

科学研究費助成事業 研究成果報告書

平成 26 年 6 月 6 日現在

機関番号：17104

研究種目：基盤研究(C)

研究期間：2011～2013

課題番号：23500280

研究課題名(和文) 経験から高次知識を自己組織化する学習理論

研究課題名(英文) Learning theory for higher-knowledge self-organization from experiences

研究代表者

古川 徹生 (FURUKAWA, Tetsuo)

九州工業大学・生命体工学研究科(研究院)・教授

研究者番号：50219101

交付決定額(研究期間全体)：(直接経費) 3,200,000円、(間接経費) 960,000円

研究成果の概要(和文)：高次の知識を得るための階層的学習法に関する研究を行った。第1次階層での学習は、個々のケースにおけるデータを学び、モデル化を行う。第2次学習では1次学習の結果をデータとみなして、すべてのケースに共通する普遍モデル、すなわち高次知識を得る。この学習タスクはテンソル方程式で書くことを見出し、自己組織化写像(SOM)の階層構造で解けることを見出した。さらに多視点的数据解析を可能にするテンソルSOMを開発し、その有効性を示した。

研究成果の概要(英文)：The purpose of this work is to establish the learning theory and the algorithm acquiring the general knowledge through experiences. The proposed method has a hierarchical structure consisting of a set of first learners and the second learner, in which the first learners estimate models of individual experiences and the second learner learns the learning result of the first learners. We found that the task is described as tensor equations, which can be solved by the higher-rank of self-organizing map (SOM). Furthermore, we developed tensor SOM based on the higher-rank of SOM.

研究分野：総合領域

科研費の分科・細目：情報学、感性情報学・ソフトコンピューティング

キーワード：ニューラルネット 自己組織化写像 高次知識獲得 学習理論 多様体 テンソル 自己言及

## 1. 研究開始当初の背景

近年、知的ロボットによる自律的なダイナミクス学習や知識獲得が研究テーマとして注目されるようになってきた。しかし古典的なニューラルネット的アプローチでは、多様なダイナミクスやモデルの集合を安定に学習することは原理的に困難がある。その理由として (a) 異なるダイナミクス間やモデル間で記憶の混乱や干渉が生じて学習プロセスが不安定化する、(b) 膨大なバリエーションを学習する必要がある一方で、同種の経験は回数が限られる、等が挙げられる。さらに階層性を必要とする高次の知識獲得については (c) 低次の学習により獲得したモデルやダイナミクスを高次の学習対象とする必要があるが、それらはニューラルネットの結合係数として埋め込まれておりそのままでは取りだせないという問題点も存在する。

一方、本申請者はモジュラーネット型自己組織化マップ (mnSOM) やメタ多様体 (=ファイバー束) モデルを自己組織化する高階化自己組織化写像 (SOM<sup>n</sup>) の研究に取り組んでおり、特に SOM<sup>n</sup> の研究を通じて上記(a)~(c)の問題点を解決する学習理論の確立に従事してきた。このような背景の下で高次知識を自己組織化するための学習理論とアルゴリズム開発を行うことを本研究の目的とした。

## 2. 研究の目的

私たち人間は経験を積み重ねることを通じて知識を獲得する。それは単なるデータの丸暗記ではなく、ひとつひとつ異なるデータの中から共通性を見出し、普遍的な規則として知識を獲得する。本研究の目的は、データから普遍性を抽出して高次知識を自己組織化するための、中核的な学習理論を確立することである。すなわち本研究の中心テーマは (1) 経験集合から普遍的なモデルを自己組織化すること、(2) 階層的あるいは多視点的な知識表現を実現すること、(3) それらの原理を数学的・アルゴリズム的に記述すること、の3点である。これを実現するため、ニューラルネットワーク的アプローチ・ベイズ的アプローチ・情報幾何的アプローチを組み合わせてアルゴリズム開発と理論確立を行う。また発展的テーマとして、階層的知識表現を用いた自己言及性も試みる。

