligned} \tag{E.65}$$

と表現される。

今1つのKullback-Leiblerの情報量

$$\begin{aligned} & KLI(\{p(a_j/b_k)\}_{j=1\sim n}\ ;\ \{p_1(a_j)\}_{j=1\sim n}) \\ & = \sum_{j=1}^{n} p(a_j/b_k) \cdot \log_e [p(a_j/b_k)/p_1(a_j)] \\ & = -\sum_{j=1}^{n} p(a_j/b_k) \cdot \log_e p_1(a_j) \\ & -[-\sum_{j=1}^{n} p(a_j/b_k) \cdot \log_e [p(a_j/b_k)]] \end{aligned} \tag{E.66}$$

を導入すれば、平均相互情報量 AMI(A, B) は、

$$\begin{aligned} & AMI(A, B) \\ & = \sum_{k=1}^{m} p_2(b_k) \cdot \\ & KLI(\{p(a_j/b_k)\}_{j=1\sim n}\ ;\ \{p_1(a_j)\}_{j=1\sim n}) \end{aligned} \tag{E.67}$$

と再表現される。

パターン認識において最大類似度認識法 [B3], [B4] と知られている**最大事後確率決定法則**(maximum a-posteriori-probability decion rule) では、b_ℓ が受信されたとき、入力番号

$$j = \mathrm{argmax}_{i=1\sim n}\ p(a_i/b_\ell) \tag{E.68}$$

を見つけ、a_jが送信されたと決定する。このときの誤決定確率は

$$\sum_{i=1(i\neq j)}^{n} p(a_i/b_\ell) \tag{E.69}$$

であるが、すべての b_k $(k=1\sim k)$ について平均化した値（**平均誤決定確率**；average error-probability）Pe は

$$Pe=\sum_{k=1}^{m} p_2(b_k)\cdot\sum_{i=1(i\neq j)}^{n} p(a_i/b_k) \tag{E.70}$$

と表される。このとき、曖昧度としてのエントロピー H(A/B)は、誤差エントロピー

$$H_{error}=-P_e\cdot\log_e P_e-(1-P_e)\cdot\log_e(1-P_e) \tag{E.71}$$

を導入すれば、R.M.Fanoの不等式

$$0\leqq H(A/B)\leqq H_{error}+P_e\cdot\log_e(n-1) \tag{E.72}$$

によって評価される。

$P_e\in\{0,1\}$ならば、$H_{error}=0$ であり、不等式

$$0\leqq H(A/B)\leqq P_e\cdot\log_e(n-1) \tag{E.73}$$

が成り立っている。

Fanoの不等式は次のように証明される。

式(E.1)を考慮し、補助定理E1の(ｉ)を適用すれば、不等式

$$\begin{aligned}&H(A/B)\\&=\sum_{k=1}^{m} p_2(b_k)\cdot\\&\Big[-\sum_{i=1}^{n} p(a_i/b_k)\cdot\log_e p(a_i/b_k)\Big]\\&\leqq\sum_{k=1}^{m} p_2(b_k)\cdot\\&\Big[-\sum_{i=1}^{n} p(a_i/b_k)\cdot\log_e q_{ik}\Big]\end{aligned} \tag{E.74}$$

が、確率条件

$$\sum_{i=1}^{n} q_{ik}=1 \tag{E.75}$$

を満たす任意の確率分布$\{q_{ik}\}_{i=1\sim n}$について成立するから、特に、

$$q_{ik}=\begin{cases}1-P_e & \text{if } i=j\\ P_e/(n-1) & \text{if } i\neq j\end{cases} \tag{E.76}$$

