解説

脳を創るために脳の機能を再現して構造を評価する

上 村 卓 也

NTT コミュニケーション科学基礎研究所*

Reproducing Brain Functions and Evaluating Its Structure for Creating the Brain

Takuya Koumura

NTT Communication Science Laboratories*

概要

脳を創れるかどうかを考える前に、まずは身近な例としてケーキを創れるかどうかを考えてみる。創るべきケーキの要素を考え、それらをケーキの「構造」と「機能」の大きく2つに分類してみる。すると、ケーキを創るための方法も、ケーキの構造を再現してから機能を評価する方法と、ケーキの機能を再現してから構造を評価する方法の、大きく2つに分類できることがわかる。次に、本題である脳を創ることについて考える。ケーキを創るときと同じ議論を適用し、脳の構造を再現してから機能を評価する方法と、脳の機能を再現してから構造を評価する方法の、2通りの方法を考える。脳を計算機上に創るための具体的な方法を紹介し、最後に、脳の機能を再現してから構造を評価する方法の意義や妥当性などについて述べる。レストラン級のケーキを家で創ることが簡単ではないが可能であるように、脳を創ることも簡単ではないが、本稿で紹介した2通りの方法がお互いにうまく補完しあっていけば、いずれは可能になるだろう。

1. ケーキの構造と機能を創る

脳を創れるか？と言われるとなんだか難しく感じてしまうかもしれない。そこで、まずはケーキを創れるかどうか考えるところから始めてみる。

先日あなたはレストランでケーキを食べ、その美味しさに感動したため家でも作ってみたいとなった。何ができればレストランのケーキを創ったことになるだろう？例えば、小麦粉や砂糖をレストランと同じ分量だけ用意し、レストランのケーキと同じ高さのスポンジを焼き、同じ位置に同じ種類の果物を配置し、クリームを同じ厚さに塗って同じ形に造形すると、レストランのケーキを作ったといえるだろうか。たぶんこれだけではダメで、できた自作ケーキを食べてみて、レストランと同じ味や香りや柔らかさを感じられて初めて、レストランのケーキを作ったといえるだろう。

逆に、次のような方法も可能だろう。まずは小麦粉や砂糖の分量をレストランと同じにしたり、スポンジの高さをレストランと合わせたりすることにはこだわらない。手に入る材料を様々に組み合わせる色々な種

類のケーキを作って食べてみて、食べたときの味や香りや柔らかさがレストランのケーキと一致するような作り方を探索する。ただし、レストランのケーキと同じ味や香りや柔らかさのケーキを作ったからといって、それはレストラン風の自作ケーキであり、レストランのケーキを作ったとはいえない。レストランのケーキと自作ケーキを同じ角度で切ってみて、スポンジの高さやスポンジに挟まっている果物の配置が同じになっていることを確認して初めて、レストランのケーキを作ったといえるだろう。

実は、これら2つの方法ではどちらも、ケーキを2つの要素に分けて考えている。1つめは、果物の種類と配置やクリームの厚さなど、ケーキを構成している組成やその構造である。ここでは、これをケーキの「構造」と呼ぶ。2つめは、完成したケーキの味や香りなど、それが食べる人にとってどう感じられるか、である。これはケーキの「機能」といいいいだろう。このように名付けると、レストランのケーキを創るためには、その構造と機能を両方とも再現しなければならないことがわかる(図1)。上述の2つの方法は、次のように書ける。

* 〒243-0198 神奈川県厚木市森の里若宮 3-1

図1 ケーキの構造と機能の例

- I. まずレストランのケーキの構造を調べ、その構造を再現するようなケーキを作る．できた自作ケーキの機能を調べ、レストランのケーキの機能を再現できているか、評価する．
- II. まずレストランのケーキの機能を調べ、その機能を再現するようなケーキを作る．できた自作ケーキの構造を調べ、レストランのケーキの構造を再現できているか、評価する．

