類似度関数の選定に関する適切さの検証法

鈴　木　昇一

A Method for a Test of a Statistical Hypothesis about Whether or Not a Selected Similarity-Measure is Proper

Shoichi Suzuki

あらまし

不動点探索形構造受精に関する多段階帰納推論を使ったパターン認識アルゴリズムを内蔵している認識システム RECOGNITRON においては、モデル構成作用素T、類似度関数SM、大分類関数BSCを各変換段階で用い、入力パターンの帰属するカテゴリを仮決定しながら、次の段階へと進むとき、訂正できる機能がある多段階構成がとられている。この多段階構成アルゴリズムは、処理の対象とする問題のパターンφについてのパターンモデルの列（知覚的記憶表彰の列）を再生しながら、“パターンのモデルとその帰属する可能性のあるカテゴリの番号のリストとの対として定義されるカテゴリ帰属知識”に関する構造受精変換の不動点として、最終的には、“その帰属するカテゴリの代表パターンのモデルとそのカテゴリ番号のみからなるリストとの対”を確保する認識手法である。

処理の対象としたパターンφの有限部分集合$\Phi_0(\subseteq \Phi)$に対する認識率が良好でない場合、その原因として、3構成要素 T, SM, BSC がΦ_0に関し、適切に選定されていなかったことが先ず、挙げられる。

本論文の主目的は、T, BSCが適切に選ばれている状況の下で、有限部分パターン集合Φ_0について、SMが適切に選定されているかどうかを赤池情報量基準AICを用いて判定する方法を、数理的に研究することである。本認識アルゴリズムがSS公理系で構築されており、従って、axiom 2を満たすSMが多数存在することを考慮すると、選ばれた1つのSMが適切かどかを検証するこのような方法は必要なものである。

キーワード

モデル構成作用素　類似度関数　大分類関数　カテゴリ帰属知識の不動点
構造受精変換　適合度検定　χ^2確率分布　赤池情報量基準

Abstract

The recognition system RECOGNITRON having a pattern-recognition algorithm of a multistage induction inference using structural fertilization transformation seems to seek for highly probable categories at a new stage determining temporary categories of an input pattern φ in question at an old stage.A given input pattern φ is transformed into a sequence of categorical membership-knowleges which is defined as a pair of a pattern-model and a list of its category-numbers obtained at each stage using model-construction operator T, similarity-measure function SM,and rough classifiers BSC, and at the final stage is reproduced as the first half(the model corresponding to the prototypical pattern $T\omega_j$ of the j-category $\mathfrak{C}_j$ to which φ may belong) of the fixed-point knowledge of a selected structural-fertilization transformation.

If a probability of misrecognition is low for a finite subset $\Phi_0 \subseteq \Phi$(a set of patterns to be recognized) , three fundamental constituents T, SM and BSC of RECOGNITRON are measures inadequate to the situation Φ_0.

The main purpose of this paper is to mathematically investigate a method of testing whether or not using AIC(the Akaike information criterion) SM is adequate to Φ_0 on the assumption that T and BSC were selected adequately.In consideration of that the recognition algorithm holds good under a system of SS-axioms and therefore there are many SMs satisfying axiom 2 of SS-axioms,such a method of testing the selected SM for a adequacy is necessary.

Key words model-construction operarors similarity-measure function rough classifier fixed point of categorical membership-knowledge structural-fertilization transformation test of goodness of fit χ^2-probability distribution Akaike information criterion

1.まえがき

処理の対象とする問題のパターン（入力パターン）φについて、第2章、式（11）のカテゴリ事前確率分布を多段階的に第2章、式(45)のカテゴリ**事後確率分布**に変換するのが、**不動点（探索形構造受精多段階帰納推理によるパターン）認識の働き**を備えている認識システム **RECOGNITRON**の情報処理機能である[36]，[37]。

カテゴリ番号の全集合をJとする。

不動点探索形構造受精多段階変換によって、初期のカテゴリ帰属知識 $\langle\varphi, J\rangle$ に関し、帰納推理を実行するパターン認識の働きを備えている認識システムRECOGNITRONは、モデル構成作用素T, 類似度関数SM, 大分類関数 BSC などで代表されるその持っている知識の範囲で、処理の対象とする問題のパターン（入力パターン）φについて、

各カテゴリ番号リスト$\mu \subseteq J$を各認識段階でその都度適切に選んで得られる

各構造受精変換 $TA(\mu)T$による変換

を、認識の初期段階のカテゴリ帰属知識 $\langle T\varphi, J\rangle$ に対し、カテゴリ帰属知識に関する**不動点方程式**

$$TA(\mu)T\langle\psi, \lambda\rangle =_{\triangle} \langle\psi, \lambda\rangle$$

が成立するまで繰り返し、φに対応する終局的な**カテゴリ極限分布（カテゴリ事後確率分布）**

$$p_j(\psi;\underline{\mathfrak{C}}), j \in J$$

を、推定しようとする。ここに、各 $p_j(\psi;\underline{\mathfrak{C}})$ は、不動点パターン ψ が第 $j\in J$ 番目のカテゴリ $\mathfrak{C}_j$ に帰属している程度（**帰属度**）である。

不動点認識の働きからもたらされる3認識結果 (i) 認識処理可能, (ii) 認識処理不能, (iii) 認識処理不定において、入力パターン φ に対応し想起される**不動点パターン** ψ と、φ に対応する終局的なカテゴリ極限分布がどのように表現されるかについては、文献［37］にその説明があるが (付録Gの定理14)、不動点探索形構造受精多段階帰納推理によるパターン認識の働きによる“**認識結果の分類定理**［37］”においては、(ii) の認識処理不能の場合を除いて、(iii) の認識処理不定の場合は、強制的に、カテゴリ番号

$$j=\mathrm{argmax}_{k\in\lambda}\ sk(\varphi;\underline{\mathfrak{C}})\in\lambda\in J$$

を求め、(i) の認識処理可能の場合のごとく、認識確定させることが出来ることに、注意しておく。

不動点探索形構造受精多段階パターン認識法で使われるの、類似度関数［36］,［37］SMが、処理の対象とする問題のパターン φ の集合 Φ に対し、適切に選ばれているかどうかのを検定（SMの適合度検定 (test of goodness of fit)）することを考えよう。モデル構成作用素［36］,［37］Tと、大分類関数 BSC とが適切に選ばれているものとし、SMが適切に選ばれているかどうかを、**仮説検定**してみよう。

本研究の主目的は、認識システムRECOGNITRONの認識性能を左右する類似度関数SMを評価するために、統計的な赤池情報量基準AIC を応用する数理を展開することである。

不動点探索形構造受精に関する多段階帰納推論を使ったパターン認識アルゴリズムを内蔵しているRECOGNITRONの3構成要素は、

①モデル構成作用素 T：$\Phi\rightarrow\Phi$

②類似度関数 SM：$\Phi\times\Omega\rightarrow\{s\mid 0\leqq s\leqq 1\}$

③大分類関数BSC：$\Phi\{0,1\}$

であり、この3構成要素 T,SM,BSC を用い、

④カテゴリ選択関数CSF：$\Phi\times 2^J\rightarrow 2^J$

⑤構造受精作用素 $A(\gamma)$：$\Phi\times 2J\rightarrow\Phi,\ \gamma\in 2^J$

が構成される。ここに、2^J は J のすべての部分集合（集合とリストとを同一視して、リストとして扱われることがある）のなす集合である。

処理の対象としたパターン集合 $\Phi_0(\subseteq\Phi)$ に対する認識率が良好でない場合、その原因として、3構成要素 T, SM, BSC が Φ_0 に関し、適切に選定されていなかったことが先ず、挙げられる。本論文は、主として、SM がパターンの有限部分集合 $\Phi_0\subseteq\Phi$ に関し、適切に選定されているかどうかを赤池情報量基準 AICを用いて、判定する手法（以下の（三））を提案するものである。

本論文では、RECOGNITRON での仮説検定として、次の（一）,（二）,（三）が研究される：

（一）不動点探索形構造受精多段階パターン認識法において、T, BSC が適切に選ばれているものとし、SM が適切に選定されているかどうかの ε-検定が説明される（3.2節）。

（二）認識システム RECOGNITRON の備わっているパターン認識の働きが正常に機能しているどうか、つまり、抜き取られた n 個のパターンが、対象とするパターン集合 Φ からの無作為標本であるどうかを、χ^2分布（カイ2乗分布）を使って、検定する手法が説明される（3.3.5節）。

（三）認識システム RECOGNITRON が採用した類似度関数 SM について正常に有限部分パターン集合 Φ_0 に関し機能しているどうかを、赤池情報量基準 AIC で判定する方法が研究される（4.5節）。

このような研究を行った諸論文はこれまで存在していないが、認識システムがSS公理系（文献［37］の4公理 axiom 1～4 のこと）で構築されている以上、T, BSCを選定・固定した状況下で、処理の対象とする問題のパターンの有限部分集合Φ_0に関し、SM が適切に選定されているかどうかを検証する手法は必要なものである。上記の（一）～（三）が実際にどの程度信頼性のある検証手法であるかどうかは、計算機シミュレーションを繰り返しみなければ判明しないのが残念である。

2.認識システム RECOGNITRON と、カテゴリ分布の変換、認識結果の分類

本章では、認識システム RECOGNITRON の性能を 赤池情報量基準 AIC を用いて評価するにあたって必要な前提としての“**カテゴリ分布の変換と認識結果の分類**”に関する説明がなされる。

2.1 事前カテゴリ分布

想定している**全カテゴリ集合**（a set of all categories）

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

に注目する。処理の対象とする問題のパターン（pattern in questionto to be recognized）φの集合 Φ は、或る可分なHilbert空間$\mathfrak{H}$の、零元 0 を含む或る部分集合である。正常なパターン$\varphi \in \Phi$は$\underline{\mathfrak{C}}$内の1つのカテゴリ、例えば、**第j∈J番目のカテゴリ**$\mathfrak{C}_j$に帰属しているものとされる。Φ内には、全く、どのカテゴリにも帰属しないパターン、2つ以上のカテゴリのどれにも帰属しているパターンなどの異常なパターンが存在しているかもしれない。

任意のパターン$\varphi \in \Phi$は事前に、共通して、事前**カテゴリ生起分布**（a priori probability of occurrences of any $\mathfrak{C}_j$; a priori categorical grade-of -membership distribution）

$$p(\mathfrak{C}_j),\ j \in J \tag{2}$$

を持っている。ここに、$\mathfrak{C}_j$の生起確率$p(\mathfrak{C}_j)$は、**確率条件**（probability condition）

$$[\forall j \in J, 0 < p(\mathfrak{C}_j) < 1] \wedge \sum_{j \in J} p(\mathfrak{C}_j) = 1 \tag{3}$$

を満たしている。

2.2 モデル構成作用素 T と、再帰領域方程式の解

パターンφの集合Φに対し文献［37］の付録Aで説明されているaxiom 1を満たすように、適切に選んだ**モデル構成作用素**（model-construction operator）

$$T: \Phi \rightarrow \Phi \tag{4}$$

を固定し、先ず、この式（4）の写像（モデル形式）T について、3性質

①$0 \in \Phi$

②$\forall \varphi \in \Phi, a \cdot \varphi \in \Phi$ for any positive real number a

③$\forall \varphi \in \Phi, T\varphi \in \Phi$

を満たすようなパターン集合Φを想定すれば、Φ_Bをパターンと判明しているφの基本集合（**基本領域**; basic domain）として、パターン集合 Φ は、**再帰領域方程式**（reflective domain equation）

$$\Phi = \Phi_B \cup T \cdot \Phi \cup R^{++} \cdot \Phi,$$

where

$$T \cdot \Phi \equiv \{T\varphi \mid \varphi \in \Phi\}$$

$$R^{++}\cdot\Phi \equiv \{a\cdot\varphi \mid \varphi\in\Phi,\ a\in R^{++}\,(\text{a set of positive real numbers})\} \tag{5}$$
を満たさなければならない。

再帰領域方程式（5）の解としてのパターン集合Φは、**初期条件**

$$\Phi(t)\mid_{t=0} = \Phi_B \ni 0 \tag{6}$$
の下で、**反復式**（an iterative equation）

$$\Phi(t+1) = \Phi_B \cup T\cdot\Phi(t) \cup R^{++}\cdot\Phi(t) \tag{7}$$
の解が明らかに、

$$\Phi = \lim_{t\to\infty}\Phi(t) \tag{8}$$
であるから、Tの満たすべきaxiom 1を使えば、

$$\Phi = R^{++}\cdot(\Phi_B \cup T\cdot\Phi_B) \tag{9}$$
と求められ、表示される。

2.3 類似度SMから定まるパターンの事前生起確率分布

さて、文献［37］の付録Bで説明されているaxiom 2を満たすように、類似度関数（similarity－measure function）

$$SM: \Phi\times\Omega \to \{s \mid 0\leqq s\leqq 1\} \tag{10}$$
を選ぶと、式（2）のカテゴリ生起確率分布より精密なφの事前生起確率分布（a prior distribution of occurrences of φ）

$$SM(\varphi,\omega_j),\ j\in J \tag{11}$$
が得られる。ここに、$\omega_j(\neq 0)\in\Phi$は第$j\in J$番目のカテゴリ$\mathfrak{C}_j$の持つ諸性質を典型的に所有している代表パターン（prototypical pattern）であり、その適応的決定法は文献［37］の付録Iで説明されている。式（1）での全カテゴリ集合$\underline{\mathfrak{C}}$に関し、**代表パターン集合**

$$\Omega \equiv \{\omega_j \mid j\in J\} \tag{12}$$
が導入されたことに注意しておく。Ω は1次独立でなければならないが、第$j\in J$番目のカテゴリ$\mathfrak{C}_j$は、典型としてのω_jを中心とした緩やかな類概念であることを仮定した状況に注意しておこう。

文献［37］の付録Bのaxiom 2の（ii）より、**確率条件**

$$[\forall j\in J, 0\leqq SM(\varphi,\omega_j)\leqq 1 \wedge \sum_{j\in J} SM(\varphi,\omega_j) = 1 \tag{13}$$
が成立しており、

SM(φ,ω_j) は、パターンφが第$j\in J$番目のカテゴリ$\mathfrak{C}_j$に
帰属している確率（φがω_jに似ている確率）である

と解釈される。

2.4 大分類関数BSC

大分類関数（binary-state classifier）と呼ばれる写像

$$BSC: \Phi\times J \to \{0,1\} \tag{14}$$
は、文献［55］の付録Cで説明されているaxiom 3を満たすように、選ばれていなければならない。その解釈として、

