

階層構造を持つ連想記憶モデルの構造発見へのデータ駆動型アプローチ

Data Driven Approach for Structure Discovery of Associative Memory Model with a Hierarchical Structure

若杉 健介[†]
Kensuke Wakasugi桑谷 立[‡]
Tatsu Kuwatani永田 賢二[†]
Kenji Nagata麻生 英樹[§]
Hideki Asoh岡田 真人[†]
Masato Okada

1. まえがき

連想記憶モデル[1]では相関を持つ記憶パターンを埋め込んだ場合、記憶パターンとその混合状態が定常状態となる。また、混合状態が記憶パターンの上位概念を表し、これらが階層構造をなすことが知られている[2]。しかしながら、その構造を客観的に示すことができる手法は確立しておらず、データ駆動によって構造の種類を決定することが求められる。一方、機械学習により構造の種類をデータのみから学習するモデルが提案されており[3]、同モデルはデータの構造決定に利用できる。そこで本研究では、上記の提案手法を連想記憶モデルの解析に適用し、データ駆動型アプローチによって連想記憶モデルが階層構造を持つことを示す。

2. 構造解析手法

2.1. 評価関数

評価関数はKempら[3]の提案手法を連想記憶モデルに合わせてパラメータ調整し、解析に用いた。評価関数はデータ \mathbf{D} が与えられたもとの構造の種類 F と構造 \mathbf{S} の事後確率 $P(\mathbf{S}, F | \mathbf{D})$ で表される。ここで \mathbf{D} は、任意の状態の m 次元特徴量ベクトル $\mathbf{d}^i (1 \leq i \leq n)$ を状態の数 n だけ並べた $n \times m$ の行列である。本研究では後述する連想記憶モデルのニューロンの状態を \mathbf{d}^i とした。また、各状態のことを entity と呼び、一つ以上の entity の集まりを node と呼ぶ。 \mathbf{S} はデータの構造を表す隣接行列である。すべての entity はいずれかの node に含まれるものとし、データの構造は node の接続によって表される。 F は Fig. 1 に示す五種類のいずれかの構造を表す。構造の種類ごとの評価値を比較することで構造の種類決定が可能となる。

$P(\mathbf{S}, F | \mathbf{D})$ はベイズの定理により以下のように展開できる。

$$P(\mathbf{S}, F | \mathbf{D}) \propto P(\mathbf{D} | \mathbf{S}) P(\mathbf{S} | F) P(F) \quad (1)$$

右辺は以下の通りである。ただし簡単のため、特徴量一つの場合 ($\mathbf{D} = \mathbf{d}^1$) の場合について示す。

$$P(\mathbf{d}^1 | \mathbf{S}) \propto \int d\mathbf{E} d\sigma P(\mathbf{d}^1 | \mathbf{E}, \sigma, \mathbf{S}) \alpha \exp(-\alpha \mathbf{E}) \alpha \exp(-\alpha \sigma) \quad (2)$$

$$P(\mathbf{d}^1 | \mathbf{E}, \sigma, \mathbf{S}) = -\frac{1}{4} \sum_{i,j} \left(w_{i,j} (d_i^1 - d_j^1)^2 + 2\delta_{i,j} \sigma^2 d_i^1 d_j^1 \right) \quad (3)$$

$$w_{ij} = \begin{cases} 1/E_{ij} & (S_{ij} = 1) \\ 0 & (S_{ij} = 0) \end{cases} \quad (4)$$

$$P(\mathbf{S} | F) \propto \frac{1}{Z} \theta^n \quad (5)$$

$$P(F) \propto 1 \quad (\text{i.e. uniform distribution}) \quad (6)$$

\mathbf{E} は node 間の距離、 σ はデータのバラつきに対する制約である。 \mathbf{E} , σ はともに、指数分布を事前分布にとり、周辺化

[†] 東京大学新領域創成科学研究科

[‡] 東北大学環境科学研究科

[§] 独立行政法人産業技術総合研究所

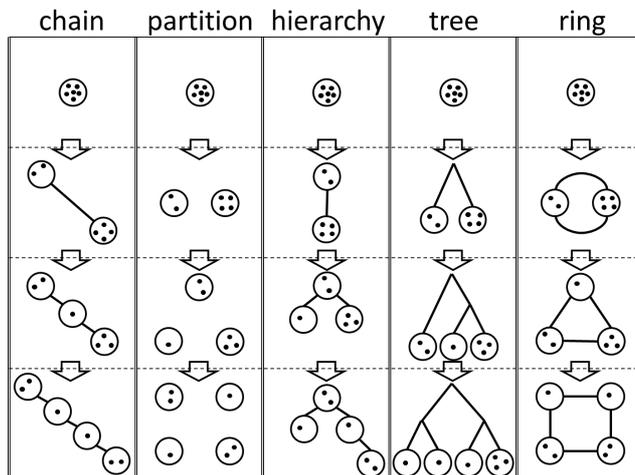


Fig. 1. グラフの複雑化. chain, partition, hierarchy, tree, ring のそれぞれについて node の数を増やしながらグラフを複雑化していく。円が node を表し、黒点が entity を表す。