2. ケーキの数理モデル

自然言語で記述するよりもある程度数式で記述したほうがわかりやすいかもしれない．あるケーキを z とする．これ自体はこのままでは観測できない． z についての情報を得るには、それに対して何らかの操作 g を適用し、 $y = g(z)$ を得る必要がある．例えば、 g が半分に切って断面を観察するという操作のとき、 y はそこに見えた果物の種類や配置となる．また、 g が食べる・見るという行為のとき、 y は z の味や飾り付けの見た目である．このように、適切な g を用いると、観測不可能であった z の構造や機能を調べることができる．

レストランのケーキを z_{true} 、家で作るケーキを z_{model} とすると、ここでの目的は z_{model} を z_{true} に近づけることである．しかしながら、 z は観測不可能であるため、どのように z_{model} を動かせばよいか、そのままではわからない．そこで、ある g を用いて $y_{model} = g(z_{model})$

が $y_{true} = g(z_{true})$ に近づくように、 z_{model} を動かしていく．ここで用いた g と y を、 g_{train} 、 y_{train} と呼ぶことにする． $y_{model,train}$ と $y_{true,train}$ がどのくらい近いを表す誤差関数を E_{train} とすると、この操作は次のように書ける．

$$z_{model} = \arg \min E_{train}(y_{model,train}, y_{true,train})$$

$$y_{model,train} = g_{train}(z_{model})$$

$$y_{true,train} = g_{train}(z_{true})$$

実際は g_{train} と y_{train} をそれぞれ複数個用いてケーキを作ることが多い．構造については例えば、スポンジの高さと果物の種類とクリームの高さの全てを考慮し、機能については、味と香りと柔らかさの全てを考慮すると、よりレストランに近いケーキを作れそうである．よって上記の式は実際は次のようになる．

$$z_{model} = \arg \min \sum_i E_{train,i}(y_{model,train,i}, y_{true,train,i})$$

次に、 g_{train} とは異なる g_{test} を用いて得られた $y_{model,test}$ が $y_{true,test}$ に似ているかどうかを E_{test} によって評価する．

$$\sum_i E_{test,i}(y_{model,test,i}, y_{true,test,i})$$

$$y_{model,test,i} = g_{test,i}(z_{model})$$

$$y_{true,test,i} = g_{test,i}(z_{true})$$

上述の2つの方法に当てはめると、構造を再現して機能を評価する方法では、 $(train, test) = (\text{構造}, \text{機能})$ となる．機能を再現して構造を評価する方法では、 $(train, test) = (\text{機能}, \text{構造})$ となる．

ここでは議論のために便宜的にケーキを変数で表し

図2 脳の構造と機能の例

たが、実は、変数や数式で表した時点で、これは真のケーキではなくケーキの数理モデルになってる。しかしながら、本稿では、便宜的に上記 z_{true} を真のケーキと考え、これを創る（＝再現するような z_{model} を作る）ためにどうするかという議論をすることにする。

z_{model} を評価するための g_{test} や y_{test} に g_{train} や y_{train} と同じものを使わないことが重要である。例えば、スポンジの高さ (y_{train}) がレストランと同じになるように作った結果、スポンジの高さ (y_{test}) がレストランと同じになりました、といわれても、それは自明でしょう、と思われてしまう。

3. 脳の機能と構造を創る

ここまでで、レストランのケーキを家でも創れることがわかった。あとは同じようにして脳を創るだけである。ここまでの議論の「ケーキ」を「脳」に置き換えるだけで脳を創れると思うと、あまり難しくはなさそうだ。脳を創るときも、ケーキのときと同じように、構造と機能という2つの側面に分けて考える（図2）。脳の場合、例えば次のようなものを構造と考えることができるだろう。

- 神経細胞やグリア細胞などの細胞内分子の種類と数。例えば、細胞内や膜状に存在するタンパク質や、その複合体からなるチャネルや受容体など。また、それらが引き起こす生化学反応。
- 細胞などの組織学的な性質。形状や配置や、神経細胞間の接続からなるネットワーク。どの細胞とどの細胞が接続しているかという状態など。
- 神経細胞の生理学的な性質。例えば、どのような神