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

が採用される。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)

$$\forall j \in J, \forall i \in J-\{j\}, BSC(\omega_i, j)=0 \qquad (15)$$

を公理として要請していない事実に注意しておこう。

大分類関数 BSC は、付録Bでのaxiom 2を満たす**類似度関数 SM のカテゴリ抽出能力のあいまい性を2値化する形で、そのカテゴリ抽出能力を補う**ものである。

2.5 カテゴリ選択関数CSFと構造受精作用素 A(γ)

集合 2^J は全カテゴリ番号集合（a set of category-numbers） J のすべての部分集合の成す集合（ベキ集合；power set）である。

先ず、**カテゴリ選択関数**（category-selection function）

$$CSF: \Phi \times 2^J \rightarrow 2^J \qquad (16)$$

については、その定義から、

①$\varphi=0 \vee \gamma=\phi$ の場合、

$$CSF(\varphi, \gamma)=\phi \qquad (17)$$

②$\varphi \neq 0 \wedge \gamma \neq \phi$ の場合、

$$CSF(\varphi, \gamma)=\{k \in \gamma \mid [1+BSC(\varphi, k) - psn(\sum_{k \in \gamma} BSC(\varphi, k)-b)]\cdot SM(\varphi, \omega_k)>0\}. \qquad (18)$$

と表現されることが導かれている(文献[37]の、付録Eのaxiom 4の(i)、並びに、付録Eの定理E2)。

次に、候補カテゴリの番号のある部分集合 $\gamma \in 2^J$ を助変数とする"**構造受精作用素**"(structure-fertilization operaror）と呼ばれる写像

$$A(\gamma): \Phi \rightarrow \Phi \qquad (19)$$

については、実変数 u の、関数（positive-sign function）

$$psn(u) \equiv 0 \text{ if } u<0, \equiv 1 \text{ if } u \geqq 0 \qquad (20)$$

と、不等式

$$0<b \leqq 1 \qquad (21)$$

を満たす閾値（threshold value） b を用意する。その定義から、

①$\varphi=0 \vee \gamma=\phi$ の場合、

$$A(\gamma)\varphi=0 \qquad (22)$$

②$\varphi \neq 0 \wedge \gamma \neq \phi$ の場合、

$$A(\gamma)\varphi= \sum_{k \in \gamma}[1+BSC(\varphi, k)-psn(\sum_{k \in \gamma} BSC(\varphi, k)-b)]\cdot SM(\varphi, \omega_k)\cdot T\omega_k. \qquad (23)$$

であることが導かれている（文献［37］の、式（6.12)、並びに、付録Eの定理E2)。

2.6 カテゴリ帰属知識 ⟨φ, γ⟩

「パターン $\varphi \in \Phi$ が、式(1)の全カテゴリ集合 $\mathfrak{C}$ の部分集合

$$\underline{\mathfrak{C}}(\gamma) \equiv \{\mathfrak{C}_j \mid j \in \gamma\} \subseteq \underline{\mathfrak{C}} \qquad (24)$$

の何れか1つのカテゴリ $\mathfrak{C}_j$ に帰属する可能性がある」

という"パターン $\varphi \in \Phi$ の**カテゴリ帰属知識**（categorical membership-knowledge)"を、認識システム RECOGNITRON がパターン $\varphi \in \Phi$ に対し持っているとする。この知識を、

$$\langle \varphi, \gamma \rangle \in \langle \Phi, 2^J \rangle \qquad (25)$$

と表す。全カテゴリ番号集合$\gamma \in 2^J$は、要素（カテゴリ番号）を並べて得られるリスト（list）として、取り扱われている。具体的には、γは、

$$\gamma \equiv [j_1, j_2, \cdots, j_k] \in 2^J \tag{26}$$

はk個の要素$j_1, j_2, \cdots, j_k \in J$からなるリストを表し、集合$\{j_1, j_2, \cdots, j_k\}$と同一視することがある。特に、要素を1つも持たないリストを空集合ϕで表す。ここに、

$$\langle \Phi, 2^J \rangle \equiv \{\langle \varphi, \gamma \rangle \mid \varphi \in \Phi, \gamma \in 2^J\} \tag{27}$$

は、**カテゴリ帰属知識空間**（categorical membership-knowledge space）と呼ばれ、すべてのパターン$\varphi \in \Phi$と、すべてのカテゴリ番号集合$\gamma \in 2^J$との成す対であり、**対リスト**（pair list）と呼ばれることがある。

2.7　カテゴリ帰属知識間の等形式関係 $=_\triangle$ と構造受精変換 TA(μ) T

2つのカテゴリ帰属知識

$$\langle \varphi, \gamma \rangle, \langle \psi, \lambda \rangle \in \langle \Phi, 2^J \rangle \tag{28}$$

間の、2元関係（a binary relation on $\langle \Phi, 2^J \rangle$）としての、等形式関係（equi-form relation）$=_\triangle$は、次の定義1で与えられる。

［定義1］（カテゴリ帰属知識間の等形式関係）

$\langle \varphi, \gamma \rangle =_\triangle \langle \psi, \lambda \rangle$(恒等的に等しい)

$\Leftrightarrow \varphi = \psi \wedge \gamma = \lambda.$ □

さて、カテゴリ帰属知識$\langle \varphi, \gamma \rangle$から$\langle \psi, \lambda \rangle$への変換（**構造受精変換**；structure-fertilization transformation）

$$TA(\mu)T : \langle \Phi, 2^J \rangle \rightarrow \langle \Phi, 2^J \rangle \tag{29}$$

の定義は、次の定義2で与えられる。

［定義2］（**構造受精変換 TA**(μ)T）

$$TA(\mu)T\langle \varphi, \gamma \rangle =_\triangle \langle \psi, \lambda \rangle \tag{30}$$

,where

$$\psi \equiv TA(\mu)T\varphi \tag{31}$$

$$\lambda \equiv CSF(\varphi, \mu \cap \gamma) \tag{32}$$

と与えられる。

2.8　カテゴリ帰属知識間の等構造関係 =

$$\langle \varphi, \gamma \rangle =_\triangle \langle \psi, \lambda \rangle \Rightarrow \langle \varphi, \gamma \rangle = \langle \psi, \lambda \rangle \tag{33}$$

が成り立つという意味で、上の定義1の等形式関係$=_\triangle$より弱い**等構造関係**（a equi-structure relation）= を、次の定義3のごとく定義する。式（33）に示されているように、形式が同じであれば、構造も同じであるが、構造が同じだからいって、形式が同じとは限らないカテゴリ帰属知識$\langle \varphi, \gamma \rangle$が存在する事実に注意しておこう。

［定義3］（**カテゴリ帰属知識間の等構造関係**）

$\langle \varphi, \gamma \rangle = \langle \psi, \lambda \rangle$

$\Leftrightarrow CSF(\varphi, \gamma) = CSF(\psi, \lambda) \wedge$

$[\forall j \in CSF(\varphi, \gamma) \cap CSF(\psi, \lambda),$

$SM(\varphi, \omega_j) = SM(\psi, \omega_j)].$ □

2.9 カテゴリ帰属知識 $\langle \varphi, \gamma \rangle$ の直交分解（SS分解）

等構造関係 ＝ に関するカテゴリ帰属知識 $\langle \varphi, \gamma \rangle$ の直交分解（**SS分解**）について、説明しておこう。

2カテゴリ帰属知識 $\langle \varphi, \gamma \rangle$, $\langle \psi, \lambda \rangle$ 間の**内積** $(\langle \varphi, \gamma \rangle, \langle \psi, \lambda \rangle)$ を、

$$(\langle \varphi, \gamma \rangle, \langle \psi, \lambda \rangle) \equiv \Sigma_{j \in CSF(\varphi,\gamma) \cap CSF(\psi,\lambda)} \sqrt{SM(\varphi, \omega_j)} \cdot \sqrt{SM(\psi, \omega_j)} \tag{34}$$

と定義し、その**ノルム** $\| \langle \varphi, \gamma \rangle \|$ を、

$$\| \langle \varphi, \gamma \rangle \| \equiv \sqrt{(\langle \varphi, \gamma \rangle, \langle \varphi, \gamma \rangle)} \tag{35}$$

と定義すれば、カテゴリ帰属知識間の等構造関係 ＝ について、カテゴリ帰属知識 $\langle \varphi, \gamma \rangle$ の直交直和分解を与える次の定理1が成り立つ。パターン $\varphi \in \Phi$ の**カテゴリ依存構造**を明らかにするこのSS分解は、

$\langle \varphi, \gamma \rangle \in \langle \Phi, 2^J \rangle$ の**標準分解**（canonical decomposition）

とも呼ばれる。ここに、$s \overline{\in} S$ は、要素 s が集合 S の元であることの意である。

［定理1］（カテゴリ帰属知識の直交直和分解（標準分解）定理）[37]

(i) ($\langle \Phi, 2^J \rangle$ の**基底** $\langle \Omega, J \rangle$ の存在)

$$(\langle \omega_i, [i] \rangle, \langle \omega_j, [j] \rangle) = \begin{cases} 1 & \text{if } i = j \\ 0 & \text{if } i \neq j \end{cases} \tag{36}$$

が成り立ち、

$$\langle \Omega, J \rangle \equiv \{\langle \omega_j, [j] \rangle \mid j \in J\} \tag{37}$$

は、カテゴリ帰属知識空間 $\langle \Phi, 2^J \rangle$ の**完全正規直交系**（complete orthonormal system）である。

(ii)（**直交展開係数** $(\langle \varphi, \gamma \rangle, \langle \omega_j, [j] \rangle)$ の決定）

$$(\langle \varphi, \gamma \rangle, \langle \omega_j, [j] \rangle) = \begin{cases} \sqrt{SM(\varphi, \omega_j)} & \text{if } j \in CSF(\varphi, \gamma) \\ 0 & \text{if } j \overline{\in} CSF(\varphi, \gamma) \end{cases} \tag{38}$$

このとき、3写像 SM, BSC, γ について、**全射性** [37] が成立するという仮定の下で、

(iii)（**直交展開式**；SS展開）

$$\forall \langle \varphi, \gamma \rangle \in \langle \Phi, 2^J \rangle, \ \langle \varphi, \gamma \rangle$$
$$= \Sigma_{j \in J} (\langle \varphi, \gamma \rangle, \langle \omega_j, [j] \rangle) \cdot \langle \omega_j, [j] \rangle \tag{39}$$
$$= \Sigma_{j \in CSF(\varphi,\gamma)} \sqrt{SM(\varphi, \omega_j)} \cdot \langle \omega_j, [j] \rangle \tag{40}$$

(iv)（ノルム $\| \langle \varphi, \gamma \rangle \|$ の表現）

$$\forall \langle \varphi, \gamma \rangle \in \langle \Phi, 2^J \rangle, \ \| \langle \varphi, \gamma \rangle \|$$
$$= [\Sigma_{j \in J} | (\langle \varphi, \gamma \rangle, \langle \omega_j, [j] \rangle) |^2]^{1/2} \tag{41}$$
$$= [\Sigma_{j \in CSF(\varphi,\gamma)} SM(\varphi, \omega_j)]^{1/2} \tag{42}$$

が成り立つ。 □

2.10 不動点探索形構造受精多段階帰納推論に基づくパターン認識過程の終了条件としての不動点方程式と、それから得られる入力パターン φ のカテゴリ極限生起分布

処理の対象とする問題のパターン（入力パターン）φ について、式（E10）の事前確率分布を多段階的に以下の式（E44）の**事後確率分布**（a posterior distribution）に変換するのが、**不動点探索**

形構造受精多段階変換によるパターン認識の働きを備えている認識システム **RECOGNITRON** の情報処理機能である。

不動点探索形構造受精多段階変換によるパターン認識の働きを備えている認識システムRECOGNITRONは、モデル構成作用素 T, 類似度関数 SM, 大分類関数BSC などで代表されるその持っている知識の範囲で、処理の対象とする問題のパターン（入力パターン）φについて、事前に、カテゴリ帰属知識〈$T\varphi,\gamma$〉を持っているならば、

各カテゴリ番号リスト $\mu \subseteq J$ を各認識段階でその都度適切に選んで

得られる各構造受精変換 $TA(\mu)T$による変換 (43)

を、認識の初期段階のカテゴリ帰属知識〈$T\varphi,\gamma$〉に対し、**不動点方程式**

$TA(\mu)T\langle\psi,\lambda\rangle =_{\triangle} \langle\psi,\lambda\rangle$ (44)

が成立するまで繰り返し（**不動点探索形構造受精多段階帰納推論に基づくパターン認識過程**）、φに対応する終局的なカテゴリ極限生起分布（パターンφが与えられたときの、全カテゴリ集合$\underline{\mathfrak{C}}$の事後生起確率分布；a posteriori probability of occurrences of $\underline{\mathfrak{C}}$, given φ）

$p_j(\psi;\underline{\mathfrak{C}}), j\in J$ (45)

を、推定しようとする。各$p_j(\psi;\underline{\mathfrak{C}})$は、

$[\forall j\in J, 0\leqq p_j(\psi;\underline{\mathfrak{C}})\leqq 1 \wedge$

$\Sigma_{j\in J}\, p_j(\psi;\underline{\mathfrak{C}})=1$ (46)

を満たし、不動点パターン ψ が第$j\in J$番目のカテゴリ$\mathfrak{C}_j$に帰属している程度（**帰属度**; grade of membership）である。実は、$p_j(\psi;\underline{\mathfrak{C}})$ は、

$$p_j(\psi;\underline{\mathfrak{C}})=\begin{cases} SM(\psi,\omega_j)/\Sigma_{j\in\lambda}\, SM(\psi,\omega_j) & \cdots j\in\lambda \text{ のとき} \\ 0 & \cdots j\in J-\lambda \text{ のとき} \end{cases} \quad (47)$$

と求められる。

2.11　不動点認識の働きと、それから得られる知覚的記憶表象の列

不動点探索形構造受精多段階帰納推論に基づくパターン認識過程を説明しよう。

処理の対象とする問題のパターン（入力パターン）$\varphi\in\langle\Phi, 2^J\rangle$に関する認識過程とは14式 (50)～(63) のことである。そして、不動点方程式 (44) におけるψ,λとは、実は、不動点方程式 (62) での$\psi_t,\lambda_t\in\langle\Phi, 2^J\rangle$のことであり、パターン$\varphi$の帰属するカテゴリ候補に関する絞り込み性質

$J\supset\lambda_0\supset\lambda_1\supset\lambda_2\supset\cdots\supset\lambda_{t-2}\supset\lambda_{t-1}\supset\lambda_t$ (48)