とおいて、更に変形すると、

$$\begin{aligned}&H(A/B)\\&\leqq\sum_{k=1}^{m} p_2(b_k)\cdot\Big[-\sum_{i=1(i\neq j)}^{n} p(a_i/b_k)\cdot\\&\log_e\{P_e/(n-1)\}\Big]\\&+\sum_{k=1}^{m} p_2(b_k)\cdot\big[-p(a_j/b_k)\cdot\log_e(1-P_e)\big]\end{aligned} \tag{E.78}$$

$$\begin{aligned}&=-Pe\cdot\log_e\{P_e/(n-1)\}\\&-\log_e(1-P_e)\cdot\sum_{k=1}^{m} p_2(b_k)\cdot p(a_j/b_k)\\&=-P_e\cdot\log_e\{P_e/(n-1)\}-[\log_e(1-P_e)]\cdot\sum_{k=1}^{m}\\&p_2(b_k)\cdot\Big[1-\sum_{i=1(i\neq j)}^{n} p(a_i/b_k)\Big]\end{aligned} \tag{E.79}$$

$$\begin{aligned}&=-P_e\cdot\log_e P_e+P_e\cdot\log_e(n-1)-[\log_e(1-P_e)]\cdot\\&(1-P_e)\end{aligned}$$

$=H_{error}+P_e\cdot\log_e(n-1)$
を得、証明が終わる。

付録F．平均認識歪みとしての、平均相互情報量AMI の意味、その下界の直接的な証明と、認識歪み関数 ρ の諸例

本付録Fでは、2カテゴリ平均相互情報量 AMI($\underline{\mathfrak{C}}[j]$；$\Phi[t_1, t_2]$)に平均認識歪みの意味を付与した場合、その下界を求め、その後に、本研究内容に具体性をもたらすため認識歪み関数 ρ を数例構成し、最後に、処理の対象としている問題のパターン $\varphi\in\Phi$ につき曖昧度としての条件付きエントロピー H($\underline{\mathfrak{C}}$/Tφ)を最大にするパターン情報処理を採用している認識システム RECOGNITRON［B3］，［B4］ が、結局、基本パターン集合領域 Φ_B が高々可算集合である場合、多カテゴリ平均相互情報量としてのシャノン相互情報量 I($\underline{\mathfrak{C}}$, Φ)を最小にしていることを指摘し、シャノン相互情報量をパターン認識分野に導入することの重要性が強調される。

F1．平均認識歪みとしての、2カテゴリ平均相互情報量 AMI($\underline{\mathfrak{C}}[j]$；$\Phi[t_1, t_2]$)の下界の直接的な証明

$\underline{\mathfrak{C}}[j]$と$\Phi[t_1, t_2]$ との間の、式(3.17)の2カテゴリ平均相互情報量 AMI($\underline{\mathfrak{C}}[j]$；$\Phi[t_1, t_2]$)は、axiom 2を満たす式(2.12)の類似度関数 SM を使用し、2つの条件付き $p(\mathfrak{C}_1[j]/T\varphi)$，$p(\mathfrak{C}_2[j]/T\varphi)$を2式(3.63)，(3.64)の如く設定し、式(3.17)で定義されている。この非負量 AMI($\underline{\mathfrak{C}}[j]$；$\Phi[t_1, t_2]$)は、

$$①\forall\varphi\in\Phi,\ \forall n\in\{1, 2\},$$

$$q(T\varphi)\cdot p(\mathfrak{C}_n[j]/T\varphi)$$

$$=p(\mathfrak{C}_n[j])\cdot p(T\varphi/\mathfrak{C}_n[j]) \tag{F.1}$$

$$②\forall\varphi\in\Phi,\ \sum_{n=1}^{2}p(\mathfrak{C}_n[j])\cdot p(T\varphi/\mathfrak{C}_n[j])$$

$$=q(T\varphi) \tag{F.2}$$

を考慮すれば、

$$\mathrm{AMI}(\underline{\mathfrak{C}}[j]\,;\,\Phi[t_1, t_2])$$

$$=\sum_{\varphi\in\Phi[t_1,t_2]}\sum_{n=1}^{2}p(\mathfrak{C}_n[j])\cdot p(T\varphi/\mathfrak{C}_n[j])$$