する。本研究では事前分布のパラメータ $\alpha = 0.4$ とした。また、 n は node の数、 θ は $0 \leq \theta \leq 1$ の定数であり、node の数に対する制約を表す。本研究では $\theta = \exp(-1000)$ とした。 Z はデータの構造 \mathbf{S} に関する規格化項である。

2.2. アルゴリズム

Eq. 1 の評価関数をもとに最適な構造を探索した。一つの node から探索を始め、評価関数を最大化しつつ、node 分割していく (Fig. 1)[3]。また、node 分割法に規則性を持たせることで構造の種類を表現する。本研究では、chain 構造、partition 構造、hierarchy 構造、tree 構造、ring 構造について解析を行い、それぞれの最適な構造同士を比較し、最適な構造の種類 F と構造 \mathbf{S} を決定した。

3. 連想記憶モデル

ニューロン数が N の系を考える。各ニューロンは ± 1 のいずれかの値をとり、互いに相互作用している[1]。記憶させるパターン数を p 、記憶パターンを ξ^p とすると、系の自由エネルギーと自由エネルギーの停留点を表す鞍点方程式は以下のように表される。

$$f(\beta, \mathbf{m}) = \frac{1}{2} \sum_{\mu=1}^p (m^\mu)^2 - \frac{1}{\beta} \left[\ln 2 \cosh \left(\beta \sum_{\mu} m^\mu \xi^\mu \right) \right]_{\xi} \quad (7)$$

$$m^\mu = \left[\xi^\mu \tanh \left(\beta \sum_{\mu} m^\mu \xi^\mu \right) \right]_{\xi} \quad (8)$$

ただし、 $[\dots]_{\xi}$ は ξ についての期待値を表す。また、 $\beta = 1/T$ であり、物理学的知見から逆温度と呼ぶ (T は温度)。さらに、記憶パターン ξ^p と任意の状態 \mathbf{x} とのオーバーラップ m^p を、

$$m^p = \frac{1}{N} \sum_{i=1}^N \xi_i^p x_i \quad (9)$$

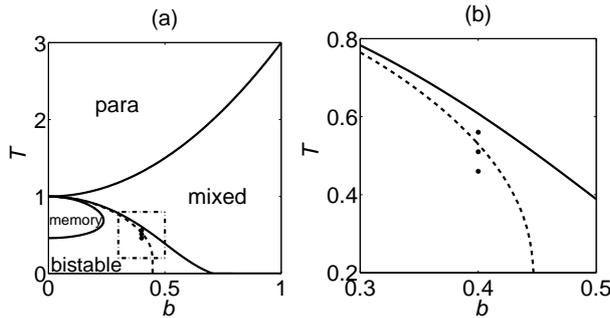


Fig. 2. (a):親パターン1つ、子パターン3つの場合の相図。破線はmemoryとmixedの自由エネルギーの大小関係が入れ替わる境界線を表す。(b):(a)の一点鎖線枠内の拡大図。点は解析したデータの生成条件を表す。

とおいた。鞍点方程式を満たす m は自由エネルギーの停留点であるので、この m における自由エネルギーの安定性を調べることで、平衡状態かどうか確認される。

本研究では、親パターン ξ に対し、相関 b を持つ子パターンを3つ生成し、この3つを記憶パターンとして解析を行った。この場合、子パターンと子パターン三つを足し合わせた混合状態が平衡状態になり得る。これら平衡状態の安定性は、温度 T および相関 b の値によって変化する。Fig. 2(a)に示す相図がその様子を表しており、para相は記憶パターン、混合状態ともに不安定な領域である。memory相は記憶パターンのみ、mixed相は混合状態のみが安定な領域である。bistable相は記憶パターン、混合状態ともに安定な領域である。また、bistable相の破線は、記憶パターンと混合状態の自由エネルギーの大小関係が逆転する境界線を表しており、破線より右方で混合状態が、左方で記憶パターンが自由エネルギー的に小さくなる[4]。

本研究では、レプリカ交換モンテカルロ法によりシミュレーションを行った。温度ごとにニューロン N 個に対し想起過程シミュレーションし、ニューロンの状態 d^i を100回サンプリングして D とした。この時、破線を境にサンプリングされやすい状態が入れ替わるため、 $b=0.4$ において、温度を変えて解析を行い、破線を境にした構造の変化を解析した(Fig. 2(b))。