を備えたカテゴリ番号の列

$\lambda_0, \lambda_1, \lambda_2, \cdots, \lambda_{t-2}, \lambda_{t-1}, \lambda_t\in 2^J$ (49)

を発見することである：

$\langle\psi_0, \lambda_0\rangle\in\langle\Phi, 2^J\rangle$

,where $\psi_0\equiv T\varphi, \lambda_0\equiv J$ (50)

$\to TA(\mu_0)T\langle\psi_0, \lambda_0\rangle =_{\triangle} \langle\psi_1, \lambda_1\rangle\in\langle\Phi, 2^J\rangle$ (51)

,where $\mu_0\in\langle\Phi, 2^J\rangle$ (52)

$\to TA(\mu_1)T\langle\psi_1, \lambda_1\rangle =_{\triangle} \langle\psi_2, \lambda_2\rangle\in\langle\Phi, 2^J\rangle$ (53)

,where $\mu_1\in\langle\Phi, 2^J\rangle$ (54)

$\rightarrow TA(\mu_2)\,T\langle \psi_2, \lambda_2\rangle =_{\triangle} \langle \psi_3, \lambda_3\rangle \in \langle \Phi, 2^J\rangle$ (55)

,where $\mu_2 \in \langle \Phi, 2^J\rangle$ (56)

$\rightarrow$

…

$\rightarrow TA(\mu_{t-2})\,T\langle \psi_{t-2}, \lambda_{t-2}\rangle$ (57)

$=_{\triangle} \langle \psi_{t-1}, \lambda_{t-1}\rangle \in \langle \Phi, 2^J\rangle$ (58)

,where $\mu_{t-2} \in \langle \Phi, 2^J\rangle$ (59)

$\rightarrow TA(\mu_{t-1})\,T\langle \psi_{t-1}, \lambda_{t-1}\rangle$

$=_{\triangle} \langle \psi_t, \lambda_t\rangle \in \langle \Phi, 2^J\rangle$ (60)

,where $\mu_{t-1} \in \langle \Phi, 2^J\rangle$ (61)

$\rightarrow TA(\mu_t)\,T\langle \psi_t, \lambda_t\rangle$

$(=_{\triangle} \langle \psi_{t+1}, \lambda_{t+1}\rangle) =_{\triangle} \langle \psi_t, \lambda_t\rangle \in \langle \Phi, 2^J\rangle$ **(fixed-point equation)** (62)

,where $\mu_t \in \langle \Phi, 2^J\rangle$ (63)

□

パターンφの帰属するカテゴリ候補に関する式 (48) の**絞り込み性質**を備えた14式 (50) ～ (63) のパターン認識過程には、パターンφを作業処理してから得られた短期記憶系（a system of short-term memory）での、次々と書き換えらたパターン列

$\psi_0, \psi_1, \psi_2, \cdots, \psi_{t-2}, \psi_{t-1}, \psi_t$ (62)

も得られており、この式（62）が多段帰納推論での**知覚的記憶表象**（perceptual and memorial representative）$\psi_s (0 \leqq s \leqq 1)$ の列である。入力パターンφは、不動点方程式 (63) の解 $\langle \psi_t, \lambda_t\rangle \in \langle \Phi, 2^J\rangle$ の前半 $\psi_t \in \Phi$ として再現され、入力パターンφの帰属するカテゴリは、後半$\lambda_t \in 2^J$に注目して、

φ belongs to one of the obtained set $\underline{\mathfrak{C}}(\lambda_t) = \{\mathfrak{C}_j \mid j \in \lambda_t\}$ (63)

と、断定される。

2.12 認識結果の分類

不動点方程式 (62) の$CSF(\psi_t, \mu_t \cap \lambda_t) \subseteq J$ はカテゴリ帰属知識 $\langle \psi_t, \mu_t \cap \lambda_t\rangle$ の帰属する有効なカテゴリ候補の番号リストであり、不動点方程式（62）が成立しているならば、構造受精変換 TA(μ)T の定義2から、パターンψ_tと候補カテゴリ番号リストλ_tに関する2つの不動点方程式

$TA(\mu_t \cap \lambda_t)\,T\psi_t = \psi_t$ (64)

$\wedge\, CSF(\psi_t, \mu_t \cap \lambda_t) = \lambda_t$ (65)

が成立している。

ところで、文献［37]，付録Gの定理G14と同様に、次の定理2が成り立つ。

［定理2]（**連想形認識不動点解の分類定理**）

入力パターンφから想起されるパターンは、不動点方程式（64）を満たすという意味での不動点パターンψ_tであり、前節で説明された不動点探索形構造受精多段階認識の、入力パターンφに関する結果は、

(i)（**認識確定**） $\lambda_t = [j]$ ならば、

φ belongs to the j-th category $\mathfrak{C}_j$ (66)

(ii)（**認識不能**） $\lambda = \phi$ ならば、

φ belongs by no means to one of the category set

$\underline{\mathfrak{C}} = \{\mathfrak{C}_j \mid j \in J\}$ (67)

(iii) **(認識不定)** $\lambda_t = [j_1\, j_2 \cdots j_k]$ ならば、

φ belongs to one of the category subset (68)

$\underline{\mathfrak{C}}\ (\lambda_t) = \{\mathfrak{C}_j \mid j \in \lambda_t\}$

と**分類**される。 □

上述の分類定理においては、(ii) の場合を除いて、(iii) の場合は、強制的に、カテゴリ番号

$j = \mathrm{argmax}_{\,i \in \lambda}\ p_i(\psi_t; \underline{\mathfrak{C}}) \in \lambda \in J$ (69)

を求め、式 (66) のごとく、認識確定させることが出来よう。ここに、argmax $_{a \in A}\cdots$は、変数 $a \in A$ に依存する量 … を $a \in A$ にわたって変動させた場合得られる最大値を与える $a \in A$ の内、最も若いものを指す。

3. χ^2分布を利用したSM-認識性能の仮説検定

本章では、仮説検定によって類似関数 SM の性能を評価する手法が研究される。

3.1 仮説検定による類似関数SMの性能評価

不動点探索形構造受精多段階パターン認識法で使われるの、式 (10) の類似度関数 SM が、処理の対象とする問題のパターン φ の集合 Φ に対し、適切に選ばれているかどうかのを検定 (SM の適合度検定 (test of goodness of fit)) することを考えよう。第2章の式 (4) のモデル構成作用素 T と、第2章の式 (14) の大分類関数 BSC とが適切に選ばれているものとし、SM が適切に選ばれているかどうかを、χ^2分布を利用して、**仮説検定** (test of statistical hypothesis) してみよう。

3.2 適切さの簡単な判定と、正認識率s_j

本節では、SM が適切に選定されているか、その選定適切さに関する簡単な判定法を指摘してみよう。

不動点探索形構造受精多段階パターン認識法において、T, BSC が適切に選ばれているものとし、SM が適切に選定されているかどうかの ε-検定は、次のように規定される：

Φ 内のすべてのパターンを認識してみて、カテゴリ $\mathfrak{C}_j$ に帰属している m_j 個のパターンの内、同一のカテゴリ $\mathfrak{C}_j$ に帰属すると実際に認識されたパターンの個数が $n_j\,(\leqq m_j)$ であると判明した場合、

$s_j \equiv n_j / m_j$ (1)

は Φ 内に第$j \in J$ 番目のカテゴリ $\mathfrak{C}_j$に帰属しているパターン φ が存在している割合を表すパターン集合Φについての**正認識率**であり、

$[\forall j \in J, 0 \leqq s_j \leqq 1] \wedge \sum_{j \in J} s_j \leqq |J|$

ここに、$|J|$ は集合 J に含まれている要素の総数 (2)

が成立していなければならないが、このとき、不等式

$\varepsilon > |J| - \sum_{j \in J} s_j \geqq 0$ (3)

が成立していれば、SM が適切に選定されている。 □

3.3 χ^2分布（カイ2乗分布）による仮説検定

前節に引き続き、パターン集合 Φ から抜き取られてきた n 個のパターンの内、第$j \in J$ 番目のカテゴリ $\mathfrak{C}_j$に帰属しているパターンが m_j 個存在する場合、

第 $i \in J$ 番目のカテゴリ $\mathfrak{C}_i$に帰属しているパターンが不動点探索形構造受精多段階パターン認識法を用いて、$\mathfrak{C}_j(i \neq j)$ に帰属すると誤認識されても、$\mathfrak{C}_j(i \neq j)$ に帰属すると認識されたパターンと数える (4)

と約束して、

$\mathfrak{C}_j$に帰属すると認識されたパターンの数が n_jである場合、
選定されているSMの適切さを検定する方法 (5)

を説明しよう。

3.3.1 危険率, 有意水準による仮説検定

危険率による仮説検定に関する次の規約において、危険率,有意水準（level of significance）と称される α の値は0.05とすることもあるが、0.01とすることも多い。

［統計的な仮説検定の、実践上の規約］

"母集団 P の分布（分布関数）F に関して、或る仮説 H を立てる。別に、危険率 α を定めておく。仮説 H の下における分布 F を定める確率変数 Xに対して、或る一定のの領域（棄却域, **危険域**; critical regin） CR を、

prob$\{X \in R\}$（Xが CR の元である確率）$= \alpha$ (6)

を定めておく。今1回の試行によって、標本値 xを得た場合に、次の2つの行動 (i), (ii) の何れか 1つを取る。

(i) $x \in CR$ ならば、我々は仮説 H を棄てる（正しい仮説 H を捨てることにより、誤りを犯すこの第1種の誤りを確率 α で生じる)。

(ii) $x \overline{\in} CR$ ならば、我々は仮説 H を棄てない、又は、仮説 H を採る（正しくない仮説 H を採択するという第2種の誤りが生じる)。" □

上の規約に従って、同じ種類の検定を十分多数回行ったならば、例えば、$\alpha = 0.01$とすると、我々が誤りを犯すのは（仮説 H が正しいのに、仮説 H を棄てると判断するのは)、大体100回に1回の割合であるということである。

3.3.2 χ^2分布（カイ2乗分布）

平均値 m, 分散 σ^2の正規分布を $N(m, \sigma^2)$ と表す。$N(m, \sigma^2)$ の確率密度 g(x)は、

$$g(x) = [2\pi\sigma^2]^{-1/2} \cdot \exp(-(x-m)^2/(2\sigma^2)) \quad (-\infty < x < +\infty) \tag{7}$$

である。

(1) 確率変数 X の分布が標準分布 N(0, 1) であれば、確率変数

$$Y = X^2 \tag{8}$$

の確率密度 $k^1(y)$ は、次の式 (9) で表される:

$$k_1(y) = \begin{cases} 0 & \cdots y \leqq 0 \text{ のとき} \\ [2\pi]^{-1/2} \cdot y^{-1/2} \cdot \exp(-y/2) & \cdots y > 0 \text{ のとき} \end{cases} \tag{9}$$

□

(2) n個の確率変数 $X_1, X_2, \cdots, X_n$ が独立で、同一の正規分布 $N(0,1)$ を持つならば、確率変数

$$Y = X_1^2 + X_2^2 + \cdots + X_n^2 \tag{10}$$

の確率密度 $k_n(y)$ は、次の式(12)で表される:

正の定数 C_2 は、

$$\int_0^{+\infty} dy\ k_n(y) = 1 \tag{11}$$

と定めると、

$$k_n(y) = \begin{cases} 0 & \cdots y \leqq 0 \text{ のとき} \\ C_2 \cdot y^{n/2-1} \cdot \exp(-y/2) & \cdots y > 0 \text{ のとき} \end{cases} \tag{12}$$

□

特に、

$$\int_0^{+\infty} dy\ y \cdot k_n(y) = n\text{(平均値)} \tag{13}$$

$$\int_0^{+\infty} dy\ (y-n)^2 \cdot k_n(y) = 2n\text{(分散)} \tag{14}$$

となる。確率密度 $k_n(y)$ を持つ分布を自由度 n の χ^2 分布という。

3.3.3 χ^2 分布による適合度検定

k個の排反なクラス

$$C_1, C_2, \cdots, C_k \tag{15}$$

を考える。1回の試み (trial) の結果起こり得るあらゆる場合が、式 (15) で表されるという訳である。これらは互いに排反し、式 (15) のどれか1つ C_i は必ず起こるものとする。

この試みを n 回行う。n回の試みのうち、式 (15) の $C_1, C_2, \cdots, C_k$ が実現された回数 (実測度数) を、各々、

$$n_1, n_2, \cdots, n_k \tag{16}$$

とする。ここに、等式

$$n_1 + n_2 + \cdots + n_k = n \tag{17}$$

が成立している。この n 個が対象とするパターン集合 Φ からの無作為標本であるとする。

第i (=1～n) のクラス C_i の起こり得る確率を p_i とし、

$$m_i = n \cdot p_i \tag{18}$$

とおく。m_i は C_i の起こり得る期待度数であって、実測度数 n_i の、期待度数 m_i からの差の度合いとして、

$$\chi_0^2 \equiv \sum_{i=1}^{k} (n_i - m_i)^2 / m_i \tag{19}$$

を考える。χ_0^2 は、各クラス C_i における期待度数 m_i と実測度数 n_i との差 $n_i - m_i$ の平方 $(n_i - m_i)^2$ の、期待度数 m_i に対する比 $(n_i - m_i)^2/m_i$ の総和であり、各差 $n_i - m_i$ が酷ければ、χ_0^2 も大きくなり、χ_0^2 の大きさの程度によって、仮説

$H_0: p_1, p_2, \cdots, p_n$ (理論分布と実測値から得た度数分布が一致しているという仮説) (20)

を棄却すればよい。

一般に、$F(x)$ を分布関数とする確率変数 X と、$F_i(x_i)$ を分布関数とする確率変数 X_i の列

$$X_1, X_2, \cdots, X_n \tag{21}$$

とがあって、任意の2実数 a,b について、

$$\int_a^b dF_i(x_i) \rightarrow \int_a^b dF(x) \tag{22}$$

が成立するとき（式（21）の確率変数列の分布が確率変数 X の分布に限りなく近づくとき）、式（21）の確率変数列は確率変数 X に**法則収束**するという。

［定理3］［17］**（χ_0^2 の分布の、χ^2 分布への法則収束定理）**

式（19）の確率変数 χ_0^2 の分布 F(y) は、$n \to \infty$ に対して、自由度 $k-1$ の χ^2 分布 G(y) に

$$\int_a^b dy\, F(y) \to \int_a^b dy\, G(y) \quad \text{for any } a, b\,(a \leqq b) \tag{23}$$