$$\cdot\log_e[p(T\varphi/\mathfrak{C}_n[j])/q(T\varphi)] \tag{F.3}$$

$$=-\sum_{\varphi\in\Phi[t_1,t_2]}q(T\varphi)\cdot\mathrm{loge}q(T\varphi)$$

$$-\sum_{n=1}^{2}p(\mathfrak{C}_n[j])\cdot$$

$$【-\sum_{\varphi\in\Phi[t_1,t_2]}p(T\varphi/\mathfrak{C}_n[j])$$

$$\cdot\log_e p(T\varphi/\mathfrak{C}_n[j])】 \tag{F.4}$$

と変形出来る。

カテゴリ $\mathfrak{C}_n[j]$ がパターンモデル Tφ として再生されるときの費用(**歪み測度**)

$$\rho(\mathfrak{C}_n[j],\ T\varphi)(\equiv \mathrm{dtr}(T\varphi/\mathfrak{C}_n[j]))\geqq 0 \tag{F.5}$$

を考え、その平均値(**平均歪み測度**)

$$D(\mathfrak{C}_n[j]/T\varphi\ ;\ n=1, 2)$$

$$\equiv\sum_{\varphi\in\Phi[t_1,t_2]}\sum_{n=1}^{2}p(\mathfrak{C}_n[j])$$

$$\cdot p(T\varphi/\mathfrak{C}_n[j])\cdot\mathrm{dtr}(T\varphi/\mathfrak{C}_n[j]) \tag{F.6}$$

を定義すれば、次の定理F1は、その下界が

$$\mathrm{LB}(\mathrm{AMI}\ ;\ \underline{\mathfrak{C}}[j],\ \Phi[t_1, t_2]) \equiv s\cdot D(\mathfrak{C}_n[j]/T\varphi\ ;\ n=1, 2) + \sum_{n=1}^{2} p(\mathfrak{C}_n[j])\cdot \log_e \lambda_n \tag{F.7}$$

であることを指摘している。

［定理F1］**(2カテゴリ平均相互情報量 AMI ($\underline{\mathfrak{C}}[j]$; $\Phi[t_1, t_2]$) の下界定理)**

2つの助変数 s, λ_n を 1 より大きくない非負条件

$$0 \leqq \sum_{n=1}^{2} p(\mathfrak{C}_n[j])\cdot \lambda_n \cdot \exp(s\cdot d_{tr}(T\varphi/\mathfrak{C}_n[j])) \leqq 1 \tag{F.8}$$

を満たすように選ぶことが出来るならば、不等式

$$\mathrm{AMI}(\underline{\mathfrak{C}}[j]\ ;\ \Phi[t_1, t_2]) \geqq \mathrm{LB}(\mathrm{AMI}\ ;\ \underline{\mathfrak{C}}[j],\ \Phi[t_1, t_2]) \tag{F.9}$$

が成立する。

（証明）

$$\begin{aligned}
&\mathrm{AMI}(\underline{\mathfrak{C}}[j]\ ;\ \Phi[t_1, t_2]) - s\cdot D(\mathfrak{C}_n[j]/T\varphi\ ;\ n=1, 2) - \sum_{n=1}^{2} p(\mathfrak{C}_n[j])\cdot \log_e \lambda_n \\
&= \Sigma_{\varphi\in\Phi[t_1, t_2]} \sum_{n=1}^{2} p(\mathfrak{C}_n[j])\cdot p(T\varphi/\mathfrak{C}_n[j]\cdot \log_e[p(T\varphi/\mathfrak{C}_n[j])/q(T\varphi)] \\
&\quad - s\cdot D(\mathfrak{C}_n[j]/T\varphi\ ;\ n=1, 2) - \sum_{n=1}^{2} p(\mathfrak{C}_n[j])\cdot \log_e \lambda_n \\
&\qquad \because\ \text{式(F.3)}
\end{aligned} \tag{F.10}$$

