

解説

大脳皮質とベイジアンネットワーク

The Cerebral Cortex and Bayesian Networks

一 杉 裕 志* *産業技術総合研究所

Yuuji Ichisugi* *National Institute of Advanced Industrial Science and Technology (AIST)

1. はじめに

大脳皮質は脳の中で知能をつかさどるもっとも重要な部分であり、その情報処理原理の解明が強く望まれている。最近、複数の計算論的神経科学者が、大脳皮質の本質的メカニズムをベイジアンネットワーク [1] で説明する大変有望なモデルを提案している。大脳皮質に関する膨大な知見が、ベイジアンネットワークを核としたモデルにより急速に統一されつつあり、遠くない将来に人間のような高い知能を持ったロボットが出現することも、あり得ない話ではなくなっている。

2. 大脳皮質に関する神経科学的知見

大脳皮質とベイジアンネットワークとの関係を説明する前に、大脳皮質に関して知られている神経科学的知見について概観しよう。

大脳皮質とは、脳の表面にある厚さ 2 [mm] 程度の薄い組織で、高度な知能に最も深く関与している。以下に、大脳皮質を構成する要素を、マクロなスケールからミクロなスケールに向かって説明する (図 1)。

大脳皮質は、解剖学的な違いや他の組織との接続の仕方の違いによって領野と呼ばれる約 50 個の領域に区別されている。今日では、どの領野がどのような機能を持っているかについてはおおよそ分かっている。領野ごとに、視覚、聴覚、運動制御、行動計画、言語理解など様々な機能が担当されている。

約 50 個の領野は、相互に接続して階層的なネットワークを形作っている。領野の接続関係については、解剖学にかなり分かっている。接続されている二つの領野の間には上下関係がある場合がある。上位の領野と下位の領野は双方向に接続されており、しかもその接続には上下非対称な、ある規則性がある (後述)。なお、同じ階層に属する領野の間にも双方向の結合がある。

感覚器から入力される情報は、領野の階層によって処理

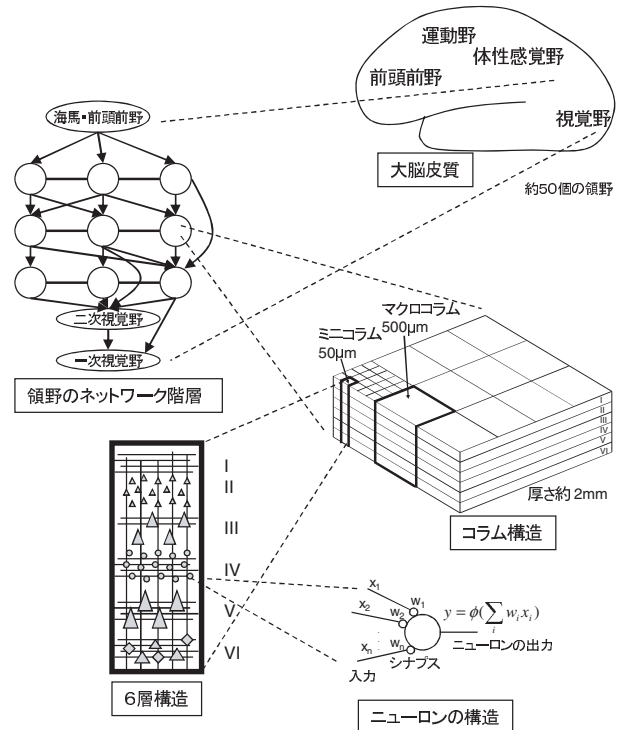


図 1 大脳皮質の構造

される。視覚刺激の場合、まず大脳皮質の一次視覚野とよばれる領野に入力される。一次視覚野で処理された視覚情報は、海馬・前頭前野に近いより上位の領野に向かってさらに階層的に処理されていく。上の階層の領野ほど、より抽象度の高い情報を表現するようになる。運動制御や行動計画に関する領野もまた階層構造をしており、上位の領野ほど抽象度の高い意思決定をする。

個々の領野は、直径 500 [µm] 程度の細長いマクロコラムとよばれる柱状の機能単位の集合である。マクロコラムの中には、入力刺激に対して同じように応答する多くのニューロンが存在する。さらにマクロコラムは、直径 50 [µm] 程度のより細いミニコラムとよばれる機能単位の集合である。ヒトの脳には 100 万個程度のマクロコラム、1 億個程度のミニコラムがある。