というように、法則収束する。 □

定理3の適用により、

$$k = |J| \tag{24}$$

とおくと、式(19) の確率変数 χ_0^2 は漸近的に、自由度 $|J|-1$ の χ^2 分布をなすことがわかった。

仮説 H_0 が正しいならば、

$$E(n_i)\,(n_i \text{の期待値}) = np_i \tag{24}$$

であるから、n 個の観測値の内、およそ

$$m_i = np_i \tag{25}$$

個は第 $i \in J$ 番目のカテゴリ $\mathfrak{C}_i$ に帰属するものであろう。従って、仮説 H_0 と実測度数値 n_i がよく合っているどうかの目安としては、n_i と m_i との差を用いることが出来る。

この立場から、実測値から計算した式（19）の χ_0^2 の値が大きいほど、仮説 H_0 は実測値と一致しないと見なされてよい。即ち、

［危険率 α での、χ^2 分布に関する仮説検定］

α を不等式

$$0 < \alpha < 1 \tag{26}$$

を満たすように設定し、棄却域として、

$$\chi_0^2 > t \tag{27}$$

を採用し、χ^2 分布の表から

$$\text{prob}\{\chi_0^2 > t \mid H_0\} = \int_t^\infty dy\, k_n(y)$$

（H_0 の下で、不等式 $\chi_0^2 > t$ が成立する確率） (28)

を満たす非負実数 t を求めて、

実測値から計算した χ_0^2 の値が t より大ならば、仮説 H_0 を捨て、この χ_0^2 の値が t より小ならば仮説 H_0 を捨てない。 □

かくして、危険率 α で、（仮に立てたが、先にいって誤っていると判定され棄却されるかも知れないという意味の）式（20）の**帰無仮説** (null hypothesis) H_0 の検定を行うことが出来る。

3.3.4 差の分布と多項分布

パターン φ の集合 Φ_j を、

Φ_j：第 $j \in J$ 番目のカテゴリ $\mathfrak{C}_j$ に帰属している
　パターン $\varphi \in \Phi$ の集合 (29)

と導入し、標本 $\cup_{j \in J} \Phi_j$ の大きさ (size) を n として、

$$s_j \equiv (1/n) \cdot \Sigma_{\varphi \in \Phi_j} SM(\varphi, \omega_j) \,/\, [(1/n) \cdot \sum_{j \in J} \Sigma_{\varphi \in \Phi_j} SM(\varphi, \omega_j)] \tag{30}$$

を定義すると、確率的性質

$$[\forall j\in J, 0\leqq s_j\leqq 1]\wedge\sum_{j\in J}s_j=1 \tag{31}$$
を満たす1つの組
$$\underline{s}=\{s_j\mid j\in J\} \tag{32}$$
が得られる。
$$m_j\equiv n\cdot s_j \tag{33}$$
を用いて、確率変数 X_j を、
$$X_j\equiv(n_j-m_j)/\sqrt{m_j} \tag{34}$$
と定義する。

確率変数 X_j の n 組 $\underline{X}=(X_1, X_2,\cdots,X_n)$ は、等式
$$\sum_{j\in J}\sqrt{s_j}\cdot X_j=0 \tag{35}$$
を満足する。即ち、点 $\underline{X}$ は超平面式（35）の上に乗っており、母集団（population）Φ から抽出（sampling）された標本（sample）$\cup_{j\in J}\Phi_j$ が乗っている。

式（35）の証明は、次の通りである：
$$\sum_{j\in J}n_j=n \tag{36}$$
$$\wedge\sum_{j\in J}m_j=n \tag{37}$$
であるから、
$$\begin{aligned}
&\sum_{j\in J}\sqrt{s_j}\cdot X_j\\
&=\sum_{j\in J}\sqrt{s_j}\cdot(n_j-m_j)/\sqrt{m_j}\\
&=\sum_{j\in J}\sqrt{s_j}\cdot(n_j-m_j)/\sqrt{n\cdot s_j} \qquad \because \text{式}(33)\\
&=(1/\sqrt{n})\cdot\sum_{j\in J}(n_j-m_j)\\
&=(1/\sqrt{n})\cdot[\sum_{j\in J}nj-\Sigma_{j\in J}m_j]\\
&=(1/\sqrt{n})\cdot[n-n]=0. \qquad \because 2\text{式}(36), (37)
\end{aligned}$$
□

さて、改めて、x_j, $\underline{x}$ を、
$$x_j=(n_j-ns_j)/\sqrt{ns_j} \tag{38}$$
$$\underline{x}=(x_1, x_2,\cdots,x_n) \tag{39}$$
と定義すると、

n 個のパターンの内、

カテゴリ $\mathfrak{C}_1$ に帰属するパターンが n_1 個

カテゴリ $\mathfrak{C}_2$ に帰属するパターンが n_2 個

……

カテゴリ $\mathfrak{C}_{|J|}$ に帰属するパターンが $n_{|J|}$ 個

であるような確率（n回の独立な試みにおいて、カテゴリ $\mathfrak{C}_1,\mathfrak{C}_2,\cdots,\mathfrak{C}_{|J|}$ に帰属するパターンが各々、丁度 $n_1, n_2,\cdots, n_{|J|}$ 回現れる確率） (40)

は、多項分布であり、
$$\begin{aligned}
&q_n(n_1, n_2,\cdots, n_{|J|})\\
&\equiv\{n!/(n_1!\cdot n_2!\cdots\cdots n_{|J|}!)\}\cdot s_1^{n1}\cdot s_2^{n2}\cdot\cdots\cdot s_{|J|}^{n|J|}
\end{aligned} \tag{41}$$
と表される。ここに、2式（31),（36）が成立している。

文献［17], 4.1節, 定理4.1（局所極限定理, p.111）の適用によって、式（41）は、
$$\begin{aligned}
&q_n(n_1, n_2,\cdots, n_{|J|})\\
&=p_n(\underline{x})
\end{aligned}$$

$$\equiv[1/\{(2\pi n)^{(|J|-1)/2}\cdot\sqrt{s_1\cdot s_2\cdot\cdots\cdot s_{|J|}}\}]\cdot$$
$$\exp(-(1/2)\cdot\sum_{j\in J}x_j^2)\cdot[1+c/\sqrt{n}] \qquad (42)$$

と評価される。ここに、c は $n\to\infty$ に対し有界である。

3.3.5 RECOGNITRON での仮説検定

認識システムRECOGNITRONの備わっているパターン認識の働きが正常に機能しているどうか、つまり、抜き取られた n 個のパターンが、対象とするパターン集合 Φ からの無作為標本であるどうかを検定してみよう。

式（30）のように定義される各 s_j を用いて、式（31）の確率的性質を満たす式（32）の $\underline{s}$ に注目する。

パターン集合 Φ から任意に n 個選ぶと、m_j 個は、第 $j\in J$ 番目のカテゴリ $\mathfrak{C}_j$ に帰属するパターンであったとする。

不動点探索形構造受精多段階パターン認識の働きによって、n 個のパターンの内（m_j個のパターンの内でないことに注意)、n_j 個が第 $j\in J$ 番目のカテゴリ $\mathfrak{C}_j$ に帰属すると、実際に認識されたとする。ここに、2式 (36), (37) が成立している。

式（74）を満たし、k を｜J｜と設定した式（19）の

$$\chi_0^2\equiv\sum_{j=1}^{|J|}(n_j-m_j)^2/m_j \qquad (43)$$

は、自由度 ｜J｜－1 の χ^2 分布に法則収束する。

そこで、仮説

$H_0: s_1, s_2, \cdots, s_{|J|}$

(理論上のカテゴリ分布 $<m_1, m_2, \cdots, m_{|J|}>$ と、不動点探索形構造受精多段階パターン認識の働きで実際に得られたカテゴリ分布 $<n_1, n_2, \cdots, n_{|J|}>$ とが一致している) (44)

を用意する。

$$\alpha=\mathrm{prob}\{\chi_0^2>t\mid H_0\} \qquad (45)$$

を満たす t を求めて、危険率 α で、

(i) $\chi_0^2>t$ ならば、仮説 H_0 を棄却する

(ii) $\chi_0^2\leqq t$ ならば、仮説 H_0 を棄却しない (46)

という検定を行う。

ここで、式 (33) の m_j が式 (30) の s_j を用いて、定義されていること、並びに、式（39）の Φ_j を考慮すると、

式（46）で仮説 H_0 を棄却しないと結論されたならば、

T, BSCを固定した条件の下で、SMが適切に選定されている (47)

と検定できる。

4.赤池情報量基準 AIC によるSM－認識性能の検定

本章では、認識システム RECOGNITRON が正常に機能しているどうかを、赤池情報量基準 AIC で判定する方法が研究される。先ず、最尤推定量について説明した後（4.1節)、多項分布の最尤推定量（4.2節)、並びに、最尤推定量の漸近正規性、一致性、漸近不偏性、漸近有効性（4.3節）についても説明する。その後、赤池情報量基準 AIC をK-L情報量と関連し説明し（4.4節)、最後

に、類似度関数 SM の、有限部分集合 Φ_0 についての適切さを、AIC を用いて判定する手法（4.5節）が研究される。

4.1 最尤推定量

確率密度を持つ連続確率分布の場合で、説明しよう。確率密度を持たない離散確率分布、例えば、4.2節の多項分布などについては、以下の記述において、確率密度の代りに、確率そのものを採用すればよい。

今後、x の関数はすべて（$-\infty, +\infty$）において区分的に連続とする。

$$I_1 \equiv \{x \mid -\infty < x < +\infty\} \tag{1}$$

$$I_2(j) \equiv \{\theta \mid \theta'(j) < \theta < \theta''(j)\},\ j = 1 \sim m \tag{2}$$

を定義すると、$x \in I_1$, $\underline{\theta} \equiv (\theta_1, \theta_2, \cdots, \theta_m) \in I_2$ で定義され、

$$\text{“}\underline{\theta} = (\theta_1, \theta_2, \cdots, \theta_m) \in I_2 \equiv I_2(1) \times I_2(2) \times \cdots \times I_2(m) \tag{3}$$

を助変数とする $x \in I_1$ の母集団確率密度関数

$$f(x, \underline{\theta})\text{”} \tag{4}$$

が与えられたとする。**大きさ n の標本**

$$\underline{x} = (x_1, x_2, \cdots, x_n) \tag{5}$$

を得たとすれば、一般に標本値がこの標本値$\underline{x}$を中心とする近傍の値を実現する確率は、

$$L(\underline{x}, \underline{\theta})\, dx_1 dx_2 \cdots dx_n \equiv f(x_1, \underline{\theta}) \cdot f(x_2, \underline{\theta}) \cdot \cdots \cdot f(xn, \underline{\theta})\, dx_1 dx_2 \cdots dx_n \tag{6}$$

によって与えられる。この式(6)の形を確率要素といい、係数

$$L(\underline{x}, \underline{\theta}) \equiv f(x_1, \underline{\theta}) \cdot f(x_2, \underline{\theta}) \cdot \cdots \cdot f(xn, \underline{\theta}) \tag{7}$$

を**尤度関数**（likelihood function）と呼ぶ。

母数$\underline{\theta}$を含まない統計量

$$\underline{\theta}^{\wedge} = \underline{\theta}^{\wedge}(x_1, x_2, \cdots, x_n) \tag{8}$$

を推定する点推定の最も重要な方法の1つは、フィッシヤーによって導入された**最尤法**（method of maximum likelihood）がある。

大きさ n の式（5）の標本 $x_1, x_2, \cdots, x_n$ に対して、式（7）の尤度関数 L（x, θ）は、標本値$\underline{x}$が1点の近傍をとる確率に比例するから、θ の真値$\underline{\theta*}$は L($\underline{x}$, $\underline{\theta}$）に大きな値を与える可能性を持つ。そこで、L($\underline{x}$, $\underline{\theta}$）を最大にする$\underline{\theta}$を推定量とみなす立場を取ろう。そのような式（6）で登場している$\underline{\theta}^{\wedge} = \underline{\theta}$は、**最尤方程式**と呼ばれる

$$\partial \log_e L(\underline{x}, \underline{\theta}) / \partial \underline{\theta} \ (= L(\underline{x}, \underline{\theta})^{-1} \cdot \partial L(\underline{x}, \underline{\theta}) / \partial \underline{\theta}) = 0 \tag{9}$$

を満たす。方程式（9）を満たす式（8）の最尤解 $\theta^{\wedge}$ は、**最尤推定量**（maximum likelihood estimate）と呼ばれる。

4.2 多項分布の最大対数尤度をもたらす最尤推定量

n 個の観測値

$$\{k_1, k_2, \cdots, k_n\} \tag{10}$$

が与えられたとき、c 個の事象

$e_1, e_2, \cdots, e_c$ (11)

が生起した度数を、各々、

$n_1, n_2, \cdots, n_c$ (12)

とする。ここに、

$k_i \in \{e_1, e_2, \cdots, e_c\}\ (i=1\sim n)$ (13)

であり、等式

$n_1+n_2+\cdots+n_c=n$ (14)

が成立している。

1より大きくない非負量 q_i を事象 e_i の生起確率とすると、多項分布

$\underline{q}=\{q_1, q_2, \cdots, q_c\}$ (15)

ここに、

$0 \leqq q_i \leqq 1\ (i=1\sim c) \wedge q_1+q_2+\cdots+q_c=1$ (16)

が定義され、その対数尤度

$\ell(\underline{q})$

$\equiv \log_e [n!/(n_1!\cdot n_2!\cdot \cdots n_c!)]\cdot q_1^{n1}\cdot q_2^{n2}\cdot \cdots \cdot q_c^{nc}$ (17)

は、

$\ell(\underline{q})$

$=\log_e [n!/(n_1!\cdot n_2!\cdot \cdots n_c!)]+\sum_{i=1}^{c} n_i\cdot \log_e q_i$ (18)

で、与えられる。

式（12）の度数分布が与えられたとき、パラメータ（助変数）q_i の組

$q_1, q_2, \cdots, q_c$ (19)

の最尤推定量 $q_i\hat{}$ の組

$q_1\hat{}, q_2\hat{}, \cdots, q_c\hat{}$ (20)

を、式（18）の $\ell(\underline{q})$ を最大とするように、求めよう。式（16）から得られる

$q_c=1-\sum_{i=1}^{c-1} q_i$ (21)

を、式（18）の $\ell(\underline{q})$ に代入すると、表現

$\ell(\underline{q})$

$=\log_e [n!/(n_1!\cdot n_2!\cdot \cdots n_c!)]$

$+\sum_{i=1}^{c-1} n_i\cdot \log_e q_i$

$+n_c\cdot \log_e (1-\sum_{i=1}^{c-1} q_i)$ (22)