$$\begin{aligned}
&= -\Sigma_{\varphi\in\Phi[t_1, t_2]} \sum_{n=1}^{2} p(\mathfrak{C}_n[j])\cdot p(T\varphi/\mathfrak{C}_n[j])\cdot \log_e[q(T\varphi)/p(T\varphi/\mathfrak{C}_n[j])] \\
&\quad - \Sigma_{\varphi\in\Phi[t_1, t_2]} \sum_{n=1}^{2} p(\mathfrak{C}_n[j])\cdot p(T\varphi/\mathfrak{C}_n[j])\cdot s\cdot d_{tr}(T\varphi/\mathfrak{C}_n[j]) \\
&\quad - \Sigma_{\varphi\in\Phi[t_1, t_2]} \sum_{n=1}^{2} p(\mathfrak{C}_n[j])\cdot p(T\varphi/\mathfrak{C}_n[j])\cdot \mathrm{loge}\ \lambda_n \\
&\qquad \because\ \text{式(F.6)と、等式}\ \Sigma_{\varphi\in\Phi[t_1, t_2]}\ p(T\varphi/\mathfrak{C}_n[j]) = 1
\end{aligned} \tag{F.11}$$

$$\begin{aligned}
&= \Sigma_{\varphi\in\Phi[t_1, t_2]} \sum_{n=1}^{2} 【-p(\mathfrak{C}_n[j])\cdot p(T\varphi/\mathfrak{C}_n[j])\cdot \log_e[\{q(T\varphi)\cdot\lambda_n\cdot \exp(s\cdot \mathrm{dtr}(T\varphi/\mathfrak{C}_n[j]))\}/p(T\varphi/\mathfrak{C}_n[j])]】 \\
&\geqq \Sigma_{\varphi\in\Phi[t_1, t_2]} \sum_{n=1}^{2} 【-p(\mathfrak{C}_n[j])\cdot p(T\varphi/\mathfrak{C}_n[j])\cdot[\{q(T\varphi)\cdot\lambda_n\cdot \exp(s\cdot \mathrm{dtr}(T\varphi/\mathfrak{C}_n[j]))\}/p(T\varphi/\mathfrak{C}_n[j]) - 1] \\
&\qquad \because\ \log_e x \leqq x - 1\ \text{if}\ x > 0
\end{aligned} \tag{F.12}$$

$$= \Sigma_{\varphi\in\Phi[t_1, t_2]} \sum_{n=1}^{2} p(\mathfrak{C}_n[j])\cdot$$

$p(T\varphi/\mathfrak{C}_n[j])\cdot$
$[1-\{q(T\varphi)\cdot\lambda_n\cdot$
$\exp(s\cdot dtr(T\varphi/\mathfrak{C}_n[j]))\}/p(T\varphi/\mathfrak{C}_n[j])]$
$=1-\Sigma_{\varphi\in\Phi[t_1,t_2]}\ \sum_{n=1}^{2}\ p(\mathfrak{C}_n[j])\cdot$
$p(T\varphi/\mathfrak{C}_n[j])\cdot\{q(T\varphi)\cdot\lambda_n\cdot$
$\exp(s\cdot dtr(T\varphi/\mathfrak{C}_n[j]))\}/p(T\varphi/\mathfrak{C}_n[j])$
$\because\ \ \Sigma_{\varphi\in\Phi[t_1,t_2]}\ \sum_{n=1}^{2}\ p(\mathfrak{C}_n[j])\cdot$
$p(T\varphi/\mathfrak{C}_n[j])=1$ (F.13)
$=1-\Sigma_{\varphi\in\Phi[t_1,t_2]}\ q(T\varphi)\cdot\sum_{n=1}^{2}p(\mathfrak{C}_n[j])\cdot$
$\lambda_n\cdot\exp(s\cdot dtr(T\varphi/\mathfrak{C}_n[j]))$
$\geqq 0\quad\because$ 式(F.8) (F.14)
を得、証明が終わる。 □