経伝達物質に反応してどのような膜電位変化や発火のパターンを示すのか。

これらの一部は「機構（もしくは神経機構）」と呼ばれることもあるが、ここではまとめて「構造」と呼んでしまうことにする。

一方、脳の機能には、例えば次のようなものがあるだろう。

- 脳の活動によって得られる認知的な機能。例えば、刺激の認識や、それと記憶に基づいた意思決定。
- 認知に基づいた身体運動。例えば、物体に手を伸ばす、など、身体の一部を介した反応。

ケーキのときと同様に、脳を創るためには、これら全てを再現しなくてはならない。逆に言うと、これらの全てを再現できたとき、脳を創れた、といえるだろう。つまり、脳を創れるか？ という問いは、これらを全て再現できるか？ という問いに置き換わる。

脳を構造と機能の2つに分けて考えると、ケーキを創るときと同じように、脳を創るための方法を2通り考えることができる（図3）。

I. 構造を再現してから機能を評価する。まず、神経細胞の性質やその組織学的なつながりなどの構造を再現する。その際、作った脳がどのような機能を持つかは（もしくは、そもそも何らかの機能を持つのかは）、一旦は考えない。脳の構造を再現できたら、次に、刺激が認識されるパターンなどの認知機能や生成される運動パターンを真の脳と比較することにより、その機能を評価する。

II. 機能を再現してから構造を評価する。まず、刺激

図3 構造を再現して機能の評価する方法（右下の実線）と、機能を再現して構造を評価する方法（左上の実線）の模式図。右上に到達できたとき、つまり、脳の構造も機能も再現できたとき、脳を創れた、といえるだろう。実際の研究ではこれらの中間に位置するものが多い（破線）。

の認識などの認知機能や生成される運動パターンについて、真の脳の機能を再現する。その際、構造を再現することは考えず、機能を再現するためにはどのような方法を用いてもよいことにする。次に、できた脳の構造に対して解剖学的方法や生理学的な計測手法を適用し、結果を真の脳と比較する。

もちろん、脳の構造・機能は、周囲の環境や生体内の脳以外の器官と相互に依存しているので、脳単体を創ただけでは生体内の脳を再現することはできないだろう。例えば、脳の置かれている環境として、周囲の物質中の分子濃度や脳自体の物理的加速度、脳以外の神経系の活動などがある。さらに、それらの時間変化も重要である。そのため、脳単体を創ることの意義は小さいと思う人もいるかもしれないが、そんなことはない。これは、ケーキを創るときにも同様の問題が存在することを考えればわかるだろう。ケーキの状態は皿の温度や周囲の気温と相互に依存している。味や香りなどは食べる人の生理的・心理的状態とも相互に依存するだろう。しかしながら、これらの問題を一旦おいておいて、家でケーキ単体を創ることに異議を唱える人は少ないだろう。同様に、脳単体を創ることに、十分意義はある。ここでは、一旦、脳と他の器官を切り離せると仮定して、脳を一つのシステムとし、脳とそこへの入力を切り離して考える。仮に脳が他の身体器官と切り離せるとすると、想定できる全ての入力について脳と同じ状態を再現できたとき、脳を創ったと言ってよいのではないだろうか。

脳を創るための方法を構造と機能に分けるという考

え方は私が勝手に提案しているだけだが、類似の概念は他の論文でも提案されている¹⁾。本稿で構造と機能と呼んでいるものは、ここでの認知的妥当性と生物学的妥当性と呼ばれている。また、構造を再現して機能で評価する方法は、ボトムアップと呼ばれ、機能を再現して構造を評価する方法は、トップダウンと呼ばれている。微妙に異なる点もあるが、考え方は概ね類似している。自分以外にも類似の考えを持つ仲間がいることはとても心強い。