大脳皮質は、領野ごとに多少の違いはあるものの、どの

原稿受付 2011 年 3 月 28 日

キーワード: Cerebral Cortex, Bayesian Networks, Computational Neuroscience, Artificial Intelligence

*〒 305-8568 つくば梅園 1-1-1 中央第 2

*Tsukuba-shi, Ibaraki

領野もほぼ同じような**6層構造**をしている。層ごとに、存在するニューロンの数や種類、他の領野との接続の仕方が異なる。

個々のニューロンは他のニューロンからの入力を受け取り、演算結果を他のニューロンに送る。ニューロンとニューロンを接続する部分をシナプスとよぶ。シナプスは**結合の重み**を持っており、この重みが大きいほど情報が強く伝達される。ニューロンは、基本的には、入力のベクトルと重みのベクトルの内積計算程度の簡単な演算しか行わない。

以上が大脳皮質の基本的な構造である。パターン認識や運動制御などの脳の様々な高次機能が、同じようなコラムの繰り返し構造を持つ、たった50個程度の領野のネットワークで実現されている。これは大変不思議なことである。しかし逆に言えば、大脳皮質の基本的な動作原理さえ分かれば、脳の様々な高次機能を人工的に再現できる可能性が出てくることになる。

3. 大脳皮質とベイジアンネットの類似性

ベイジアンネットは、複数の事象の間の因果関係をネットワーク構造で表現し、同時に因果関係の強さを表す数値(条件付確率)を保持する、一種のデータ構造である。ベイジアンネットを用いれば、確率論に基づいた推論を効率的に行うことができ、近年様々な分野で応用されている。この推論は、脳の機能の一つである「直観」と似た働きをする。

大脳皮質とベイジアンネットの動作には、以下のように多くの類似性がある。

- 推論時に上下双方向かつ非対称に情報が流れる。
- 局所的かつ非同期な情報のやりとりだけで動作する。
- 非負の値のみが流れる。
- 値がしばしば正規化される。
- 学習も局所的な情報だけを使って行える。
- 文脈や事前知識を利用したロバスタな推論が行える。
- 確率論に基づいた合理的な動作をする。

このような類似性に着目し、大脳皮質の本質的機能をベイジアンネットで表現したモデルで理解する動きが広がりつつある[2]~[9]。これらのモデルは、単に脳に関する一部の実験事実を再現するだけの現象論的モデルにとどまらず、大脳皮質の学習・認識の計算論的目的、アルゴリズムとデータ構造、それを實現する神経回路に至るまで、幅広く詳細な説明を与えつつある。

以下の章で、大脳皮質とベイジアンネットの間のより詳細な対応関係について、紹介していく。

4. 視覚的注意とニューロン応答の意味

視覚刺激の情報処理過程には、「目の前の物体が何か」を処理する側頭葉に向かう腹側経路と、「物体がどの位置に

あるか」を処理する頭頂葉に向かう背側経路の二つがある。Raoはこの処理過程をノード四つの簡単なベイジアンネットで素直にモデル化した[4]。このネットワークは当然、提示された視覚刺激の特徴と位置を認識できる。しかしそれだけではなく、驚くべきことに、視覚刺激に対する注意を向けたときに、視覚野のニューロンの応答がどのように変化するかという電気生理実験の知見をも定性的に再現できるのである。このモデルでは、注意は**事前確率**、ニューロン応答は**事後確率**であると解釈している。

RaoのモデルはさらにChikkerurらによって拡張され、より広範囲の電気生理実験の知見が再現されている[9]。それに加え、同一のモデルで、「画像のどの場所に視線が行きやすいか」を予測するまったく異なる心理物理学の知見の再現にも成功している。Chikkerurらのモデルは、視覚に関する過去の複数の有名な現象論的モデルを簡単な一つのベイジアンネットを用いて統合したことになる。ベイジアンネットを核にして、多くの脳のモデルが急速に統合されていく時代の幕開けを確信させる。

5. コラム構造・6層構造の意味

5.1 6層の間の結合の規則性

すでに述べたように、上下関係にある二つの領野は双方向結合を持っており、その結合の仕方に規則性がある[11]。ボトムアップの結合は主に3層から4層に向かう。トップダウンの結合は主に5・6層から1層に向かう。

一方、コラム内では4層に入力された情報は2・3層を経由し5層から皮質以外の組織に出力されると考えられている[10]。5層の情報はさらに6層を経由して4層に入力される。