## 3. 研究の方法

(1) 本研究では、現象の普遍的法則を記述する上で必要な真の説明変数が存在すると仮定し、

観測されたデータから説明変数を推定するという立場を取る。この説明変数は直接観測することができない潜在変数である。また潜在変数と観測データの間には位相同型な非線形写像が存在すると仮定し、潜在変数と位相同型写像の同時推定問題に帰着する。本研究ではこの学習の枠組みを位相保存写像学習と呼んで、その学習理論を構築することを第一の目的とした。位相保存写像学習の代表は Kohonen の自己組織化写像 (Self-Organizing Map: SOM) である。また SOM をベイズ的枠組みで学習理論を構築したのが Bishop の生成位相モデル (Generative Topographic Mapping: GTM) である。本研究では GTM 同様の方法論を取りつつも、高次知識推定に必要な学習アルゴリズムが満たすべき要件を明らかにすることに取り組む。すなわち、単に位相保存写像学習ができるだけでなく、より普遍的な学習理論として記述できることを重要視する。そのような理想的な学習理論に基づくアルゴリズムを本研究では正準位相保存写像学習 (Canonical Topographic Mapping: CTM) と呼び、CTM の原理とアルゴリズムの発見を目指した。

(2) 本研究の取り組みの第二点は、階層的学習による高次の知識獲得を行う学習理論とアルゴリズムの研究である。これは高階 SOM (SOM<sup>n</sup>) を元にしており、下位の学習装置 (1<sup>st</sup> Learners) が個々の現象に関するモデルを作るとともに、上位の学習装置 (2<sup>nd</sup> Learner) が各現象に共通する普遍的なモデルを学習するものである。これはヘテロドメインなメタ学習課題になっている。本研究では SOM<sup>n</sup> 同様に SOM を基本構成要素としたアルゴリズムで行うが、最終的には CTM を構成要素として置き換えることが可能である。

本研究では、特に2つの課題に関して階層的学習の研究を行った。第一の課題は潜在システム空間学習である。これは複数の異なるダイナミカルシステムによって生成された時系列集合を学習することを通して、その背後にあるシステム集合に共通した普遍的法則性を発見するタスクである。

第二の課題は形状空間法の理論とアルゴリズム開発である。これは複数の異なる物体形状を多様体でモデル化することを通して、それらの物体に共通する形状特徴を発見する課題であり、形状を表現する下位学習と形状集合のモデル化を行う上位学習の階層性を持つ。

(3) 本研究の取り組みの第三点は、関係データ

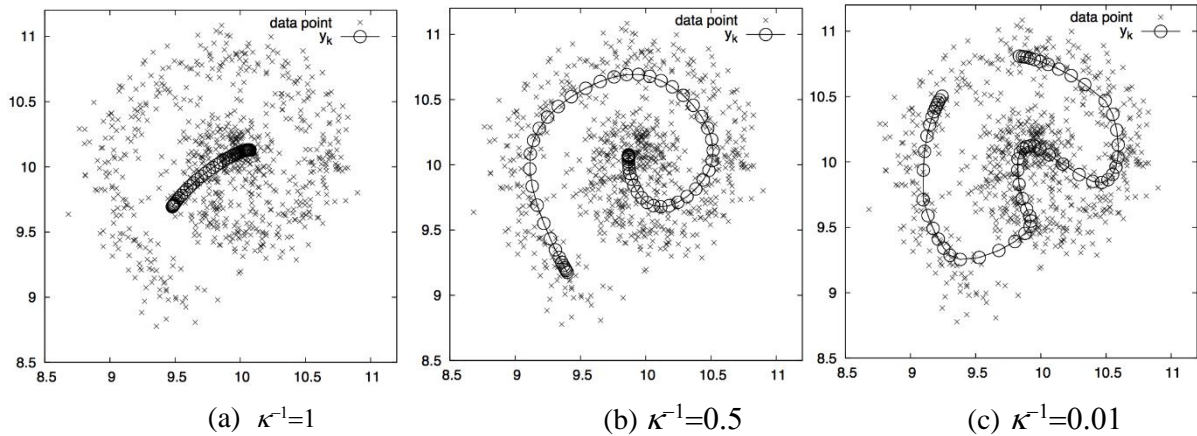


図 1: 一般化した位相保存写像学習における競合因子  $\kappa^{-1}$  の効果. (1) 競合小: GTM に相当 (2) 競合中: 提案理論に基づき導出したアルゴリズム (3) 競合大: SOM に相当. 最適な学習アルゴリズムを理論的に導出できる.