4. 脳の数理モデル

次に、ケーキのときと同様に、数式で表してみる。ケーキを創るときは、ケーキ自体を z とおいた。脳を創るときは、前段落で述べたように、脳への入力とそのときの脳の内部状態に分けて考えたほうがわかりやすい。そこで、脳を、刺激 x が来た時に内部状態が z になるシステム f と考える。これは、 $z = f(x)$ と書けるだろう。そして、脳の内部状態 z について知るために、何かしらの観測方法 g を用いて結果 y を得る。これは、 $y = g(z)$ と書ける。例えば、 g が神経活動計測の場合、 y は膜電位変化や発火時刻などである。 g が生体の行動の場合、 y は例えば選択されたボタンの番号やその際の身体運動の軌跡などが考えられる。

ケーキの場合と同様に、数式で表した時点で、これは真の脳ではなく、脳のモデルになっている。しかしながら、本稿では、 f_{true} を真の脳と考え、これを創る（＝再現するような f_{model} を作る）ためにどうするかという議論をすることにする。

また、ここまで、脳をどのように創るか、具体的な方法は議論していなかった。脳を創るというと、まず思い浮かぶのは、臓器培養だろうか。上述のように他の身体器官と切り離せないという考えのもと、ロボットの一部として脳を創るという試みもあるだろう。本稿では、脳を計算機上に創ることを考える。計算機上では、現実ではあり得ないような環境を仮定したり、現実の脳より構造や機能を単純化して本質的な要素を探したりできる。そのため、どうすれば脳を創れるかを考えるために、計算機はとても便利な道具となる。

真の脳 f_{true} を計算機上に再現するためには、何かしらの関数 f_{model} を f_{true} のモデルと考え、 f_{model} を f_{true} に近づけることを目指すとよいだろう。そのために、これもケーキの時と同様に、 $y_{model,train} = g_{train}(f_{model}(x))$ が $y_{true,train} = g_{train}(f_{true}(x))$ に近くなるように f_{model} を決めることになる。実際は、何らかの誤差関数 $E_{train}(y_{model,train}, y_{true,train})$ を考え、それを最小化することが多い。次に、 g_{train} とは異な

る g_{test} による $y_{true,test}$ と $y_{true,model}$ を比較することでモデルを評価する。この際、ケーキのときと同じく、複数の g と y を用いることが多い。ケーキのときと同じく、次のように書ける。

モデルの最適化：

$$f_{model} = \arg \min \sum_i E_{train,i}(y_{model,train,i}, y_{true,train,i})$$

$$y_{model,train,i} = g_{train,i}(z_{model})$$

$$z_{model} = f_{model}(x)$$

$$y_{true,train,i} = g_{train,i}(z_{true})$$

$$z_{true} = f_{true}(x)$$

モデルの評価：

$$\sum_i E_{test,i}(y_{model,test,i}, y_{true,test,i})$$

$$y_{model,test,i} = g_{test,i}(z_{model})$$

$$z_{model} = f_{model}(x)$$

$$y_{true,test,i} = g_{test,i}(z_{true})$$

$$z_{true} = f_{true}(x)$$

$(train, test) = (\text{構造}, \text{機能})$ または $(train, test) = (\text{機能}, \text{構造})$

実際は、 f_{model} を何かしらのパラメータ θ を持った関数 $f_{model}(x, \theta)$ とすることが多い。こうすることで、 f_{true} のモデリングのために、関数 f_{model} の形の決定とそのパラメータ θ の推定に分けて考えることができるようになる。ある関数 f_{model} を仮定できたとすると、パラメータ推定は $\theta = \arg \min \sum_i E_{train,i}(y_{model,train,i}, y_{true,train,i})$ と書ける。

5. 脳を創るための方法の具体例

ここから、計算機上に脳を創るための方法の例を具体的に紹介する。ボトムアップ（構造の再現→機能の評価）な方法と、トップダウン（機能の再現→構造の評価）な方法に分けて紹介する。