ここで、次の問題を設定してみよう。

【条件付き確率 $p(T\varphi/\mathfrak{C}_n[j])$, $n\in\{1,2\}$ を変えた場合の、2カテゴリ平均相互情報量 $AMI(\mathfrak{C}_n[j];\Phi[t_1,t_2])$ の**最小化問題】**

助変数 s は非正値をとるものとする。

$D(\mathfrak{C}_n[j]/T\varphi;n=1,2)\leqq d$ (F.15)

を満たす条件付き確率

$p(T\varphi/\mathfrak{C}_n[j]),\ n\in\{1,2\}$ (F.16)

の内で、$AMI(\underline{\mathfrak{C}}[j];\Phi[t_1,t_2])$ を最小にするものを求めよ。 □

上記の最小化問題の解は、次のように求められる。

【上記の最小化問題の**解**】

$s\cdot D(\mathfrak{C}_n[j]/T\varphi;n=1,2)\geqq s\cdot d$
$\because\ \ s\leqq 0$ (F.17)

であるから、

$AMI(\underline{\mathfrak{C}}[j];\Phi[t_1,t_2])$
$\geqq s\cdot D(\mathfrak{C}_n[j]/T\varphi;n=1,2)$
$+\sum_{n=1}^{2}\ p(\mathfrak{C}_n[j])\cdot\log_e\lambda_n$ (F.18)
$\geqq s\cdot d$
$+\sum_{n=1}^{2}\ p(\mathfrak{C}_n[j])\cdot\log_e\lambda_n$
$\because\ \ s\leqq 0$ (F.19)

を得、

$D(\mathfrak{C}_n[j]/T\varphi;n=1,2)=d$ (F.20)

のとき、$AMI(\underline{\mathfrak{C}}[j];\Phi[t_1,t_2])$ は最小値になり、

$AMI(\mathfrak{C}_n[j];\Phi[t_1,t_2])$
$=s\cdot d$
$+\sum_{n=1}^{2}\ p(\mathfrak{C}_n[j])\cdot loge\ \lambda\ n$ (F.21)

を与える式(F.6)の条件付き確率 $p(T\varphi/\mathfrak{C}_n[j])$, $n\in\{1,2\}$ が求めるものである。 □

F2. 認識歪みの諸例

明らかに成立する2種類の条件付き確率 $p(T\varphi/\mathfrak{C}_n[j])$, $p(\mathfrak{C}_n[j]/T\varphi)$ 間の等式

③$\forall j \in J,\ \forall \varphi \in \Phi[t_1, t_2]$,

$$p(\mathfrak{C}_n[j]) \cdot p(T\varphi/\mathfrak{C}_n[j]) = q(T\varphi) \cdot p(\mathfrak{C}_n[j]/T\varphi) \tag{F.22}$$

に注意すれば、式(F.6)の平均歪み測度 $D(\mathfrak{C}_n[j]/T\varphi\ ;n=1, 2)$ は、

$$D(\mathfrak{C}_n[j]/T\varphi\ ;n=1, 2) = \Sigma_{\varphi \in \Phi[t_1, t_2]} \sum_{n=1}^{2} q(T\varphi) \cdot p(\mathfrak{C}_n[j]/T\varphi) \cdot dtr(T\varphi/\mathfrak{C}_n[j]) \tag{F.23}$$

と変形されるが、ここで、**相違度関数**(dissimilarity-measure function)

$$dmf: T\cdot\Phi \times T\cdot\Phi \rightarrow R^+ \tag{F.24}$$

を導入して、

$$dtr(T\varphi/\mathfrak{C}_n[j]) = \begin{cases} dmf(T\omega_j, T\varphi) & \text{if } n=1 \\ \max_{i \in J-\{j\}} dmf(T\omega_i, T\varphi) & \text{if } n=2 \end{cases} \tag{F.25}$$