を多視点的に解析する位相保存写像学習のアルゴリズムを開発することである. ここで言う関係データとは, 複数の因子の組み合わせで観測データが得られるような場合を指す. 典型的な例がマーケティングにおける商品-顧客データであり, どの顧客がどの商品をどう評価したかという表としてデータが与えられる. このとき商品と顧客の双方について同時解析するアルゴリズムである. これは SOM<sup>9</sup> による階層的学習の延長で実現できる.

(4) 本研究の第四点は, 階層型学習の延長として可能と思われる新たな学習スタイルを模索・挑戦することである. そのひとつとして, 観測空間がヘテロであるような関係データの取り扱いである. これは商品-顧客データの例で言えば, 商品の種類によって評価項目が異なるような場合に該当する. さらにその延長として, コミュニティにおけるメンバー間の相互評価を通じた自己認識・自己言及の可能性についても研究する.

#### 4. 研究の成果

(1) 位相保存写像学習においては, アルゴリズムの何が位相を保存させるかが最も重要な点である. SOM の場合は近傍学習がその役割を担い, GTM の場合は RBF 基底がその役割を担う. 結果的にはどちらも学習が行われるものの, 位相保存性の理論的根拠を求め, かつ学習性能も SOM, GTM を上回るアルゴリズムの開発に取り組んだ. 本研究では潜在空間におけるノイズが重要な役割を果たすとみなし, 確率的生成モデルを作り, 変分ベイズ法でアルゴリズムを導出した. その結果, 位相保存写像学習の基本となる学習理論を構築でき, GTM や SOM を一元的に理解できるようになった. また導出したアルゴリズム

は GTM を上回る学習安定性が得られ, かつ SOM の問題点であった近傍学習の自動化や最適化も可能になった.

さらにわれわれは, 位相保存写像学習には非正則問題が付随することを見出した. そのため EM 法や変分ベイズ法をそのまま適用しただけでは, しばしば適切な解が得られず, 局所最適解に陥ることがわかった. これは変分近似によって生じる問題ではないため, モンテカルロ法等を用いても解決しない. その原因は等価な置換解同士が拮抗することによる擬似最適解が生じることがわかり, 置換解の間に競合を加えることで解決できることを示した. 置換解間の競合の大きさを  $\kappa$  とすれば, 通常の変分ベイズ法や EM 法は競合のもっとも小さい状況 ( $\kappa^{-1}=1$ ) であるのに対し, SOM は競合が最大 ( $\kappa^{-1}=0$ ) に相当し, 過剰な競合によって学習が困難になる場合があることがわかった. 理論的に最適な競合は  $\kappa^{-1}=0.5$  の場合であり, このとき学習は非常に安定することが示された(図1).

本研究は潜在空間推定の基本的な理論基盤を構築する上で重要な位置を占めると考えている. そのためには今後, 位相保存写像の理論体系構築を行い, 国内外の他の研究を体系内にまとめることが必要である. 特に確率過程としての多様体学習や無限混合分布との関係などの位置づけや, 非正則問題への対応などが重要になる.

(2) 階層的学習による潜在システム空間推定法の研究では, 複数のシステムを学習する課題が以下のテンソル方程式を解く問題であることを示した.

$$\begin{pmatrix} z_{t+1} \\ x_t \end{pmatrix} \simeq \begin{pmatrix} \mathbf{W} & \mathbf{O}_1 \\ \mathbf{O}_2 & \mathbf{V} \end{pmatrix} \times_1 \begin{pmatrix} \psi_\Omega(\omega) \\ \psi_\Theta(\theta) \end{pmatrix} \times_2 \varphi(z_t)$$

ここで  $x_t$  が観測時系列,  $z_t$  がシステム内部の状態変数,  $\omega, \theta$  はそれぞれシステムと観測系を決めるパラメータベクトルである. また  $\mathbf{W}, \mathbf{V}$  はパラメータからシステムと観測系を生成するテンソルである. この中で  $x_t$  のみが既知で, 他の  $z_t, \omega, \theta, \mathbf{W}, \mathbf{V}$  は未知変数である. この5重推定問題を解くテンソル方程式が高次知識を抽出するための方程式であることを示した. これはヘテロドメインなメタ学習課題に相当する困難な課題であるが, 高階 SOM による解法を提案した.