が得られる。

$\underline{q}$ がこの式（22）の対数尤度 $\ell(\underline{q})$ を最大とするための必要条件は、

$\partial \ell(\underline{q})/\partial q_i$

$=n_i/q_i-n_c/[1-\sum_{i=1}^{c-1} q_i]=0\ (1 \leqq i \leqq n)$ (23)

である。この式（23）より、

$n_i/q_i=n_c/[1-\sum_{i=1}^{c-1} q_i]$

$=n_c/q_c\ (1 \leqq i \leqq n)$ (24)

が得られる。この式（24）から、

Kをある定数として、

$n_i/q_i = K\ (1 \leqq i \leqq n)$ (25)

が成り立つことがわかる。ここで、

$$
\begin{aligned}
n &= \sum_{i=1}^{c} n_i \qquad \because \ \text{式 (14)} \\
&= K \cdot \sum_{i=1}^{c} q_i \qquad (26) \\
&= K \quad \because \qquad \text{式 (16)}
\end{aligned}
$$

であることから、$K = n$ が判明し、最尤推定量 $q_i\hat{}$ として、

$q_i\hat{} = n_i/n\ (1 \leqq i \leqq n)$ (27)

が得られる。

結論として、最尤モデルは、

$q\hat{} = \{q_1\hat{}, q_2\hat{}, \cdots, q_c\hat{}\}$,

ここに、$q_i\hat{} = n_i/n\ (1 \leqq i \leqq n)$ (28)

であることがわかり、最大対数尤度は、

$$
\begin{aligned}
&\ell(\underline{q}) \\
&= \log_e[n!/(n_1! \cdot n_2! \cdot \cdots n_c!)] + \sum_{i=1}^{c} n_i \cdot \log_e n_i - n \cdot \log_e n \qquad (29)
\end{aligned}
$$

となる。

4.3 最尤推定量に関し知られている諸事実

最尤推定量 $\theta\hat{}$ について、次の事実が知られている（文献［19], p.41（3.3節最尤法, E節））

助変数 $\underline{\theta}$ の真値が $\underline{\theta*}$ である確率密度関数

$f(x, \underline{\theta*})$ (30)

を持つ統計的母集団

を考えよう。この母集団が大きいために、式（3）の標本 $\underline{x}$ を抜き出す過程において母集団の構成は変わらないと考え、その第 i（＝1～n）番目の実現値が x_i である確率変数 X_i の列

$\underline{X} = (X_1, X_2, \cdots, X_n)$ (31)

の各 X_i は、互いに独立で、同じ分布 $f(x, \underline{\theta})$ に従い、各 X_i の統計分布密度は $f(x_i, \underline{\theta})$ であると考えることができる。

$f(x, \underline{\theta})$ に関するある条件の下では、**最大対数尤度を与える最尤推定量**（対数尤度を最大とするような助変数を選ぶことによって（最尤法）、近似的には真の分布に近いモデル（最尤モデル）を得ようとして推定された推定量）$\underline{\theta\hat{}}$ は、大きな実験回数（母集団のサイズ）n に対しては、

平均値ベクトル $\underline{\theta*}$, 分散行列 $(1/n) \cdot J^{-1}$ の正規分布 $N(\underline{\theta*}, (1/n) \cdot J^{-1})$ (32)

に近似的に従う（漸近正規性）。但し、J は、その第i行第j列の成分が

$$
\begin{aligned}
&Ex(\partial \log_e f(x, \underline{\theta}) / \partial \theta_i \\
&\qquad \cdot \log_e f(x, \underline{\theta}) / \partial \theta_j) \mid_{\underline{\theta} = \underline{\theta*}} \qquad (33)
\end{aligned}
$$

で与えられる**Fisher行列**である。Ex（…）は … の、x に関する期待値である。これより、

$n \to \infty$ のとき、

（イ）（**一致性**）$\underline{\theta\hat{}}$ は真の値 $\underline{\theta*}$ に収束すること

（ロ）（**漸近不偏性**）$\underline{\theta*}$ の推定量として偏りが無くなること

（ハ）（**漸近有効性**）不偏推定量の中で最も分散が小さいこと

などが判明する。

以上で、最尤推定量としてパラメータを推定した後、初めて計算可能なAICを次の4.4節で導出する準備的考察が得られた。

4.4 赤池情報量基準 AIC

2つの関数 $g(x), f(x)$ を、

$g(x)$：真の確率分布の確率密度関数 (34)

$f(x)$：設定するモデルに対応する確率密度関数 (35)

とすると、**平均対数尤度**

$$\int_{-\infty}^{+\infty} dx\, g(x) \cdot \log_e f(x) \tag{36}$$

を大きくすれば、$g(x), f(x)$ の違いを表し、$g(x)$が $f(x)$ に一致すると零になる非負量としての**K-L情報量**（Kullback-Leibler Information）

$$I(g;f)$$

$$\equiv \int_{-\infty}^{+\infty} dx\, g(x) \cdot \log_e [g(x)/f(x)] \tag{37}$$

$$= \int_{-\infty}^{+\infty} dx\, g(x) \cdot \log_e g(x) - \int_{-\infty}^{+\infty} dx\, g(x) \cdot \log_e f(x) \geqq 0 \tag{38}$$

は小さくなる。

式（5）のn個の独立な観測値$\underline{x}$が得られると、式（7）で表されるその対数尤度

$$\sum_{i=1}^{n} \log_e f(x_i) \tag{39}$$

のn分の1

$$(1/n) \cdot \sum_{i=1}^{n} \log_e f(x_i) \tag{40}$$

で、**式（36）の平均対数尤度が近似される**。

よって、式（38）の符号に注意すると、

式（39）の対数尤度が大きいほど、設定したモデルは
真の分布の様相に近い (41)

という結論が得られる。

以上の事実などに注意すると、以下の式（44）は、最大対数尤度が同程度のモデルがある時、その中で実際に推定しなければならないパラメータの数が最も少ないものを選ぶべきであることを示しており、“**節約の原理**”の1つの具体化と言える。

複数個のモデルがある時、各モデルの善し悪しを評価する基準として、最尤モデルの平均対数尤度である式（39）の、式（5）のデータに関する期待値（**期待平均対数尤度**）を導入する。

期待対数尤度の値が大きいほどそのモデルは良いと言える。モデルの最大対数尤度を期待対数尤度の1つの推定量と考えることができるが、詳しく調べると、**最大対数尤度そのものは期待平均対数尤度の不偏推定量に他ならない**ことがわかる。

一般に、最大対数尤度は、期待平均対数尤度の本当の値に比べて大きく出やすいという**偏り**を持つ。この傾向はモデル内の自由パラメータの数が大きいほど著しい。これは、最大対数尤度の比較によってモデルを選択すると、自由パラメータの数の大きいモデルほど、選ばれやすいことを示している。

最大対数尤度の期待平均対数尤度に対する偏りの程度と、モデル内の自由パラメータの数との間の関係を調べると、

(モデルの最大対数尤度)－(モデル内の自由パラメータの数) (42)

が近似的に、**期待平均対数尤度の不偏推定量**となること（統計量である式（42）の期待値が期待平均対数尤度となること）が導かれる。歴史的経緯を考慮して、この式（42）を（−2）倍した量

$$\mathrm{AIC} \equiv (-2) \times (\text{期待平均対数尤度の不偏推定量}) \tag{43}$$

$$= (-2) \times [(\text{モデルの最大対数尤度}) - (\text{モデル内の自由パラータの数})]$$

$$\doteq (-2) \times (\text{モデルの最大対数尤度}) + 2 \times (\text{モデル内の自由パラータの数}) \tag{44}$$

がモデル選択の基準となる。

AICを最小とするモデル（最小AIC推定値;MAICE）が最適なモデルと考えられる。

式（44）のAICは、赤池情報量基準（Akaike information criterion）と呼ばれているものであり、

AIC

=(−2)×(モデルの最大対数尤度)+2×

(モデル内の、推定すべきパラメータの数) (45)

とも書ける。

4.5 AICによる類似度関数SMの適切さの判定

いよいよ、本数理的研究の主内容に入ろう。

4.5.1 有限な部分集合 Φ_0 と、正認識された"第 $j \in J$ 番目のカテゴリ $\mathfrak{C}_j$ に帰属するパターンの総数 $n(2, j)$

第 $j \in J$ 番目のカテゴリ $\mathfrak{C}_j$ は生起確率 $p(\mathfrak{C}_j)$ を持っているとする:

$$[\forall j \in J, 0 \leqq p(\mathfrak{C}_j) \leqq 1] \wedge \sum_{j \in J} p(\mathfrak{C}_j) = 1 \tag{46}$$

□

$|\Phi_0| \times p(\mathfrak{C}_j)$ を正整数になるように、処理の対象とする問題のパターン φ からなるパターン集合 Φ から、有限な部分集合 Φ_0 を選び、カテゴリ $\mathfrak{C}_j$ に帰属している"集合 Φ_0 内のパターンの個数"は、$|\Phi_0| \times p(\mathfrak{C}_j)$ に近い正整数でなければならない。$|\Phi_0|$ は集合 Φ_0 内の要素の総数である。

ここで、

$n(i, j)$: 第 $i \in I \equiv \{1, 2\}$ 番目の問題地域で、第 $j \in J$ 番目の
カテゴリ $\mathfrak{C}_j$ に帰属しているパターンの総数 (47)

を導入する。具体的には、$n(1, j)$ はあらかじめ判明している第 $j \in J$ 番目のカテゴリ $\mathfrak{C}_j$ に帰属している"パターン集合 Φ_0 内のパターン総数"であり、

$$n(1, j) \equiv |\Phi_0| \times p(\mathfrak{C}_j) \tag{48}$$

とおく。$n(2, j)$ は認識システムRECOGNITRONによって正認識された"第 $j \in J$ 番目のカテゴリ $\mathfrak{C}_j$ に帰属するパターンの総数"である。

更に、

$p(j/i)$: 第 $i \in I$ 番目の問題地域での、第 $j \in J$ 番目ののカテゴリ $\mathfrak{C}_j$ の存在確率 (49)

を導入し、また、

$n(i)$: 第 $i \in I \equiv \{1, 2\}$ 番目の問題地域でのパターン総数 (50)

とすると、

$$n(1) = |\Phi_0| \tag{51}$$

$$n(2) = [\text{RECOGNITRON によって正認識されたパターン } \varphi \in \Phi_0 \text{ の総数}] \leqq |\Phi_0| \tag{52}$$

である。

$$\sum_{j\in J} n(i,j) = n(i),\ i\in I \tag{53}$$
$$\sum_{j\in J} p(j/i) = 1,\ i\in I \tag{54}$$
が成立している。

4.5.2　適合度検定の設定

RECOGNITRONでの、適合度検定（test of goodness of fit）とは、

各 $j\in J$ につき、$n(2,j)$ が式（48）の $n(1,j)\equiv|\Phi_0|\times p(\mathfrak{C}_j)$ に比例していれば、
RECOGNITRONの認識機能は Φ_0 について正常に機能している　(55)

とみて、文献［19］の5.2節（pp.74-77）からhintを得て、AICが小さい値を取るほど、最適と考えられる判定法を、展開しよう。

4.5.3　MODEL(0) の構築

前項4.5.2で指摘した判定法を研究しよう。

多項分布

$$p(\{n(i,j)\}/p(j/i))$$
$$\equiv \prod_{i\in I}[\{n(i)!/\prod_{j\in J} n(i,j)!\}\cdot\prod_{j\in J} p(j/i)^{n(i,j)}] \tag{56}$$

を考える。

$$K\equiv\sum_{i\in I}\log_e[\{n(i)!/\prod_{j\in J} n(i,j)!\}] \tag{57}$$

とおくと、式(56)での $\{p(j/i)\}$ をパラメータとみなしたときの対数尤度 $\ell(\{p(j/i)\})$ は、

$$\ell(\{p(j/i)\})$$
$$=\log_e p(\{n(i,j)\}/p(j/i))$$
$$=K+\sum_{i\in I}\sum_{j\in J} n(i,j)\cdot\log_e p(j/i) \tag{58}$$

となる。但し、$p(j/1)=n(1,j)/|\Phi_0|$（∵　4式（48），（49），（51），（54））は推定すべきパラメータではない。p（j/2）のみ推定すべきパラメータである。

カテゴリ番号 $j\in J$ のみの1変数関数 $\theta(j)$ を導入して、パターン分布が2問題地域で同じであるというモデル MODEL(0) は、

$$\text{MODEL}(0):\ p(j/i)=\theta(j),\ i\in I,\ j\in J \tag{59}$$

と表現され、問題地域番号 $i\in I$ とカテゴリ番号 $j\in J$ のみの2変数関数 $\theta(j/i)$ を導入して、パターン分布が2問題地域で同じでないというモデルMODEL(1) は、

$$\text{MODEL}(1):\ p(j/i)=\theta(j/i),\ i\in I,\ j\in J \tag{60}$$

で表現される。

4.5.4　MODEL(0) の AIC

MODEL(0)では、式(59)が仮定されているから、対数尤度 $\ell(\{\theta(j)\})$ は、式(58)から、

$$\ell(\{\theta(j)\})$$
$$=K+\sum_{i\in I}\sum_{j\in J} n(i,j)\cdot\log_e\theta(j)$$
$$=K+\sum_{j\in J}\{\sum_{i\in I} n(i,j)\}\cdot\log_e\theta(j) \tag{61}$$

となる。多項分布の対数尤度（4.2節）から考えて、式（61）から求まる最尤推定量 $\theta\hat{}(j)$ は、

$$\theta\hat{}(j)=\sum_{i\in I} n(i,j)/\sum_{i\in I} n(i),\ j\in J \tag{62}$$

である。

このとき、推定すべき自由パタラメータの数は、

$$\theta\hat{}(1),\ \theta\hat{}(2),\cdots,\theta\hat{}(|J|) \tag{63}$$

の内、式(54)の制約を考慮すると、$|J|-1$個であるから、MODEL(0)の赤池情報量基準 AIC(0)は、式(45)から、

$$AIC(0)$$
$$=-2\times[K+\sum_{j\in J}\{\sum_{i\in I}n(i,j)\}\cdot\log_e\{\sum_{i\in I}n(i,j)/\sum_{i\in I}n(i)\}]+2\times[|J|-1] \quad (64)$$

で与えられる。

E3.5.5 MODEL(1)のAIC

一方、MODEL(1)では、式(60)が仮定されているから、対数尤度$\ell(\{\theta(j/i)\})$は、式(58)から、

$$\ell(\{\theta(j/i)\})$$
$$=K+\sum_{i\in I}\sum_{j\in J}n(i,j)\cdot\log_e\theta(j/i) \quad (65)$$