この二つの知見を合わせると、コラム内における情報処理の途中結果である3層の情報が上位領野に送られ、最終結果の5層の情報は下位領野に戻されるという、とても奇妙な構造をしていることになる(図3上)。

この構造のかなりの部分の意味は、以下で説明するように、筆者が導いた近似確率伝播アルゴリズムによって説明できる[5]。

5.2 近似確率伝播アルゴリズム

ベイジアンネット上での推論には、確率伝播アルゴリズム[1]がよく用いられる。このアルゴリズムは神経回路と似た点があるが、少し複雑であり、それをそのまま大脳皮質内の神経回路が実行しているとは考えにくい。そこで、筆者はこのアルゴリズムを近似することを思い立った。いくつかの自然な仮定に基づいて近似して導いたアルゴリズムが図2である[5]。 W_{XY} は、ノード X と Y の間の条件付確率を表す行列である。このアルゴリズムに従って各変数の値を反復して計算し、ベクトル \mathbf{b}_X の値が収束したとき、その値はノード X の事後確率を表す。

$$\begin{aligned}
 \mathbf{I}_{XY}^{t+1} &= \mathbf{z}_Y^t + \mathbf{W}_{XY} \mathbf{o}_Y^t \\
 \mathbf{o}_X^{t+1} &= \prod_{Y \in \text{children}(X)} \mathbf{I}_{XY}^{t+1} \\
 \mathbf{k}_{UX}^{t+1} &= \mathbf{W}_{UX}^T \mathbf{b}_U^t \\
 \mathbf{p}_X^{t+1} &= \sum_{U \in \text{parents}(X)} \mathbf{k}_{UX}^{t+1} \\
 \mathbf{r}_X^{t+1} &= \mathbf{o}_X^{t+1} \otimes \mathbf{p}_X^{t+1} \\
 Z_X^{t+1} &= \sum_i (\mathbf{r}_X^{t+1})_i \quad (= \|\mathbf{r}_X^{t+1}\|_1 = \mathbf{o}_X^{t+1} \cdot \mathbf{p}_X^{t+1}) \\
 \mathbf{z}_X^{t+1} &= (Z_X^{t+1}, Z_X^{t+1}, \dots, Z_X^{t+1})^T \\
 \mathbf{b}_X^{t+1} &= (1/Z_X^{t+1}) \mathbf{r}_X^{t+1} \\
 \text{where } \mathbf{x} \otimes \mathbf{y} &= (x_1 y_1, x_2 y_2, \dots, x_n y_n)^T
 \end{aligned}$$

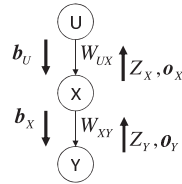


図 2 近似確率伝播アルゴリズム [5]

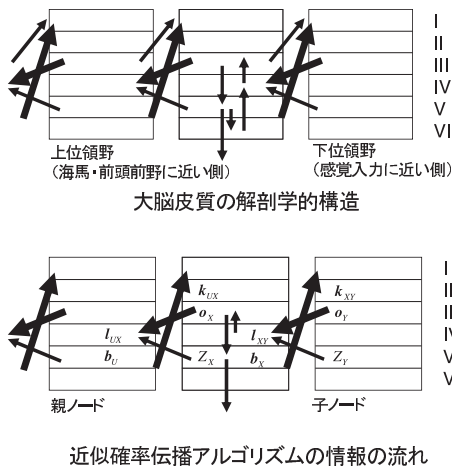


図 3 大脳皮質の 6 層の間の主な接続 (上) と、近似確率伝播アルゴリズムにおける親ノードと子ノードの間の情報のやり取りを 6 層構造に対応付けたもの (下)。近似確率伝播アルゴリズムによって、大脳皮質の層の間の主要な接続構造のかなりの部分が説明されることが分かる

5.3 アルゴリズムと解剖学的特徴との対応

アルゴリズムを単純化する目的で純粋に数学的に近似して導いた結果は、驚いたことに、前述の奇妙な解剖学的構造とよく一致する。

近似確率伝播アルゴリズムに現れる七つの変数のうち、ノード間の通信にかかわる五つの変数を、領野間の結合規則に当てはめてみると、図 3 (下) のようになる。例えば、変数 l_{XY} は子ノードの変数 Z_Y 、 o_Y から計算されるが、この三つの変数をそれぞれ 4 層、5 層、3 層に当てはめると、きれいに当てはめることができる。他の変数についても同様である。当てはめた結果は、コラム内での情報の流れのうち、4 層→2・3 層→5 層の順の情報の流れとも整合性があり、偶然とは考えにくい。