5.1 構造を再現してから機能进行评估する方法の例

細胞内の生化学反応をモデリングするためには、分子濃度を変数 z とした微分方程式や、分子数を変数 z としたマスター方程式を f_{model} として用いることができる²⁻⁴⁾。調整すべきパラメータ θ は化学反応の速度定数などである。脳組織中のイオン濃度計測などの g_{train} により真の速度定数 $y_{true,train}$ を推定する²⁾。モデル上では速度定数を明示的に記述しているので、 $y_{model,train}$ は θ そのものである。細胞に様々な刺激 x を与えたときの分子濃度を評価すべき機能 $y_{true,test}$ と考えたとき、 $y_{model,test}$ は z そのものである。 x の具体例としては、細胞への電氣的・化学的な刺激や⁵⁾、生体への物理刺激などがある。

神経細胞の生理学的な反応特性をモデリングするためには、例えば膜電位を変数 z とした微分方程式を

f_{model} として用いることができる⁶⁾。例えば有名な具体例として、Hodgkin-Huxley 方程式やそれを単純化した Integrate-and-fire モデルなどがある^{7,8)}。調整すべきパラメータ θ はイオンチャネルなどの膜タンパク質の性質などである。脳から取り出した神経細胞 f_{true} に電流 x を加えたときのイオンチャネル付近の電位変化を $y_{true,train}$ とした場合、モデルにも同じ x を加えたときの同種のイオンチャネルの内部状態 z_{model} を $y_{model,train}$ として $y_{model,train}$ が $y_{true,train}$ に近づくように θ を調整したり、 $y_{true,train}$ からイオンチャネルの反応特性 θ を直接推定したりできる⁶⁾。 g_{test} は g_{train} と異なる計測手法にすべきなので、例えば脳 f_{true} に直接電極を刺し、生体に聴覚刺激などの x を提示したときの細胞膜付近の電位変化を $y_{true,test}$ とすることができる⁹⁾。もしくは、多数の神経細胞による大局的な脳活動を $y_{true,test}$ とすることもできる。それを記録する方法 g_{test} としては、頭皮上からの電気信号の記録や脳活動の撮像などがある。モデル上では、 f_{model} 内の神経細胞の活動 z_{model} を時空間的に平均したりして統合したものを $y_{model,test}$ とすることが多い^{10,11)}。

脳部位による信号処理を再現するためには、例えば脳部位への入出力の関係を信号処理モデル f_{model} で表現する。例えば、聴覚系の信号処理モデルとしてバンドパスフィルタバンクや非線形ゲイン調整^{11,12)}、視覚系の信号処理モデルとして時空間フィルタバンクなどがある¹³⁾。調整すべきパラメータ θ はフィルタの係数や非線形関数のパラメータなどである。例えば、心理物理学的な刺激検出実験 g_{train} や、神経活動の周波数チューニングから推定されたフィルタ係数などが、 $y_{true,train}$ となる¹⁴⁾。 $y_{model,train}$ は θ そのものとなる。刺激への心理物理学的な反応特性を評価すべき機能 $y_{true,test}$ とした場合、モデルに対しても同じ刺激 x を入力し同じ方法で反応特性を評価して $y_{model,test}$ とする¹⁵⁾。他にも、モデルの内部状態 z を用いて刺激認識 g_{test} を行い、その精度 $y_{model,test}$ を用いて評価することもできる¹⁶⁾。

5.2 機能を再現してから構造を評価する方法の例

音などの外界の物理刺激 x を効率良く表現することが脳 f_{true} の機能だと考え、この機能を再現するモデル f_{model} を構築することができる。刺激を入力とし、入力と同じ値を出力するモデルを構築する。モデルに何の制約も課さないと、入力を素通りさせて出力するようなモデルが最適となってしまう、これでは意味がない。そこで、モデルに何らかの制約を課し、入力と同じ値を出力することが簡単にはできないようにする。例えば、刺激をなるべく少数の基底の重み付き線形和