で与えられ、式(54)の制約を考慮しながら、$i=1$の場合を除き、

$$\partial\ell(\{\theta(j/i)\})/\partial\theta(j/i)=0 \quad (66)$$

とおくことにより、$\theta(j/i)$の最尤推定量$\theta\hat{}(j/i)$として、

$$\theta\hat{}(j/i)=n(i,j)/n(i),\ i=2,\ j\in J \quad (67)$$

を得ることができる。

このMODEL(1)には、

$$\theta\hat{}(j/i),\ i\in I\equiv\{1,2\},\ j\in J \quad (68)$$

という$|I|\times|J|=2\times|J|$個のパラメータが含まれているが、式(54)の制約を考慮し、然も、$i=1$の場合を除くから、推定すべき自由パタラメータの数は、$|J|-1$である。

MODEL(1)の赤池情報量基準 AIC(1)は、式(45)から、

$$AIC(1)$$
$$=-2\times[K+\sum_{i\in I}\sum_{j\in J}n(i,j)\cdot\log_e\{n(i,j)/n(i)\}]+2\times[|J|-1] \quad (69)$$

で与えられる。

4.5.6 Φ_0についての適合度検定

赤池情報量基準 AICが小さいほど、より良いモデルであるから、2式(69),(64)から、その差

$$AIC(1)-AIC(0)$$
$$=-2\times[\sum_{i\in I}\sum_{j\in J}n(i,j)\cdot\log_e\{n(i,j)/n(i)\}$$
$$-\sum_{j\in J}\{\sum_{i\in I}n(i,j)\}\cdot\log_e\{\sum_{i\in I}n(i,j)/\sum_{i\in I}n(i)\}] \quad (70)$$

が1〜2以上ならば、MODEL(0)が採用されて良い。

このように、**MODEL(0)が採用されたならば、認識システム RECOGNITRON の認識機能はΦ_0($\subset\Phi$)について正常に機能していると見做される。**

ここに、モデルの自由パラメータ数$|J|-1$は、データ$n(2)$に対し、$2\sqrt{n(2)}$、或いは、高々、$n(2)/2$迄とした方が良いから、不等式

$$|J|-1\ll 2\sqrt{n(2)}\ \vee\ |J|-1\ll n(2)/2 \quad (71)$$

が成立するように、$n(2)$が選ばれていることが望ましい。

5.むすび

不動点探索形構造受精に関する多段階帰納推論を使ったパターン認識アルゴリズムを内蔵している認識システムRECOGNITRONの3構成要素は、T, SM, BSC であり、この3構成要素 T, SM, BSCを用い、カテゴリ選択関数 CSF, 構造受精作用素 $A(\gamma)$ が構成される [37]。

処理の対象としたパターン集合$\Phi_0(\subseteq\Phi)$に対する認識率が良好でない場合、その原因として、3構成要素 T, SM, BSC が Φ_0に関し、適切に選定されていなかったことが先ず、挙げられる。

本論文では、第2章で説明された不動点（探索形構造受精多段階帰納推理によるパターン）認識の働きを備えている認識システム RECOGNITRON [37] の情報処理機能が、

処理の対象とする問題のパターン(入力パターン) φ について、第2章の式(11)のカテゴリ事前確率分布を多段階的に第2章の式(45)のカテゴリ事後確率分布に変換する

ことに注目し、T, BSC が適切に選ばれている状況の下で、数理的に、

（一）SM が適切に選定されているかどうかの ε-検定（3.2節）
（二）無作為標本であるどうかの、χ^2分布を使っての検定手法（3.3.5節）
（三）類似度関数 SM の、有限部分集合Φ_0についての適切さを赤池情報量基準 AIC で判定する方法

が研究された。

本研究の主目的は、勿論、上記の（三）であったが、実際にどの程度信頼性のある検証手法であるかどうかは、計算機シミュレーションを繰り返しみなければ判明しないのが残念である。

本研究のみでないが、残された諸研究として、次の①, ②が挙げられる：

①**パターン認識の働きに関する"抽象化"**とは、類似したパターンを同一カテゴリ（類概念）として同一視する操作を指すことになるのであるが、この抽象化をSS公理系の下で、数理的に捕らえること。

認識の働きによる抽象化について、少し解説しておこう。

記号処理分野では、例えば、

$$f(\mathrm{male})=f(\mathrm{female})=\mathrm{human} \tag{1}$$

というようなソート写像

$$f: S\rightarrow S' \tag{2}$$

を導入して、ユーザに簡単な質問をして、ユーザの意図する類似関係に照らして、"類似性"の観点から領域知識を修正する手法（**知識獲得手法**）[45] と対比できるものである：

抽象化とは、類似したソート概念を同一視する操作を指すものとし、任意の式 E 中のソート記号をソート写像 f により書き換えた結果を f(E) と表記し、式 E_1, E_2 の書き換え結果 $f(E_1)$, f(E2)が等しいという関係

$$f(E_1)=f(E_2) \tag{3}$$

を満たす2式 E_1, E_2 は互いに類似していると、解釈する。f の下での E′の具体化とは、

$$f^{-1}(E')\equiv\{E\,|\,f(E)=E'\} \tag{4}$$

の各要素を指す。 □

上記の記号処理分野での知識獲得手法と対比するのは、以下に説明されるパターン認識分野での"**パターンモデル構築手法** [23], [24], [26], [37]"である。

パターン$\varphi\in\Phi$から抽出された第$k\in L$番目の特徴量を$u(\varphi,k)\in R$と表すと、特徴抽出写像

$u : \Phi \times L \to R$（実数全体） (5)

が導入され、抽出された特徴量の全体を、

$\underline{u}(\varphi) \equiv \{u(\varphi, k) \in R \mid k \in L\}$ (6)

と表す。このとき、

$\underline{v} = \{v_k \mid k \in L\} \in \underline{R}$ (7)

が与えられたとき、等式

$\underline{u}(\eta) = v$ (8)

を満たすパターン$\eta \in \Phi$を、

写像uの組$\underline{u}$：$\Phi \times L \to R$の下での$\underline{v}$の**具体化**と呼ぶ。

$\underline{u}^{-1}(\underline{v}) \equiv \{\eta \in \Phi \mid \underline{u}(\eta) = \underline{v} \in \underline{R}\}$ (9)

は、類似したパターンの集合である。

パターン$\varphi \in \Phi$の、文献[37]の付録A、axiom 1を満たすパターンモデル$T\varphi$の形式

$T\cdot = \sum_{k \in L} u(\cdot, k) \cdot \psi_k$

ここに、$\{\psi_k\}$ $k \in L$は1次独立な系、或いは、直交系 (10)

が与えられたとき、方程式（8）の解は、

$\eta = \sum_{k \in L} v_k \cdot \psi_k$ (11)

と与えられる[23]，[26]。

②カテゴリ帰属知識$\langle \varphi, \gamma \rangle$の構造受精変換

$\langle \varphi, \gamma \rangle \to TA(\mu)\, T\langle \varphi, \gamma \rangle$ (12)

におけるカテゴリ番号リスト$\mu \in 2^J$を遺伝アルゴリズム[50]で選定して、“不動点探索形構造受精多段階帰納推論を採用したパターン認識手法”を実行する手法 □

文　献

[1] 青木利夫，高橋渉：“集合･位相空間要論”，培風館，Sept.1979

[2] Angus E.Taylor, David C.Lay: “Introduction to function analysis”, John Wiley & Sons, Inc.,1980

[3] Edited by W.K.Estes: “Handbook of learning and cognitive processes（Volume 4 Attention and memory）”, Lawrence Erlbaum Associates, Inc.,1976

[4] Nils J. Nillson: “Problem-solving methods in artificial intelligence”, McGraw-Hill Book Company,Inc.,1971

[5] Elaine Rich and Kevin Knight: “Artificial intelligence（Second edition）”, McGraw-Hill, Inc.,1991

[6] Abhijit S.Pandya and Robert B.Macy : “Pattern recognition with neural networks in C++”, A CRC Book Published in Cooperation with IEEE,1996

[7] M.A.アービッブ：“脳ー思考と行動の源を探るー”，金子隆芳訳，サイエンス社，Apr.1980

[8] ルーメルハート：“人間の情報処理（新しい認知心理学へのいざない）”，御領謙訳，サイエンス社，Sept.1980

[9] 長尾真：“パターン情報処理（電子通信学会大学シリーズI-4）”，コロナ社，Mar.1983

[10] 太原育夫：“認知情報処理”，オーム社，Mar.1991

[11] Jeffrey Wood: “Invariant pattern recognition: A review”, Pattern Recognition, vol.29, no.1, pp.1-17, 1996

[12] 池田克夫，田村秀行，全炳東：“知能情報メディア ーマルチメディアの進化形ー”，電子情報通信学会誌，vol.79, no.8, p.p.788-792, Aug.1996
[13] 斎藤嘉博：“信頼性の基礎数学（信頼性工学講座第2巻)”，東京電機大学出版局，May 1975
[14] Shigeru Yamada: “Software quality/reliability measurement and assessment : Software reliability growth models and data analysis”, Journal of Information Processing, vol.14, no.3, pp.254-266, 1991
[15] 伊藤清：“確率論（現代数学14)”，岩波書店，Nov. 1966
[16] S.Yamada,M.Ohba and S.Osaki: “S-shaped reliability growth modeling for software error detectipon”, IEEE Trans. Reliability R-32, pp.475-478, 1983
[17] 丸山儀四郎：“確率および統計（基礎数学講座10)”，共立出版, Feb.1963
[18] 高井峰生，山城登久二，成田誠之助：“Synchronous Conservative Algorithmを用いた離散事象並列シミュレーションにおける性能予測”，電子情報通信学会論文誌（D-I），vol.J80-D-I, no.3, pp.237-246, Mar.1997
[19] 坂本慶行，石黒真木夫，北川源四郎：“情報量統計学（情報科学講座A･5･4)”，共立出版，June 1993
[20] 澤田清，山道弘明，藤井進：“Kullback-Leiblerの情報量に基づくソフトウェアの信頼性実証試験に関する離散型モデル”，電子情報通信学会論文誌（A），vol.J79-A, no.3, pp.830-833, Mar.1996
[21] 和田安弘，川人光男：“新しい情報量基準とCross Validationによる汎化能力の推定”，電子情報通信学会論文誌（D-Ⅱ），vol. J74-D-Ⅱ, no.7, pp.955-965, July 1991
[22] 鈴木昇一：“誤差確率分布を考慮した誤差逆伝播学習”，情報研究（文教大学・情報学部），vol.13, pp.173-202, Dec.1992
[23] 鈴木昇一：“認識工学（上)”，柏書房，Feb.1975
[24] 鈴木昇一：“マルチメディア情報処理のための一般化誤差逆伝播学習ニューラルネットの新数理”，近代文芸社，Sept.1996
[25] 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.115-1122, Nov.1977
[26] 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
[27] 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
[28] 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
[29] 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
[30] 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
[31] 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
[32] 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

[33] 鈴木昇一，前田英明："パターンの変形理論"，情報研究（文教大学・情報学部），vol.16, pp.209-267, Dec.1995
[34] 鈴木昇一，佐久間拓也，前田英明："数理形態学における諸演算とモデル構成作用素"，情報研究（文教大学・情報学部），vol.17, pp.133-170, Dec.1996
[35] 鈴木昇一："Radial-basis function networks, wavelet-based networksを用いたモデル構成作用素の構成法"，情報研究（文教大学・情報学部），vol.17, pp.71-131, Dec.1996
[36] 鈴木昇一："パターン認識の数学的理論"，電子情報通信学会技術研究報告［パターン認識と学習，パターン認識と理解］，PRL84-6（第1部），PRL84-30, PRL84-38, PRL85-27, PRU86-8, PRU86-35, PRU86-69, PRU87-1, PRU87-28, PRU88-30, PRU89-1, PRU89-27,PRU89-40,PRU89-66, PRU89-77, PRU89-136, PRU90-5, PRU90-15,PRU90-29, PRU90-125,PRU91-1, PRU91-29, PRU91-42, PRU92-1, PRU92-18, PRU92-25, PRU92-89, PRU92-102（第28部），May 1984～Jan.1993
[37] 鈴木昇一："知能情報メディア情報処理のためのカテゴリ帰属知識を用いたパターン認識問題の数理的一般解決"，近代文芸社，June 1997
[38] 河田敬義，丸山文行："基礎課程 数理統計"，裳華房，p.59, p.88, Mar.1963
[39] S.Kullback and R.A.Leibler: "On information and sufficiency", Annals of mathematical statistics,, vol.22, pp.79-86, 1951
[40] Hirotugu Akaike: "A new look at the statistical model identification", IEEE Trans. on Automatic Control, vol.AC-19, no.6, Dec.1974
[41] 竹内啓："AIC基準による統計的モデルの選択をめぐって"，計測と制御，vol.22, no.5, pp.445-453, May 1983
[42] H.Tong: "Determination of the order of a Markov chain by Akaike's information criterion", J.Applied Probability, vol.12, 1975
[43] I.Csiszar: "I-Divergence geometry of probability distributions and minimization problems", The Annals of Probability, vol.3, no.1, pp.146-158, 1975
[44] Gideon Schwarz: "Estimating the dimension of a model", The Annals of Statistics, vol.6, no.2, pp.461-464, 1978
[45] 大久保好章，原口誠："ゴールに依存した抽象化を利用した知識修正法"，コンピュータソフトウェア，vol.14, no.5, pp.60-66, Sep.1997
[46] Paul Dagum, Michael Luby: "An optimal approximation algorithm for Bayesian inference", Artificial Intelligence, vol.93,, pp.1-27, 1997
[47] Luc Devroye, Laàszlô Györhi, Gàbor Lugosi: "A probabilistic theory of pattern recognition", Springer-Verlag New York, Inc., 1996
[48] Gilbert G. Walter: "Wavelets and other orthogonal systems with applications", CRC Press, Inc., 1994
[49] A.Cichocki,R.Unbehauen: "Neural networks for optimization and signal processing", John Wiley & Sons, Inc., Mar.1994
[50] David E.Goldberg: "Genetic algorithms in search,optimization,and machine learning", Addison-Wesley publishing company, Inc., Jan.1989

付録（文献［37］（パターン認識問題の数理的一般解決）の訂正）

［訂正1］p.23↓ℓ.10

（誤　り）程式の求解過程が視覚されやすく追跡されやすい形式なので、認識

→（正しい）程式の求解過程が視覚化されやすく追跡されやすい形式なので、認識

［訂正2］p.31↑ℓ.6

（誤　り）3°抽出された特徴量の組を用いて、その帰属するであろうカテ

→（正しい）3°抽出された特徴量の組を用いて、如何なる手法でその帰属するであろうカテ

［訂正3］p.36↓ℓ.6

（誤　り）（三）（T-不変性）$\forall \varphi \in \Phi, SM(T\varphi, \omega_j) = SM(\varphi, \omega_j)$