さらにわれわれはこのテンソル方程式が, 形状集合を学習する形状空間法に応用できることを見出した. これは同じ形状クラスに属する形状集合から, そのクラスに共通する形状性質を学習するとともに, クラス内の形状バリエーションを表現するマップを得る方法である(図2). われわれはこの形状空間法を火星探査航空機の翼形状設定に応用することも試みた.

(3) 高階 SOM を元に, 関係データを多視点解析・可視化する Tensor SOM を開発した. これは SOM による非線形テンソル解析法であり, きわめて用途の広い手法である. 図3は Tensor SOM をマーケティングデータに応用した例である. この場合はユーザー, 商品, 状況の3つのマップとして可視化されているが, これらのマップ間の関係も可視化することが可能である. またこの解析結果を利用して商品の推薦システムを構築することも示した.

またマーケティングデータ以外では社内 E-mail トラフィック解析による社員関係の可視化や多チャンネル脳波信号解析にも Tensor SOM を

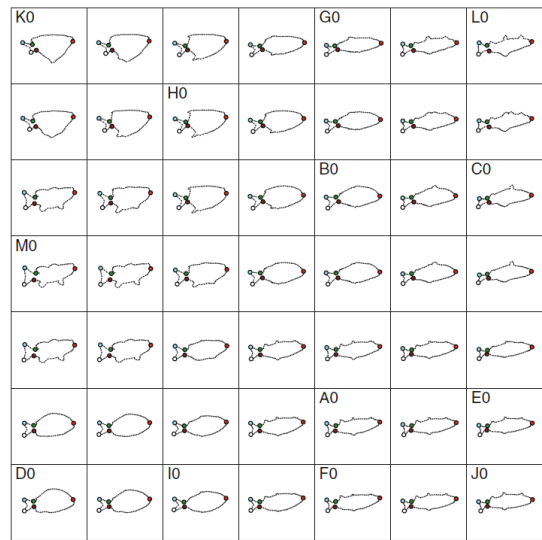


図2: 形状空間法による魚の形状マップ

応用し, その有用性を示した.

(4) 高階 SOM を元に, ヘテロな関係データの可視化アルゴリズムを開発した. ヘテロ関係データの例としては, 同一の調査対象に対して異なる統計調査を行った場合に, それぞれの統計調査同士の関係性を解析・可視化するものである.

ヘテロ関係データの解析アルゴリズムは, 統計調査集合を解析対象とする手法であるが, コミュニティにおいてメンバー同士が相互評価する場合も一種の統計調査集合とみなせば, 同様に本手法の解析対象となる. さらに, このような統計調査集合を解析することも統計調査の一つとすれば, 自己の分析結果もまた分析対象になるといえる(図4). このような自己言及性が現れる点が本手法の興味深い点であり, 「私について考える