4. 解析結果

解析結果をFig. 3に示す。クラスタ構造を見るために主成分分析を行い[4]、構造解析の結果(円および直線)と合わせて表示した。 F_{opt} は最適な構造の種類を表す。黒点はサンプリングしたentityを表し、円がnodeを表す。またnode間の接続の有無によりデータの構造を示す。データ数は100個とした。Fig. 3(a)~(c)は横軸が第一主成分、縦軸が第二主成分である。Fig. 3(d)~(f)は横軸が第二主成分、縦軸が第三主成分である。ただし、第一主成分が正のデータのみについてプロットした。またグラフはすべて $(b, T) = (0.4, 0.51)$ の場合の基底で統一してプロットした。

$(b, T) = (0.4, 0.46)$ の場合(Fig. 3(a)(d))、クラスタが6つ現れており、記憶パターン3つとその反転パターン3つを表す。構造解析の結果、それぞれの状態がtreeの末端に位置する構造が最適となった。

$(b, T) = (0.4, 0.56)$ の場合(Fig. 3(c)(f))、クラスタが2つ現れており、混合状態とその反転パターンを表す。構造解析の結果、クラスタ間に関係性がないことを表すpartition構造が最適となった。

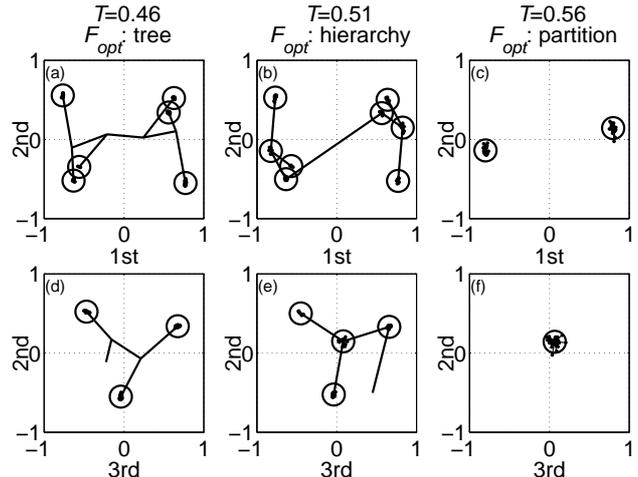


Fig. 3. Fig. 2に示した条件での解析結果。主成分分析(黒点)と構造解析(円および直線)の結果。(a)~(c)は横軸が第一主成分、縦軸が第二主成分。(d)~(f)は横軸が第二主成分、縦軸が第三主成分。ただし、第一主成分が正のデータのみ描画。(a),(d): $T = 0.46$ 。(b),(e): $T = 0.51$ 。(c),(f): $T = 0.56$ 。

$(b, T) = (0.4, 0.51)$ の場合(Fig. 3(b)(e))、クラスタが8つ現れており、前述の記憶パターンと混合状態の両方を表す。構造解析の結果、その8状態が階層的な構造を持つhierarchy構造が最適となった。特に、混合状態からすべての記憶パターンへの接続がみられ、階層性を表す結果と考えられる。

5. まとめ

本研究ではデータ駆動型アプローチによって連想記憶モデルの構造解析を行った。その結果、連想記憶モデルが階層構造を持つことを客観的手法により示すことができた。また、相図に対応して、tree構造、hierarchy構造、partition構造という、記憶パターンから混合状態への遷移の様子をグラフとして確認することができた。

今後は連想記憶モデルに埋め込む階層構造を多様化させ、複雑な構造が変化の様子を捉えることができるか検討する。

謝辞

本研究の一部は、文部科学省科学研究費補助金、課題番号25106506(永田)、課題番号23240043(麻生)、課題番号25280090、25120009(岡田)の援助の下行われた。

参考文献

- [1] J. J. Hopfield, "Neural networks and physical systems with emergent collective computational abilities", *Proceedings of the National Academy of Sciences*, 79(8), 2554-2558, 1982.
- [2] S. Amari, "Neural Theory of Association and Concept-Formation", *Biological Cybernetics*, 26(3), 175-185, 1977.
- [3] C. Kemp, J. B. Tenenbaum, "The discovery of structural form", *Proceedings of the National Academy of Sciences*, 105(31), 10687-10692, 2008.
- [4] T. Shiotsuka, K. Nagata, K. Hukushima, M. Okada, M. Inoue, "Analysis of Autocorrelation type Associative Memory Model with Hierarchical Patterns by Using PCA", *Institute of Electronics, Information, and Communication Engineers Technical Report, Neurocomputing*, 109(461), 413-418, 2010.