→（正しい）（三）（T-不変性）$\forall \varphi \in \Phi, \forall j \in J, SM(T\varphi, \omega_j) = SM(\varphi, \omega_j)$

［訂正4］p.38↑ℓ.5

（誤　り）間的内部表現が**知覚的記憶表象**（perceptual and memorial representative）と呼ばれるものである。**認**

→（正しい）間的内部表現が**知覚的記憶表象**と呼ばれるものである。**認**

［訂正5］p.47↓ℓ.4

（誤　り）　$\gamma = \langle j(1), j(2), \cdots, j(q) \rangle$　(1.3)

→（正しい）　$\gamma = [j(1), j(2), \cdots, j(q)]$　(1.3)

［訂正6］p.48↓ℓ.13

（誤　り）　②実多変連続変数の真理関数の、補間多項式による

→（正しい）　②実連続多変数の真理関数の、補間多項式による

［訂正7］p.52↑ℓ.2

（誤　り）り入れた“特徴抽出過程”の研究はS.Suzukiに先ず、行

→（正しい）り入れた“特徴抽出過程”の研究はS.Suzukiによって先ず、行

［訂正8］p.56↓ℓ.12

（誤　り）$x = \langle x_1, x_2, \cdots, x_q \rangle \in M(=R^q)$　(2.6)

→（正しい）$x = \langle x_1, x_2, \cdots, x_q \rangle \in M(\subseteq R^q)$　(2.6)

［訂正9］p.57↓ℓ.7

（誤　り）$M \equiv \{0,1,2,\cdots\}, dm(x) = 1$ if $x \in M$　(2.11)

→（正しい）$M \equiv \{0,\pm 1,\pm 2,\cdots\}, dm(x) = 1$ if $x \in M$　(2.11)

［訂正10］p.61↓ℓ.14

（誤　り）$\| \varphi \|_K \equiv (\varphi, \varphi)_K$　(2.38)

→（正しい）$\| \varphi \|_K \equiv \sqrt{(\varphi, \varphi)_K}$　(2.38)

［訂正11］p.61↑ℓ.5

（誤　り）2.1.4　代表パターンω_jの集合Φと、実際に処理の対象とするパターン集合Φ

→（正しい）2.1.4　代表パターンω_jの集合Ωと、実際に処理の対象とするパターン集合Φ

［訂正12］p.63↑ℓ.6

（誤　り）加していく形で増加していくものである。

→（正しい）加していく形で増加するものである。

［訂正13］p.64↓ℓ.4

（誤　り）①(initialization) $\varphi \in \Phi_B$はパターンである。

→（正しい）①(initialization) $\varphi\in\Phi_B$はパターンである。特に、$0\in\Phi_B$であらねばならない。

［訂正14］p.72↓ℓ.4

（誤　り）よって、定理2.1によれば、対【Φ,T】、つまり、式（2.56）の

→（正しい）よって、定理2.1によれば、対【$\Phi,T_{\circ}$】、つまり、式（2.56）の

［訂正15］p.73↓ℓ.2

（誤　り）し、よって、定理2.1によれば、対【Φ,T】、つまり、式（2.

→（正しい）し、よって、定理2.1によれば、対【$\Phi,T_{\bullet}$】、つまり、式（2.

［訂正16］p.82↑ℓ.13

（誤　り）（イ）カテゴリ間分散とカテゴリ分散の総和とのフィーシャ比

→（正しい）（イ）カテゴリ間分散とカテゴリ分散の総和とのフィシャ比

［訂正17］p.87↓ℓ.8

（誤　り）部分集合$\{j_1,j_2,\cdots,j_k\}\in 2^J$と、同一視している。よって、

→（正しい）部分集合$\{j_1,j_2,\cdots,j_k\}\in 2^J$と、同一視している。但し、

$$p\neq q\Rightarrow j_p\neq j_q$$

が成り立っているとする。また、

［訂正18］p.91↓ℓ.12

（誤　り）モデル$T\varphi\in\Phi$のカテゴリ帰属知識$\langle T\varphi,\gamma\rangle$と等しい事実を指摘

→（正しい）モデル$T\varphi\in\Phi$のカテゴリ帰属知識$\langle T\varphi,\gamma\rangle$と構造的に等しい事実を指摘

［訂正19］p.91↑ℓ.2

（誤　り）ゴリ帰属知識$\langle\varphi,CSF(\varphi,\gamma)\rangle$と等しい事実を指摘する次の定理4.2から理解できよう。

→（正しい）ゴリ帰属知識$\langle\varphi,CSF(\varphi,\gamma)\rangle$と構造的に等しい事実を指摘する次の定理4.2から理解できよう。

［訂正20］p.97↓ℓ.5

（誤　り）Now$(\langle\varphi,\gamma\rangle,\langle\psi,\lambda\rangle)$ is the projection

→（正しい）Now $(\langle\varphi,\gamma\rangle,\langle\psi,\lambda\rangle)$ is the projection

［訂正21］p.100↓ℓ.8

（誤　り）**Turing**に相当し、ありとあらゆるこれまでの多数の研究者の提案

→（正しい）**Turing machine**に相当し、ありとあらゆるこれまでの多数の研究者の提案

［訂正22］p.102↓ℓ.11

（誤　り）更新作用素（updating operator）、或いは、**構造受精作用素**（structural-fertilization operator）と呼ばれる

$$A(\gamma):\Phi\rightarrow\Phi,\ ここに、\gamma\in 2^J \tag{6.10}$$

は、3付録A, B, Cで用意された

→（正しい）更新作用素（updating operator）、或いは、**構造受精作用素**（structural-fertilization operator）と呼ばれる写像

$$A(\gamma):\Phi\rightarrow\Phi,\ ここに、\gamma\in 2^J \tag{6.10}$$

は、3付録A, B, Cで用意された3構成作用素

［訂正23］p.108↓ℓ.9

（誤　り）(Church,s thesis）と呼ばれる主張

→（正しい）(Church's thesis）と呼ばれる主張

［訂正24］ p.110 ↓ ℓ.14
（誤　り） 密化作用素 $TA(\mu)T$ はは実現していることになる（式（6.48）
→（正しい） 密化作用素 $TA(\mu)T$ は実現していることになる（式（6.48）
［訂正25］ p.113 ↑ ℓ.11
（誤　り） という、“カテゴリ選択関数CSFの全射性” の成立は、定理4.2から保証されている。
→（正しい） という、“カテゴリ選択関数CSFの全射性” の成立は、定理3.2から保証されている。
［訂正26］ p.116 ↓ ℓ.3
（誤　り） 各線形1次係数 c_k, d_ℓ が非負・単位非超過条件を満たす2つの線形1次結合
→（正しい） 各線形1次係数 c_k, d_ℓ が非負・単位非超過条件を満たす2つの1次結合
［訂正27］ p.116 ↓ ℓ.11
（誤　り） $(\langle \varphi_k, \gamma_k \rangle, \langle \psi_\ell, \lambda_\ell \rangle) = 0 \quad (k \neq \ell)$ (7.15)
→（正しい） $(\langle \varphi_k, \gamma_k \rangle, \langle \varphi_\ell, \gamma_\ell \rangle) = 0 \quad (k \neq \ell)$ (7.15)
［訂正28］ p.116 ↓ ℓ.18
（誤　り） $(\langle \varphi_k, \gamma_k \rangle, \langle \psi_k, \lambda_k \rangle) = 1$ for any $k \in K$ (7.16)
→（正しい） $(\langle \varphi_k, \gamma_k \rangle, \langle \varphi_k, \gamma_k \rangle) = 1$ for any $k \in K$ (7.16)
［訂正29］ p.117 ↑ ℓ.6
（誤　り） $\langle \psi, \lambda \rangle = \sum_{j \in K} c_j \cdot \langle \omega_j, [j] \rangle + \langle \psi_\perp, \lambda_\perp \rangle$ (7.19)
とおくと、
→（正しい） $\langle \psi, \lambda \rangle = \sum_{j \in K} c_j \cdot \langle \omega_j, [j] \rangle + \langle \psi_\perp, \lambda_\perp \rangle$ (7.19)
$\forall j \in K, (\langle \psi_\perp, \lambda_\perp \rangle, \langle \omega_j, [j] \rangle) = 0$ (7.20)
とおくと、
［訂正30］ p.117 ↑ ℓ.2
（誤　り） （ロ） $\forall j \in K, (\langle \psi_\perp, \lambda_\perp \rangle, \langle \omega_j, [j] \rangle) = 0$ (7.20)
$\Rightarrow \forall j \in K, c_j = (\langle \psi, \lambda \rangle, \langle \omega_j, [j] \rangle).$
→（正しい） （ロ） 式（7.20）の成立
$\Rightarrow \forall j \in K, c_j = (\langle \psi, \lambda \rangle, \langle \omega_j, [j] \rangle).$
［訂正31］ p.118 ↓ ℓ.2
（誤　り） （ハ） 式（7.20）の成立
$\Leftrightarrow [\forall j \in CSF(\psi_\perp, \lambda_\perp)(\neq \phi), SM(\psi_\perp, \omega_j) = 0]$
→（正しい） （ハ） 式（7.20）の成立
$\Leftrightarrow [\forall j \in CSF(\psi_\perp, \lambda_\perp)(\neq \phi) \subset K, SM(\psi_\perp, \omega_j) = 0]$
［訂正32］ p.118 ↓ ℓ.9
（誤　り） （ヘ） 添字の集合Kを $K = J$ と選んでいれば、
$\| \langle \psi_\perp, \lambda_\perp \rangle \| = 0$
$\Leftrightarrow [\forall j \in CSF(\psi_\perp, \lambda_\perp)(\neq \phi), SM(\psi_\perp, \omega_j) = 0]$
$\vee\ CSF(\psi_\perp, \lambda_\perp) = \phi.$
→（正しい） （ヘ） 添字の集合Kを $K = J$ と選んでいれば、
$[\forall j \in CSF(\psi_\perp, \lambda_\perp)(\neq \phi) \subset J, SM(\psi_\perp, \omega_j) = 0]$
$\vee\ CSF(\psi_\perp, \lambda_\perp) = \phi$
が成立し、これは $\| \psi_\perp, \lambda_\perp \| = 0$ と同値である。

［訂正33］ p.119↓ ℓ.7

（誤　り）（へ）の成立は、2式（5.3），（5.4）から明らかである。

→（正しい）（へ）の成立は、（ハ）と2式（5.3），（5.4）とから明らかである。

［訂正34］ p.121↓ ℓ.13

（誤　り）2性質（a），（b）の成立は、次のようにわかる。

→（正しい）2性質（b），（c）の成立は、次のようにわかる。

［訂正35］ p.121↑ ℓ.10

（誤　り）という（a）の成立が判明し、

→（正しい）という（b）の成立が判明し、

［訂正36］ p.121↑ ℓ.2

（誤　り）という（b）が成り立つことがわかる。

→（正しい）という（c）の成立がわかる。

［訂正37］ p.122↓ ℓ.8

（誤　り）であることに注意しておかねばならない。あくまで、直交分解式の

→（正しい）であることに注意しておかねばならない。あくまで、直交分解式（7.21）の

［訂正38］ p.130↓ ℓ.4

（誤　り）本書で言うカテゴリ帰属知識 $\langle \varphi, \gamma \rangle$ を，処理の対象とする多段

→（正しい）本書で言うカテゴリ帰属知識 $\langle \varphi, \gamma \rangle$ を処理の対象とする多段

［訂正39］ p.140↓ ℓ.14

（誤　り）$\mathrm{E}(\langle \psi_t, \lambda_t \rangle) \geqq \mathrm{E}(\langle \psi_{t+1}, \lambda_{t+1} \rangle)$ (8.16)

→（正しい）$\mathrm{E}(\psi_t, \lambda_t) \geqq \mathrm{E}(\psi_{t+1}, \lambda_{t+1})$ (8.16)

［訂正40］ p.140↓ ℓ.6

（誤　り）ギー $\mathrm{E}(\langle \psi_t, \lambda_t \rangle)$ は式（6.20）でいうカテゴリ帰属知識の変換過

→（正しい）ギー $\mathrm{E}(\psi_t, \lambda_t)$ は式（6.20）でいうカテゴリ帰属知識の変換過

［訂正40］ p.140↑ ℓ.3

（誤　り）復により、不動点方程式（6.21）が成立するための指標であることがわかる。不等式（8.16）の証明についても、G12章で指摘されているように、条件式（G118）の下で、ミックスチュア条件を満たす類似度関数 SM について、解決済みである。この際、類似度関数の再帰的構成法によれば、任意の類似度関数を必ずミックスチュア条件が満たされるように構成し直すことが可能であることに注意しておかねばならない。

→（正しい）復により、不動点方程式（6.21）が成立するための指標であることがわかる。不等式（8.16）の証明についても、G12章で指摘されているように、条件式（G118）の下で、直交条件を満たす類似度関数 SM について、解決済みである。この際、類似度関数の再帰的構成法によれば、任意の類似度関数を必ず直交条件が満たされるように構成し直すことが可能であることに注意しておかねばならない。

［訂正41］ p.144↓ ℓ.2

（誤　り）ここで、**帰納的推論規則**について、説明しておこう。

→（正しい）ここで、**帰納的推論規則**について、説明しておこう（定理6.2を参照）。

［訂正42］ p.145↑ ℓ.12

（誤　り）**理論的手法**（axiomatic approach）を駆使しようとしている。

→（正しい）**理論的手法**を駆使しようとしている。

[訂正43] p.148 ↑ ℓ.2

（誤　り）半順序関係φの代りに、2項関係としての＜を、

→（正しい）半順序関係$\sqsubset$の代りに、2項関係としての＜を、

[訂正44] p.152 ↓ ℓ.9

（誤　り）現実のパターンφから離れ過ぎても密着し過ぎても不適切というという意味で、

→（正しい）現実のパターンφから離れ過ぎても密着し過ぎても不適切という意味で、

[訂正45] p.152 ↑ ℓ.10

（誤　り）

Axiom 1 （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)（零元の不動点性; fixed-point property of zero element）

$0 \in \Phi \wedge T0 = 0.$

(ii)（錐性; cone property; 或いは、**吸収性**）

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

for any positive real number a.