と定義すれば、

$$D(\mathfrak{C}_n[j]/T\varphi\ ;n=1, 2) = \Sigma_{\varphi \in \Phi[t_1, t_2]} \left[q(T\varphi) \cdot SM(\varphi, \omega_j) \cdot dmf(T\varphi, T\omega_j) + q(T\varphi) \cdot [1 - SM(\varphi, \omega_j)] \cdot \max_{i \in J-\{j\}} dmf(T\varphi, T\omega_i) \right]$$
$$\because \quad 2式(3.63), (3.64) \tag{F.26}$$

と表される。

式(F.24)の相違度関数 dmf の4例を構成しておけば、次の①, ②, ③のようになる：

①(パターンモデル間ノルム距離 $\| T\varphi - T\omega_j \|$ を用いた例)

$$dmf(T\omega_j, T\varphi) = \| T\varphi - T\omega_j \|^2 / \max_{k \in J} \| T\varphi - T\omega_k \|^2 \tag{F.27}$$

②(パターンモデル間内積 $(T\varphi, T\omega_j)$ を用いた例)

$$dmf(T\omega_j, T\varphi) = 1 - |(T\varphi \cdot \| T\varphi \|, T\omega_j \cdot \| T\omega_j \|)|^2 \tag{F.28}$$

③(パターンモデルの1次従属係数 $d_j(\varphi)$ を用いた例)

$$dmf(T\omega_j, T\varphi) = 1 - |d_j(T\varphi)|^2 / \sum_{k \in J} d_k(T\varphi)|^2 \tag{F.29}$$

ここに、$T\omega_k, k \in J$ が1次独立な系であるとして、近似誤差 $T\varphi - \sum_{k \in J} d_j \cdot T\omega_j$ のノルム自乗

$$\| T\varphi - \sum_{k \in J} d_j \cdot T\omega_j \|^2 \tag{F.30}$$

を最小にする各1次結合係数 d_j を $d_j(\varphi)(=d_j)$ と表している [B3], [B4]。

正規直交性

$$d_j(T\omega_i) = 1 \text{ if } i=j, =0 \text{ if } i \neq j \tag{F.31}$$

と

T-不変性

$$\forall \varphi \in \Phi, \forall j \in J, d_j(T\varphi) = d_j(\varphi) \tag{F.32}$$

とが成立していることに注意しておく。

④(axiom 2を満たす類似度関数 SM を用いた例)

$$\mathrm{dmf}(T\omega_j,\ T\varphi)$$
$$=1-\mathrm{SM}(\varphi,\ \omega_j) \quad \text{(F.33)}$$
$$=1/ 【1+\exp[\mathrm{EOC}(\varphi,\ \underline{\mathfrak{C}}[j];\ \mathrm{SM})]】$$
$$\because\ \text{式}(4.5) \quad \text{(F.34)}$$

この $\rho(T\varphi,\ T\omega_j)$ は、3性質

$$\mathrm{EOC}(\varphi,\ \underline{\mathfrak{C}}[j];\ \mathrm{SM})]\to+\infty \Leftrightarrow \mathrm{dmf}(T\omega_j,\ T\varphi)\to 0$$
$$\mathrm{EOC}(\varphi,\ \underline{\mathfrak{C}}[j];\ \mathrm{SM})]\to 0 \Leftrightarrow \mathrm{dmf}(T\omega_j,\ T\varphi)\to 1/2$$
$$\mathrm{EOC}(\varphi,\ \underline{\mathfrak{C}}[j];\ \mathrm{SM})]\to-\infty \Leftrightarrow \mathrm{dmf}(T\omega_j,\ T\varphi)\to 1 \quad \text{(F.35)}$$