回路レベルではより詳細な対応がつく。近似確率伝播アルゴリズムは、簡単な演算を行う演算素子だけからなる固

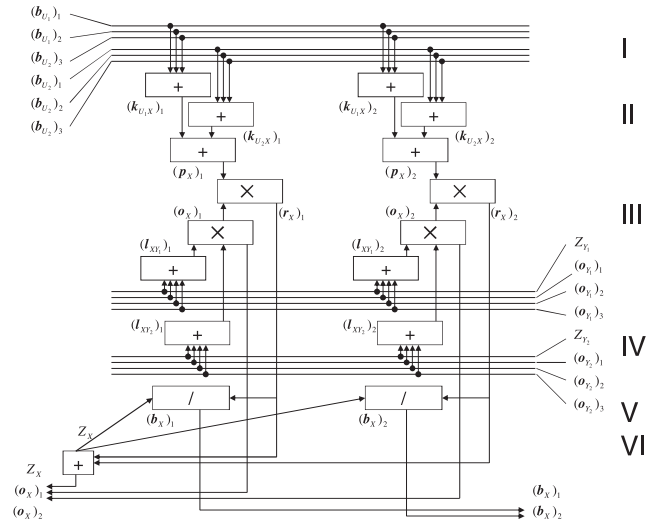


図 4 近似確率伝播アルゴリズムは、固定した回路によって実行可能であり、その構造は大脳皮質のコラム構造・6 層構造とよく一致させることができる

定した回路で実現可能である。その回路を、神経科学的知見とできるだけ合うように書いてみたものが図 4 である。この神経回路は、(1) コラム内では垂直方向だけでほとんどの情報処理が行われる、(2) 1 層、4 層には多くの水平線維が見られる、(3) 2 層および 4 層には多数のニューロンがある、という大脳皮質の主要な解剖学的知見と非常によく一致させることができている。

筆者は図 2 の近似確率伝播アルゴリズム自体の動作確認はまだ行っていないが、似た構造を持つ別のアルゴリズムの動作を確認しており、これらが大脳皮質の認識アルゴリズムの候補であると考えている。また、他の文献では大脳皮質の 6 層構造と別の推論アルゴリズムとの対応が指摘されている [3] [6] [8]。

6. BESOM モデル

人間のような高い知能を実現するためには、大規模なベイジアンネット上で動作する認識・学習アルゴリズムの詳細を明らかにする必要がある。

筆者は BESOM モデルと呼ぶ、ベイジアンネットと自己組織化マップ [12] を組み合わせた計算論的モデルを提案している [5] [13] [14]。BESOM は、入力 (観測データ) が与えられるたびに、認識ステップと学習ステップを実行する。これにより、外界を近似する確率モデルを教師なし学習する。学習ステップでは上位層の各ノードは自己組織化マップとして動作し、子ノードからの入力を圧縮し抽象化する。学習の結果はユニット間の結合の重み (条件付確率) として記憶される。認識ステップでは BESOM のすべてのノードはベイジアンネットのノード (確率変数) として動作する。

表1 BESOM, 自己組織化マップ, ベイジアンネットワーク, 大脳皮質の構成要素の間の対応

BESOM	自己組織化マップ	ベイジアンネットワーク	大脳皮質
ノード	競合層	ノード (確率変数)	マクロコラム
ユニット	入力ベクトルの要素, 競合層のユニット	確率変数が取りうる値	ミニコラム
親ノード	入力層から見た競合層	親ノード (原因)	上位領野
子ノード	競合層から見た入力層	子ノード (結果)	下位領野
ユニットの出力	入力との類似度	事後確率	コラムの発火
結合の重み	参照ベクトルの要素	条件付確率	シナプスの重み

BESOM の認識・学習アルゴリズムは未完成であり、筆者は現在、学習の安定化、アルゴリズムの理論的基礎付け、ノードの数に対してほぼ線形時間で動作する効率的なアルゴリズムの実現などに向けた取り組みを行っている [14].

