

神経活動の観測による直接的知識転写

Direct Knowledge Copy based on Neuron Behavior Observation

瀧 寛和*¹
Hirokazu Taki

三浦 浩一*¹
Hirokazu Miura

松田 憲幸*¹
Noriyuki Matsuda

曾我 真人*¹
Masato Soga

*¹ 和歌山大学 システム工学部
Faculty of Systems Engineering, Wakayama University

This paper describes the knowledge acquisition method from natural neural network to artificial neural network. We study the method of the knowledge copy using neuron behavior observation. We considered whether the human brain can be copied into the machine completely or not, and the possibility of the data mining method. In the observation, there is not complete information about the behavior. Therefore, we introduce the new functional neuron which is activated in order to statistical rule. Our target is full brain function copy and making human function on the machine.

1. はじめに

様々な知識システムの構築が行われ、限定的な分野で人の知能を超えるシステムも存在するようになってきた。しかし、普通の人のような柔軟な知性を身に付けた人工知能は未だ実現されていない。人工知能機能的な研究は、人の知的活動の機械的な実現により、多くの推論メカニズムが模倣されてきた、また、近年は、マルチエージェントなど複数の異なる行動を持つ集団としての社会知能の機能分析や機能実現が研究されてきている。エキスパートシステムの知識ベースや推論メカニズムも重要だが圧倒的に知識や情報の量が非常に重要となっている。非常に整理された知識ベースの構築には、分野の知識体系をオントロジーとしてまとめることになる、このオントロジーは、分野独立であるため、広い問題や柔軟な発想を必要とする問題への利用はできない。そこで、多くの情報を次々と記憶して、柔軟な発想につながる新たな仕組みの人工知能の枠組みが望まれている。筆者らは、この課題に対して、人の短期記憶から長期記憶への記憶の転写メカニズムの検討として、「異なる構造の神経回路網間の知識転写」[瀧 2008]を研究し、オリジナルの神経回路から例を生成し、転写先の神経回路に、その例を与えることで、知識を転写する方式での効率的な例の生成を求めた。学習において、例の生成の元となる学習する際の属性を決める仕組みが人によって異なることが考えられる。そこで、図形の分類学習での注目属性について、実験を行った。しかし、単純な学習でも個人差が大きいことが分かった。つまり、脳内の知識転写で利用される属性の差は、単一の仕組みで属性を見出す

のでなく、属性の選択にも既知の知識が利用されている可能性があると考えられる。

この知識転写は、非常に効率が悪いいため、より効率的な記憶の仕組みとして、短期記憶と長期記憶に同時に記憶する仕組みを検討した、その際の仮説として、「短期記憶が長期記憶に対するインデックスの生成を行う」[瀧 2009, 2010]を想定した。そのインデックスを生成する仕組みが、人に依存しない可能性も仮定した。画像の属性を色々な形で示し、どのインデックスが記憶に関係するかを実験したが、インデックスも複数生成されている可能性が高く、特定のインデックス生成が行われることは確認できなかった。短期記憶がインデックス生成であれば、複数のインデックス生成の仕組みを考える必要があるとの示唆である。「異なる構造の神経回路網間の知識転写」では、例を介した知識転写は、知識そのものでなく、例を介することから、間接的な知識の転写である。直接的な知識の転写を考えるにあたり、神経活動の現象を観測し、その観測から機能や記憶を推定することが考えられる。神経回路の機能を客観的に把握するには、「脳神経の活動の観測」から「人工的な神経回路網に機能転写」を行う仕組みについて、本稿では、その仕組みの要点を検討する。

2. 神経活動の観測による神経回路網の推定の可能性について

脳全体を観測して、機械にその機能を移すことができるかを考える（以下、Brain-Copy と略す）。BC では、脳の神経細胞数規模、シナプス数を見ると、実際には、観測で完全な形の転写は現時点では非常に困難なことがわかる。脳神経細胞の数は、様々な推定があるが、理化学研究所の説明[理研 1999]では、大脳で数百億個、小脳で 1000 億個と