で表現するようなモデルを考えることができる¹⁷⁻¹⁹⁾. f_{model} をこのように定義すると、基底の値が θ 、重みの値が z となる. このようなモデルは、脳が実際に刺激をどの程度効率良く表現できているかを考慮しなくても構築でき、その場合 $y_{true,train}$ は登場しない. ある刺激 x を表現するために用いる基底の個数なるべく少なくなるような制約、つまり重み z がなるべく疎になるような制約のもと、刺激 x とモデルの出力が近くなるように θ を調整する. モデルを評価するためには、例えば、モデルを刺激に対する線形フィルタとみなす. フィルタの係数は基底の値そのものである. 神経活動の周波数チューニングから推定されたフィルタ係数などを $y_{true,test}$ とし、モデルのフィルタ係数 $y_{model,test}$ と比較してモデルを評価する.

音などの外界の物理刺激 x を精度良く認識することが脳 f_{true} の機能だと考え、この機能を再現するモデル f_{model} を構築することもできる. 刺激を認識する機能、とは、例えば刺激 x が音の場合、音源の種類(車から出た音なのか犬の鳴き声なのか、など)や音源の位置などを同定する機能のことである²⁰⁾. もしくは、2つの音が同一の音源から出た音なのか異なる音源から出た音なのかを判別するという機能を考えてもよい²¹⁾. モデルを構築するためには、まず刺激 x とその正しい答え $y_{true,train}$ を用意する. そして、 x を入力として内部状態 z_{model} を計算し、 z_{model} を元に認識結果 $y_{model,train}$ を出力するモデルを用いて、 $y_{model,train}$ が $y_{true,train}$ に近づくようにモデルのパラメータを調整する. 例えば、ニューラルネットワークを教師あり学習や自己教師あり学習によって訓練すると、このような機能を比較的容易に実現することができる²²⁾. この場合、ニューラルネットワークの素子間の結合強度などが θ になる. モデルは様々な方法を用いて評価することができる. まず、機能を評価するために、モデル構築の際に使わなかった刺激 x に対する認識結果を $y_{model,test}$ とし、真の脳による認識結果 $y_{true,test}$ と比較することができる^{23,24)}. 真の脳による認識結果とは、動物が x を認識したときの回答である. これは、脳による x の表現 z_{true} を元に身体運動 g_{true} を介して出力された選択であると考えることができる. このような真の脳による認識結果は間違いを含んでいる可能性もあり、その間違いのパターンも含めて脳の機能を再現できているかを調べることは重要である. 構造を評価するためには、例えば、ある刺激 x を提示したときの神経活動 z_{true} の統計量を $y_{true,test}$ とし、同じ x を入力したときの z_{model} から計算される同じ統計量を $y_{model,test}$ とし比較できる²⁵⁻²⁷⁾. ここで、統計量としては、例えば活

動の時間平均や刺激のパターンへの同期性などを用いることができる. また、心理物理学の実験 g_{test} によって測定される刺激の検出感度を $y_{true,test}$ とすると、同じ刺激をモデルに入力したときに z_{model} から計算される刺激の検出感度を $y_{model,test}$ とし比較できる^{28,29)}. 他にも、 f_{model} が畳み込みニューラルネットワークの場合、畳み込み操作を線形フィルタであるとみなすことができる. フィルタの周波数特性を $y_{model,test}$ とし、神経活動の周波数チューニング特性を $y_{true,test}$ とし比較することもできる^{30,31)}.

5.3 実際の研究では

本稿では議論のわかりやすさのために、脳を創るための方法を2つにわけて考えている. しかし、実際の研究ではこれら両極端な方法は少なく、この中間に位置するようなものが多い(図3破線). 例えば、刺激の認識精度(機能)を動物と同等に近づけるために、末梢神経の反応特性(構造)を仮定して刺激の前処理を行い、モデルの枠組みに脳内の神経回路(構造)を模擬した神経間の接続を用いる、などの手順が、ひとつのモデルを作るために行われたりする.