が成り立ち、歪み測度 $\rho(T\varphi,\ T\omega_j)$ の意味を説明している。

F3. パターン情報処理に関し、シャノン相互情報量を導入することの意義、考察

$t_2\to\infty$ に従って、多段階認識の容易度 $\mathrm{EOC}(\varphi,\ \underline{\mathfrak{C}}[j];\ \mathrm{SM})\equiv-\mathrm{DOC}(\varphi,\ \underline{\mathfrak{C}}[j];\ \mathrm{SM})=\log^e[\mathrm{SM}(\varphi,\ \omega_j)/\{1-\mathrm{SM}(\varphi,\ \omega_j)\}]$ の上界を評価するのには、従来の手法［A1］を使えばよい。統計的独立性を仮定すれば、認識歪み(文献［B20］の付録Gを参照)に関し、φhernoffによる誤り確率の評価法などはそのまま、素直に適用できるが、この種の適用は割愛された。

本研究では、各出力出現確率 $p(T\varphi)(\varphi\in\Phi[t_1,\ t_2])$ と、出力 $T\varphi$ が観測された条件の下で各入力 $\mathfrak{C}_n[j]$ の再現条件付き確率 $p(\mathfrak{C}_n[j]/T\varphi)(n\in\{1,\ 2\})$ とを与え、式(3.17)の $\mathrm{AMI}(\underline{\mathfrak{C}}[j];\ \Phi[t_1,\ t_2])$ の最大値を求めることに関連し、式(3.72)の2カテゴリ分類困難度 $\mathrm{DOC}(\varphi,\ \underline{\mathfrak{C}}[j];\ \mathrm{SM})$、式(4.2)の2カテゴリ分類容易度 $\mathrm{EO}\varphi(\varphi,\ \underline{\mathfrak{C}}[j];\ \mathrm{SM})$, 式(3.46)の 曖昧度 $H(\underline{\mathfrak{C}}[j]/T\varphi)$ に関する解析を展開した。

曖昧度 $H(\underline{\mathfrak{C}}[j]/T\varphi)$ の、$p(\mathfrak{C}_1[j]/T\varphi)$ に関する微係数を 2カテゴリ分類困難度 $\mathrm{DOC}(\varphi,\ \underline{\mathfrak{C}}[j];\ \mathrm{SM})$ と定義することから、解析が始まっている。

付録Eにおいて、

$$A=\mathfrak{C},\ b_k=T\varphi_k \quad \text{(F.36)}$$

の場合を考え、

$$p(\mathfrak{C}_j/T\varphi_k)=\mathrm{SM}(T\varphi_k,\ \omega_j),$$
$$j\in J,\ k\in\{t_1,\ t_1+1,\ \cdots,\ t_2\} \quad \text{(F.37)}$$

と選んでいる場合(式(3.63)を参照)、

認識情報量(recognition entropy)RGetpy(φ)［B4］が

式(E.43)内のKullback-Leiblerの情報量

$\mathrm{KLI}(\ \{p(aj/bk)\}\ j=1\sim n;\ \{p1(aj)\}\ j=1\sim n)$ に等しいこと

がわかり、従って、処理の対象とする問題の個々のパターン $\varphi\in\Phi$ に関し、$H(A/b_k)$ を最小にするような認識システム RECOGNITRON による多段階認識［B3］,［B4］は、多カテゴリ平均相互情報量 $I(A,\ B)$ を最小にしていることになる。

このように考えると、RECOGNITRON による多段階想起認識の働きは、認識歪みとしての多カテゴリ平均相互情報量 $I(A,\ B)$ (式(E.42)を参照)を最大にしている。

(鈴木昇一，文教大学·情報学部·情報システム学科，“文教大学·情報学部·情報研究 no.26”投稿論文，論文題目　2カテゴリ分類困難度の情報理論，投稿年月日　2001年8月13日(月))　(完了)