BESOM を大脳皮質のモデルとして見る場合、ノードはマクロコラム、ユニットはミニコラム、ユニット間の結合の重みはシナプスの重みと解釈する。一つの領野は、多数のノードの集合である。大脳皮質、自己組織化マップ、ベイジアンネットワークの構成要素の対応をまとめると表1のようになる。

BESOM モデルにより大脳皮質の各領野が、他の領野との接続の仕方によってどのような機能を持つかが定性的に理解できるようになる [13]. 例えば、運動野は強化学習のための状態行動対を学習する領野であると素直に解釈可能である。実際に BESOM を用いた簡単なモデルで強化学習が動作可能であることは、細谷によって確認されている [7].

7. おわりに

脳の計算速度はしばしば悲観的に（大きめに）見積もられる。しかし、BESOM モデルが正しいと仮定すると、ニューロン発火のパルスのレベルまで詳細に計算する必要がないため、以下に示すように計算速度はそれほど大きくない。

仮に、ヒトの大脳皮質のニューロンをおよそ 100 億個、1 個のニューロンあたりのシナプスを 1,000 個とする、シナプスあたりの演算数が毎秒 100 回とする。すると、1PFLOPS でリアルタイムにシミュレーションできることになる。すなわち、現時点でのスーパーコンピュータですでに実現されている計算速度である。

10~20 年後には、さらなる半導体集積度の向上、アルゴリズム最適化、機能特化等により小型の計算機で人間に近い知能が実現できても不思議はない。これを搭載したロボットは将来の深刻な労働者不足を解決する手段として大きな市場が見込めるだろう。今こそ 1 人でも多くの研究者が、実用化を視野に入れた脳の情報処理の研究に取り組み始めるべき時期ではないだろうか。

参考文献

- [1] J. Pearl: Probabilistic Reasoning in Intelligent Systems: Networks of Plausible Inference. Morgan Kaufmann, 1988.
- [2] T.S. Lee and D. Mumford: "Hierarchical Bayesian inference in the visual cortex," Journal of Optical Society of America, A, vol.20, no.7, pp.1434-1448, 2003.
- [3] D. George and J. Hawkins: "A hierarchical Bayesian model of invariant pattern recognition in the visual cortex," Proc. of International Joint Conference on Neural Networks, vol.3, pp.1812-1817, 2005.
- [4] R. Rao: "Bayesian inference and attention in the visual cortex," Neuroreport vol.16, no.16, pp.1843-1848, 2005.
- [5] Y. Ichisugi: "The cerebral cortex model that self-organizes conditional probability tables and executes belief propagation," Proc. of International Joint Conference on Neural Networks, pp.1065-1070, 2007.
- [6] F. Roehrbein, J. Eggert and E. Koerner: "Bayesian Columnar Networks for Grounded Cognitive Systems," Proc. of the 30th Annual Conference of the Cognitive Science Society, pp.1423-1428, 2008.
- [7] H. Hosoya: "A motor learning neural model based on Bayesian network and reinforcement learning," Proc. of International Joint Conference on Neural Networks, pp.1251-1258, 2009.
- [8] S. Litvak and S. Ullman: "Cortical Circuitry Implementing Graphical Models," Neural Computation, vol.21, no.11, pp.3010-3056, 2009.
- [9] S. Chikkerur, T. Serre, C. Tan and T. Poggio: "What and Where: A Bayesian Inference Theory of Attention," Vision Research, vol.50, no.22, 2010.
- [10] C.D. Gilbert: "Microcircuitry of the visual-cortex," Annual review of neuroscience, vol.6, pp.217-247, 1983.
- [11] D.N. Pandya and E.H. Yeterian: 'Architecture and connections of cortical association areas,' In: Peters A, Jones EG, eds. Cerebral Cortex vol.4 Association and Auditory Cortices. pp.3-61, Plenum Press, 1985.
- [12] T. Kohonen: Self-Organizing Maps. Springer-Verlag, 1995.
- [13] 一杉裕志: 脳の情報処理原理の解明状況. 産業技術総合研究所テクニカルレポート AIST07-J00012, 2008.
- [14] 一杉裕志: 大脳皮質のアルゴリズム BESOM Ver.1.0. 産業技術総合研究所テクニカルレポート AIST09-J00006, 2009.



一杉裕志 (Yuuji Ichisugi)

1990 年東京工業大学大学院情報科学専攻修士課程修了。1993 年東京大学大学院情報科学専攻博士課程修了。博士 (理学)。同年電子技術総合研究所 (2001 年より産業技術総合研究所) 入所。プログラミング言語, ソフトウェア工学の研究に従事。2005 年より計算論的神経科学の研究に従事。日本神経回路学会, 情報処理学会, ACM の会員。(日本ロボット学会正会員)