何が構造で何が機能なのかという定義も、上記の具体例でみてきたように、普通は相対的なものである. 例えば、神経細胞の生理学的な性質という構造により、心理物理学的な刺激検出という機能が実現されている、という考えのもと、心理物理学的な実験結果を再現するモデルが機能モデルと呼ばれることもある¹⁵⁾. 一方で、刺激検出という構造により、生体が生存のために行動するという機能が実現されている、という考えのもと、心理物理学的な刺激の検出感度が構造と位置づけられることもある^{28,29)}.

6. 機能を再現してから構造を評価する

歴史的には、脳の計算機モデリングの研究は、ボトムアップ(構造の再現→機能の評価)な研究が多かったようだ. これは、ひとつには、構造を明示的に作り込んだモデルのほうが、直感的に、脳のモデルとして妥当性が高いと考える人が多かった、という理由があるのかもしれない. 機能は多少再現できていなくてもかまわないが、構造は脳と矛盾する点があってはいけない、という直感を持つ人が多かったのかもしれない. しかし、実際はそんなことは全くない. 構造と機能は対等に扱われるべきであり、脳にとってはどちらも同等に重要なものである. 一旦は機能について考えずに構造をしっかりと再現し、後から機能を評価するという方法も、一旦は構造について考えずに機能をしっかりと再現し、後から構造を評価するという方法も、学

術的にはどちらも同等に価値が高い。脳を創るためにはお互いが補完しあっていくべきだろう。

実際、(計算機モデルではなく)動物モデルの研究では、ヒトの脳の持つある機能の構造や機構を知りたいとき、構造や機構は不明でもヒトの脳と似た機能を示す動物種をモデルとして扱うことはよくある。このような研究では、まず両種の脳の機能が似ていることを確認し、その後で構造を比較する。ヒトの脳の構造とモデル動物の脳の構造の、どこが似ていてどこが似ていないのかを明らかにすることで、対象とする機能を実現するための脳の構造について考察することができるようになる。このような方法が広く受け入れられていることから、この「動物モデル」が「計算機モデル」に置き換わっても、その妥当性が下がることはないはずである。脳と機能が似ている計算機モデルを構築し、両者の構造のどこが似ていてどこが似ていないのかを調べることで、対象とする脳機能を実現するための脳の構造について考察することができる。

ボトムアップ(構造の再現→機能の評価)な研究が多かったもう一つの理由として、技術的な限界により、脳の機能を再現するほどの計算機資源が昔は得られなかった、という理由もあるだろう。最近では計算のためのハードウェアやアルゴリズムが進歩し、範囲を限定すれば脳の機能をある程度再現する計算機モデルを作ることができるようになった³²⁾。脳の機能を完全に再現するようなモデルはまだ現れていないが、それを目指して世界中で研究が進められているところである。

6.1 ニューラルネットワークを用いて脳を創る

ニューラルネットワークは内部でどのような処理が起きているのかわからない、いわゆるブラックボックスであるため、脳の構造を創るためにニューラルネットワークは適さないという主張があるかもしれない。しかしながら、そんなことは全く無い。ニューラルネットワークの内部で起きていることを理解するための方法はたくさん提案されているし³³⁾、本稿で紹介したように神経科学や心理学の実験を模擬するような分析を行い内部状態を真の脳と比較することもできる。

逆に、ニューラルネットワークは動物の脳を模して作られた計算方法なので、脳のモデルに適している、という主張もあるかもしれない。これはこれで安直すぎるような気もする。元々は神経細胞の計算原理を模して作られたものかもしれないが、現在、実際に機械学習のために使われるニューラルネットワークは、動物の脳とはとても似ていない機構で動くことも多い。ニューラルネットワークによって高次認知機能などを含めて脳の機能を完全に再現できたとしても、それだけでは