いう数である（合計 10 の 15 乗）。他の文献では、およそ 100billion（10 の 15 乗）程度はそれ以上となっている。また、シナプスは、1つの神経細胞に 1000 から 1万程度が結合している[Ramachandran 1998]とされている。つまり、各神経細胞が独立していないとして、各神経細胞の発火状態すべてを見るには、2 の 1 千数百億乗の組合せを観測する必要がある。これは、独立性を無視しているので、1つの神経細胞で見れば、入力が 100 シナプス結合なら入力 (ON/OFF) の組合せは 2 の 100 乗（10 の 31 乗）、1000 シナプス結合なら、入力の組合せは 2 の 1000 乗（10 の 302 乗程度）となる。この 1つの神経細胞についてもすべての組合せを観測するのは、事実上、困難と言える。一方、ニューラルネットワークを利用した認識を見てみると、たとえば、256x256Pixel 画像程度の認識にも利用されているが、この場合は、各 Pixel が 2 値であれば、シナプス 65536 個に相当するので、2 の 65536 乗となる。このような 1 画像の認識については、学習も認識もできている。このことから、入力すべてを考える神経細胞状態でなくても、よく現れるパターン（神経細胞の発火集合のパターン）を観測できれば、1つの神経細胞（入力シナプスが 1000 から 1万程度）は、その機能を転写（人工神経回路網上に転写）できると考えても無理はない。

3. 構造の分からない神経回路網の状態を推定

脳科学や神経科学の分野では、生きた神経細胞の断片の発火挙動を観測する技術が進んでいる、また、発火パターンの解析から、限られた数の神経間関係をデータマイニングすることで、ニューロンの結合を求めることも行われている[Takahashi 2010]。発火パターンの再現に注力したデータマイニングは、さらに、研究が進み精度向上が望めます。ここで、我々の研究の立場としては、間接的な知識転写の方針として、異なる神経回路網間での機能（情報・知識）の転写を主眼においているため、構造の再現はできなくてもよく、機能の再現（多少精度が落ちて近似的に機能が実現できる）を目指した場合の直接的な知識転写をどのようにすればよいかが課題となる。

4. 観測できない部分や不正確部分への対応

シナプス（そのシナプスへの信号供給の神経細胞）すべてが観測できなくても、おおよその神経回路の機能が推定できると考える方式について、説明する。

(1) 単純に入力の一部が欠落する場合

閾値論理回路を構成するニューラルネットワーク（1ユニ

ット）を考える

$$\sum X_i \cdot W_i (\text{入力と重みの積の和}) - \theta (\text{閾値}) = 0$$

これが、空間の線形分離の超空間（たとえば、1000 次元 $i=1, \dots, 1000$ ）となるが、ここで、いくつかの X_i が観測できなければ、少ない次元への超平面の写像を求めていることになる。次数が少なくなるほど精度は落ちるが、観測した範囲の機能は求められる（どの入力が必要であるかは、機械的には、わからない）。

(2) 欠落する情報の補間について

多くの発火する神経細胞の情報から、同じパターンで特定の神経細胞の発火が異なる場合には、情報の欠落の可能性が高い。その場合には、他の神経細胞の発火パターンとの統計的相関情報から、確率的挙動を持つ神経細胞を仮定して、神経細胞挙動を制御することで、転写の精度の低下を防ぐことができる。

(3) 発火事例からの機能推定

この神経細胞の発火集合からの機能推定（機能転写）は、間接的な神経回路網の知識転写と 1つの神経細胞であれば、実は同じことを行っているだけである（入力と出力の組を利用した学習）。異なるのは、「転写を意図した例の生成」か「例の観測」である。ただ、良く起こる例の発生パターンに注目すれば、前述の例の不足問題への解答になる。

（脳があらゆる例を学習しているとは限らない。少ない例で学習し、よく間違えるのは我々の常である）

5. まとめ

本論文では、神経回路網の挙動を観測することで、その機能を人工神経回路網に転写する方法について検討した。その中で、観測が不十分である点を考慮して、確率的挙動を持つ神経細胞の導入で、機能実現精度を向上させることを試みた。

参考文献

- [瀧 2010] 瀧, 三浦, 松田, 曾我: インデックス脳におけるスナップ記憶と短間隔記憶について, 1A4-1, 人工知能学会全国大会予稿集, 2010.
- [瀧 2009] 瀧, 三浦, 松田, 曾我, 安部: 表象と表象のない知識の相互作用による問題解決, 3C4-2, 人工知能学会全国大会予稿集, 2009.
- [瀧 2008] 瀧, 曾我, 三浦, 松田, 堀, 安部: 身体知伝達のための例示の生成, 1B2-7, 人工知能学会全国大会予稿集, 2008.
- [Takahashi 2010] N. Takahashi, T. Sasaki, W. Matsumoto, N. Matsuki, and Y. Ikegaya: Circuit topology for synchronizing neurons in spontaneously active networks, PNAS, vol.107, No.22, 2010.
- [Ramachandran 1998] Ramachandran, V.S. and Sandra Blakeslee: *Phantoms in the Brain: Probing the Mysteries of the Human Mind*. New York: William Morrow & Co., 1998.
- [理研 1999] <http://www.brain.riken.jp/jp/aware/neurons.html>