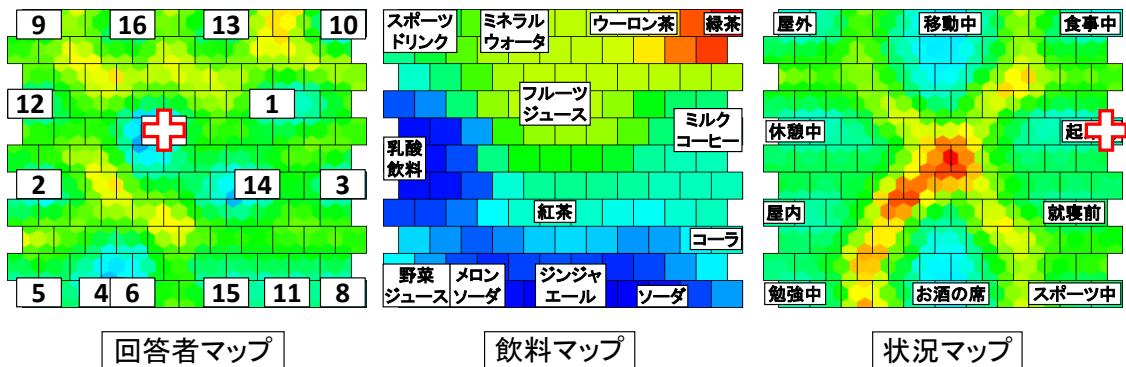


図3: Tensor SOM によるマーケティングデータの解析例. ペットボトル飲料に対するユーザーの嗜好を状況別に分析したもので, ユーザー, 飲料, 状況の3つのマップが得られる.



図 4: ヘテロ関係データ解析法を用いたコミュニティにおけるメンバー間の相互評価の可視化と自己言及. ここではプロ野球の対戦データをメンバー間のインタラクションと仮定して解析した.

私」という知能にも関わると考えている.

ヘテロ関係データ解析の研究はまだ模索段階であり, 知能との関係も含めて今後の発展課題である.

## 5. 主な発表論文等

[雑誌論文](計 5 件)

- ① 大谷 誠, 古川 徹生, 高階 SOM による形状表現マップの自己組織化 ～トポロジー拘束のない形状空間法～, 知能と情報, Vol.25, pp.701–720, 2013. (査読有)  
DOI: 10.3156/jsoft.25.701
- ② 川畑 宣之, 徳永 憲洋, 古川 徹生, 自己成長型モジュラーネットワークを用いた自律移動ロボットにおけるハイブリッド地図のオンライン構築, 知能と情報, Vol.25, pp.659–675, 2013. (査読有)  
[https://www.jstage.jst.go.jp/article/jsoft/25/2/25\\_659/\\_pdf](https://www.jstage.jst.go.jp/article/jsoft/25/2/25_659/_pdf)
- ③ S. Yakushiji, T. Furukawa, Shape space estimation by higher-rank of SOM, Neural Computing and Applications, Vol.22, pp.1267–1277, 2012. (査読有)  
DOI: 10.1007/s00521-012-1004-4
- ④ 薬師寺 翔, 古川 徹生, 高階化 SOM による形状表現マップ, 知能と情報, Vol.24, No.2, pp.648–659, 2012. (査読有)  
<https://www.jstage.jst.go.jp/article/jsoft/24/2/>

24\_2\_648/\_pdf

- ⑤ S. Yakushiji, T. Furukawa, Shape space estimation by SOM<sup>2</sup>, Lecture Notes in Computer Science, Vol.7063, pp.618–627, 2011. (査読有)  
DOI: 10.1007/978-3-642-24958-7\_72

[学会発表](計 17 件)

- ① S. Matsushita, T. Furukawa, Bayesian optimization makes GTM resemble to SOM, 21<sup>st</sup> National Convention of Japanese Neural Network Society, pp.182–183, 2011 年 12 月 17 日, 沖縄県.
- ② M. Nakano, T. Ohkubo, T. Furukawa, SOM Canonica: Establishing a standard algorithm of self-organizing maps, 21<sup>st</sup> National Convention of Japanese Neural Network Society, pp.178–179, 2011 年 12 月 17 日, 沖縄県.
- ③ T. Ohkubo, T. Furukawa, What is required for a multi-dynamical system learning task?, 21<sup>st</sup> National Convention of Japanese Neural Network Society, pp.114–115, 2011 年 12 月 16 日, 沖縄県.

[産業財産権]

○出願状況(計 1 件)

- ① 名称: マップ生成装置, その方法およびプログラム  
発明者: 古川 徹生  
権利者: 同上  
種類: 特許  
番号: 特許願 2011-186140  
出願年月日: 23 年 8 月 29 日  
国内外の別: 国内

## 6. 研究組織

- (1) 研究代表者  
古川 徹生 (FURUKAWA Tetsuo)  
九州工業大学・大学院生命体工学研究科・教授  
研究者番号 50219101