(iii)（**ベキ等性**; idempotency; 埋込性）

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

(iv)（写像 T の非零写像性; non-zero mapping property of T）

$\exists \varphi \in \Phi, T\varphi \neq 0.$ □

→（正しい）

Axiom 1 （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）

$\exists \varphi \in \Phi, T\varphi \neq 0.$ □

[訂正46] p.153 ↓ ℓ.12

（誤　り）

[命題A2]**（モデル構成作用素 T の積構成）**

各写像$T_k (1 \leq k \leq n)$が axiom 1 を満たすならば、2条件

(a)（可換性）　$T_j T_k = T_j T_k \ (j > k)$ (A5)

(b)（前段値域・後段定義域の包含性）

$\mathrm{Domain}(T_{k-1}) \supseteq \mathrm{Range}(T_k) \ (2 \leq k \leq n)$ (A6)

の下で、写像

$T \equiv T_1 \cdot T_2 \cdots T_n$ (A7)

は axiom 1 を満たす。ここに、

Domain (T_k) ≡ $\{\varphi \mid \varphi \in \Phi \wedge T\varphi \in \Phi \wedge T\varphi \neq 0\}$ (A8)

Range (T_k) ≡ $\{\eta \mid \eta = T\varphi \in \Phi \wedge \varphi \in \Phi\}$. (A9)

→（正しい）

［命題A2］**（モデル構成作用素 T の積構成）**

各写像 $T_k(1 \leqq k \leqq n)$ が axiom 1 を満たすならば、3条件

(a) $\exists \varphi_{k+1} \in \mathrm{Domain}(T_{k+1}), \eta_{k+1} \equiv T_{k+1}\varphi_{k+1} \neq 0$ であるような任意の η_{k+1} について、

$T_k\,\eta_{k+1} \neq 0\,(1 \leqq k \leqq n-1)$ (A5)

(b)（前段値域・後段定義域の同等性）

$\mathrm{Domain}(T_{k-1}) = \mathrm{Range}(T_k)\,(2 \leqq k \leqq n)$ (A6)

(c) **（可換性）** $T_jT_k = T_kT_j\,(j > k)$ (A7)

の下で、写像

$T \equiv T_1 \cdot T_2 \cdots T_n$ (A8)

は axiom 1 を満たす。ここに、

$\mathrm{Domain}(T_k) \equiv \{\varphi \mid \varphi \in \Phi \wedge T_k\varphi \in \Phi\}$ (A9)

$\mathrm{Range}(T_k) \equiv \{\eta \mid \exists \varphi \in \mathrm{Domain}(T_k), \eta \equiv T_k\varphi \in \Phi\}$. (A10)

［訂正47］ p.155 ↓ ℓ.7

（誤　り）(ii)（規格化条件; probability condition）

→（正しい）(ii)（規格化条件, 正規条件; probability condition,normality）

［訂正48］ p.156 ↑ ℓ.2

（誤　り）［命題B6］**（類似度関数 SM の指摘関数再帰的構成）**

→（正しい）［命題B6］**（類似度関数 SM の指数関数再帰的構成）**

［訂正49］ p.157 ↑ ℓ.6

（誤　り）

(ii) $SM(\varphi, \omega_j) - \max_{k \in J-\{j\}} SM(\varphi, \omega_j)$

$\geqq 1 - \delta > 0.$ □

→（正しい）

(ii) $SM(\varphi, \omega_j) - \max_{k \in J-\{j\}} SM(\varphi, \omega_j)$

$\geqq 1 - 2\delta > 0.$ □

［訂正50］ p.163 ↓ ℓ.4

（誤　り）その定義域が Φ であり、その値域が、パターン $\varphi \in \Phi$ のカテゴリ帰

→（正しい）その定義域が $\Phi \times 2^J$ であり、その値域が、パターン $\varphi \in \Phi$ のカテゴリ帰

［訂正51］ p.163 ↓ ℓ.4

（誤　り）ようなカテゴリ選択関数（category-selection func-tion）と呼ばれる

→（正しい）ようなカテゴリ選択関数（category-selection function）と呼ばれる

［訂正52］ p.169 ↑ ℓ.8

（誤　り）パターン認識とは、事例からの一般化（generaliza-tion from

→（正しい）パターン認識とは、事例からの一般化（generalization from

［訂正53］ p.187 ↓ ℓ.5

（誤　り）ミックスチュア条件を満たす類似度関数（2.7.2項）SM であれば、

→（正しい）直交条件を満たす類似度関数（2.7.2項）SM であれば、

[訂正54] p.189 ↑ ℓ.13

(誤　り)

[定理G9の系1] **(代表パターンω_jのモデル$T\omega_j$に関するカテゴリ帰属知識$\langle T\omega_j, [j]\rangle$の不動点定理)**

(iii) $j \in \mu \cap \gamma$ ならば、

$TA(\mu)\, T\langle T\omega_j, \gamma\rangle = \langle T\omega_j, [j]\rangle$

(iv) $j \in \mu$ ならば、

$TA(\mu)\, T\langle T\omega_j, [j]\rangle = \langle T\omega_j, [j]\rangle.$

(証明) 定義式 (6.18) の特殊形

$$\langle \psi, \lambda\rangle =_{\triangle} TA(\mu)\, T\langle \omega_j, \gamma\rangle \in \langle \Phi, 2^J\rangle \tag{G90}$$

でのψ, λは、

$$\psi = TA(\mu \cap \gamma) T\omega_j \tag{G91}$$

$$\lambda = CSF(\omega_j, \mu \cap \gamma) \tag{G92}$$

である。ψ, λを計算しよう。axiom 2の (i) より、

$\Sigma_{k \in \mu \cap \gamma} SM(\omega_j, \omega_k) =$

→ (正しい)

[定理G9の系1] **(代表パターンω_jのモデル$T\omega_j$に関するカテゴリ帰属知識$\langle T\omega_j, [j]\rangle$の不動点定理)**

(iii) $j \in \mu \cap \gamma$ ならば、

$TA(\mu)\, T\langle T\omega_j, \gamma\rangle =_{\triangle} \langle T\omega_j, [j]\rangle$

(iv) $j \in \mu$ ならば、

$TA(\mu)\, T\langle T\omega_j, [j]\rangle =_{\triangle} \langle T\omega_j, [j]\rangle.$

(証明) 定義式 (6.18) の特殊形

$$\langle \psi, \lambda\rangle =_{\triangle} TA(\mu)\, T\langle \omega_j, \gamma\rangle \in \langle \Phi, 2^J\rangle \tag{G90}$$

でのψ, λは、

$$\psi = TA(\mu \cap \gamma) T\omega_j \tag{G91}$$

$$\lambda = CSF(\omega_j, \mu \cap \gamma) \tag{G92}$$

である。ψ, λを計算しよう。axiom 2の (i) より、

$\Sigma_{k \in \mu \cap \gamma} SM(\omega_j, \omega_k) =$

[訂正55] p.190 ↑ ℓ.2

(誤　り)

系1の (iii) は、定理4.1を考慮すると、(i) から明らかである。系1の (iv) は、系1の (iii) から明らかである。 □

→ (正しい)

系1の (iii) は、

「式 (G90) の $\langle \psi, \lambda\rangle$ について、

$\psi = TA(\mu \cap \gamma) TT\omega_j$

$= TA(\mu \cap \gamma) TT\omega_j$　∵ axiom 2, (iii) の後半

$\lambda = CSF(T\omega_j, \mu \cap \gamma)$

$= CSF(\omega_j, \mu \cap \gamma)$　∵ 命題3.1

となって、2式（G91），(G92）に一致する」
を考慮すると、(i）の証明から明らかである。系1の（iv）は、系1の（iii）から明らかである。
□

[訂正56] p.191 ↑ ℓ.12
（誤　り）
⇒式（G98）において、$\gamma = \gamma'$としたもの.
（証明）　系1は明らかである。式（6.18）から、
→（正しい）
⇒式（G98）において、$\lambda = \gamma = \gamma'$ としたもの.
（証明）　式（G97）において、φの代りに$T\varphi$を考えると、
$TA(\mu)\,T\langle T\varphi, \gamma\rangle =_{\triangle} \langle T\varphi, \gamma'\rangle$
と書き直され、$\lambda = \gamma = \gamma'$と考えれば、系1は定理G 10から明らかである。式（6.18）から、

[訂正57] p.195 ↑ ℓ.6
（誤　り）次の定理G 11は、2式（6.42），(6.43）の$E(\varphi, \gamma)$が最小値0がと
→（正しい）次の定理G 11は、2式（6.42），(6.43）の$E(\varphi, \gamma)$が最小値0をと

[訂正58] p.203 ↑ ℓ.6
（誤　り）
［定理G16］（**代表パターンの認識困難度定理**）
代表パターンω_jの認識困難度 $grd(\omega_j)$ は、
→（正しい）
［定理G16］（**代表パターンの認識困難度定理**）
$\langle \omega_j, \lambda \rangle$ において、$j \in \lambda$である場合
代表パターンω_jの認識困難度 $grd(\omega_j)$ は、

[訂正59] p.203 ↑ ℓ.17
（誤　り）
［定理G17］（**類似度1，大分類関数値1の認識困難度定理**）
$$j \in \lambda \wedge SM(\varphi, \omega_j) = 1 \wedge BSC(\varphi, j) = 1 \tag{G163}$$
→（正しい）
［定理G17］（**類似度1，大分類関数値1の認識困難度定理**）
$\langle \varphi, \lambda \rangle$ において
$$j \in \lambda \wedge SM(\varphi, \omega_j) = 1 \wedge BSC(\varphi, j) = 1 \tag{G163}$$

[訂正60] p.207 ↓ ℓ.7
（誤　り）
（一）全体として形態的に秩序あるまとまりをなそうとする傾向、いわゆるプレグナンツ（Prägnanz-tendenz）の傾向
→（正しい）
（一）全体として形態的に秩序あるまとまりをなそうとする傾向、いわゆるプレグナンツの傾向（Prägnanz-tendenz）

[訂正61] p.227 ↑ ℓ.13
（誤　り）

先ず、定理G2に注意して、P{⟨φ,γ⟩, ⟨η,μ⟩}を、P{⟨φ,γ⟩, ⟨η,μ⟩}

→（正しい）

先ず、定理G2に注意して、P{⟨φ,γ⟩, ⟨η,μ⟩}を、P{⟨φ,γ⟩, ⟨η,μ⟩}.

［訂正62］ p.230↑ℓ.3

（誤　り）時刻 t で状態 X が prob {X (t) $=_{\triangle}$⟨ψ,λ⟩} である確率を

→（正しい）時刻 t で状態 X が⟨η, μ⟩である確率を、

［訂正63］ p.231↓ℓ.9

（誤　り）

次の補助定理 J2は、いわゆる**マルコフ連鎖式（J10）**（Markov chain）の**詳細釣合い条件**（the condition of detailed balance）が成り立つことを指摘している。

→（正しい）

次の補助定理 J2は、いわゆる**マルコフ連鎖**（Markov chain）式（J10）の**詳細釣合い条件**（the condition of detailed balance）が成り立つことを指摘している。

［訂正64］ p.258↓ℓ.15

（誤　り）一般的枠組としての多重解像度　解析（multiresolution analysis）

→（正しい）一般的枠組としての多重解像度解析（multiresolution analysis）

［訂正65］ p.259↑ℓ.14

（誤　り）　where $\mu_{A \cup B}(x) \equiv \max\{\mu A(x), \mu B(x)\}$

(b)　Aのα-レベル集合（α-level set）（非ファジィ集合）

$A(\alpha) \equiv \{x \in X \mid \mu A(x) \geqq \alpha\}$　$(0 \leqq \alpha \leqq 1)$

→（正しい）　where $\mu_{A \cup B}(x) \equiv \max\{\mu_A(x), \mu_B(x)\}$

(b)　Aのα-レベル集合（α-level set）（非ファジィ集合）

$A(\alpha) \equiv \{x \in X \mid \mu_A(x) \geqq \alpha\}$　$(0 \leqq \alpha \leqq 1)$

［訂正66］ p.263↑ℓ.7

（誤　り）

［A28］ Edited by Bruce G.Batchelor : "Pattern recog-nition (Ideas in practice)", Plenum Press, New York, 1978

→（正しい）

［A28］ Edited by Bruce G.Batchelor : "Pattern recognition (Ideas in practice)", Plenum Press, New York, 1978

［訂正67］ p.264↑ℓ.9

（誤　り）

［B1］ Claude E.Shannon: "A symbolic analysis of relay and switching circuits",, vol.57,pp.713-723, 1938

→（正しい）

［B1］ Claude E.Shannon: "A symbolic analysis of relay and switching circuits", Transactions of the American Institute of Electrical Engineers, vol.57, pp.713-723, 1938

［訂正68］ p.265↓ℓ.14

（誤　り）

［B10］ Martin E.Hellman, and Josef raviv: "Probability of error, equivocation, and the Chernoff bound", IEEE Trans. on Information Theory, vol.IT-16, no.4, July 1970

→（正しい）

[B10] Martin E.Hellman,and Josef Raviv: "Probability of error, equivocation, and the Chernoff bound", IEEE Trans. on Information Theory, vol.IT-16, no.4, pp.368-372, July 1970

[訂正69] p.267 ↓ ℓ.16

（誤　り）[B30] Jeffrey Wood： "Invariant pattern recogni-tion:A review",

→（正しい）[B30] Jeffrey Wood： "Invariant pattern recognition:A review",

[訂正70] p.275 ↓ ℓ.17

（誤　り）[H5] 水野鉄司: "多変量データ解析講義（統計ライブラリー)", 朝倉書店, NOv.1996

→（正しい）[H5] 水野欽司: "多変量データ解析講義（統計ライブラリー)", 朝倉書店, NOv.1996

[訂正71] p.275 ↑ ℓ.9

（誤　り）Cybernetics, vol.18, no.1, Jan./Feb.1988

→（正しい）Cybernetics, vol.18, no.1, Jan.Feb.1988

[訂正72] p.118 ↑ ℓ.13

（誤　り）3写像SM, BSC, γ

→（正しい）3写像SM, BSC, CSF

[訂正73] p.195 ↑ ℓ.2

（誤　り）$|\gamma|=1$　(G127)

→（正しい）$\varphi\neq 0 \wedge |\gamma|=1$　(G127)

[訂正74] p.55 ↑ ℓ.1

（誤　り）(φ, η), $\|\varphi\| \equiv (\varphi,\varphi)$　(2.1)

→（正しい）(φ, η), $\|\varphi\| \equiv \sqrt{(\varphi,\varphi)}$　(2.1)

[訂正75] p.74 ↓ ℓ.8

（誤　り）1つは非零であるような複素定数a_jの組

→（正しい）複素定数a_jの組

（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.18" 投稿論文, 論文題目類似度関数の選定に関する適切さの検証法, 投稿年月日　1997年10月9日）