脳ではなく人工知能でしかない。また、外界の物理状態を入力とした人工知能の出力に基づいて動物の身体運動を再現しただけでは、それはやはり脳ではなく動物型ロボットである。ニューラルネットワークを用いて脳を創ろうとするなら、脳の機能を再現するようなニューラルネットワークを作った後に、やはり、その構造を分析して脳としっかりと比較すべきだと思う。

7. おわりに

本稿ではまず、ケーキを創れるかどうかについて考えた。少し変な例だったかもしれないが、身近な物で例えることで、何かを創ることが(少なくとも概念的には)そこまで難しくないということが伝われば幸いである。次に、同じ議論を脳に適用し、脳を創るための二通りの方法について考えた。構造を再現して機能を評価するボトムアップな方法と、機能を再現して構造を評価するトップダウンな方法の具体例をそれぞれ紹介した。最後に、トップダウンな方法の妥当性や意義などを述べた。

既に述べたとおり、計算技術の発展に伴い、脳の機能を再現するための方法が充実してきている。また、脳の構造を調べた実験データも増える一方である。そのような状況で、機能を再現してから構造を調べる方法の有効性・有用性は、今後ますます高くなることだろう。機能を再現して構造を調べる方法と、構造を再現して機能を調べる方法の、両方をうまく組み合わせて駆使していけば、図3の右上に到達する日、つまり、脳を創れる日も、そう遠くないだろう。

謝辞

JSPS 科研費 20H05957 (学術変革領域研究 (A) 深奥質感) の助成を受けています。本稿で使用しているイラスト素材は、我々が保有するもの以外は、全てパブリックドメインです。

参考文献

- 1) Kriegeskorte, N., Douglas, P.K. (2018): Cognitive computational neuroscience, *Nat. Neurosci.*, Vol.21, pp.1148–1160; doi:10.1038/s41593-018-0210-5
- 2) Doi, T., Kuroda, S., Michikawa, T., Kawato, M. (2005): Inositol 1,4,5-trisphosphate-dependent Ca^{2+} threshold dynamics detect spike timing in cerebellar purkinje cells, *J. Neurosci.*, Vol.25, pp.950–961; doi:10.1523/JNEUROSCI.2727-04.2005
- 3) Koumura, T., Urakubo, H., Ohashi, K., Fujii, M., Kuroda, S. (2014): Stochasticity in Ca^{2+}

- increase in spines enables robust and sensitive information coding, *PLoS One*, Vol.9, p.e99040; doi:10.1371/journal.pone.0099040
- 4) Gillespie, D.T. (1976): A general method for numerically simulating the stochastic time evolution of coupled chemical reactions, *J. Comput. Phys.*, Vol.22, pp.403–434; doi:10.1016/0021-9991(76)90041-3
 - 5) Wang, S.S., Denk, W., Hausser, M. (2000): Coincidence detection in single dendritic spines mediated by calcium release, *Nat. Neurosci.*, Vol.3, pp.1266–1273; doi:10.1038/81792
 - 6) Jin, D.Z. (2009): Generating variable bird-song syllable sequences with branching chain networks in avian premotor nucleus HVC, *Phys. Rev. E*, Vol.80, p.051902; doi:10.1103/PhysRevE.80.051902
 - 7) Hodgkin, A.L., Huxley, A.F. (1952): A quantitative description of membrane current and its application to conduction and excitation in nerve, *J. Physiol.*, Vol.117, pp.500–544; doi:10.1113/JPHYSIOL.1952.SP004764
 - 8) Burkitt, A.N. (2006): A review of the integrate-and-fire neuron model: I. Homogeneous synaptic input, *Biol. Cybern.*, Vol.95, pp.1–19; doi:10.1007/S00422-006-0068-6
 - 9) 牧 勝弘, 赤木正人, 廣田 薫 (2009): 聴覚末梢系の機能モデルの提案: 聴神経の位相固定性及びスパイク生成機構のモデル化, *日本音響学会誌*, Vol.65, pp.239–250; doi:10.20697/jasj.65.5.239
 - 10) Verhulst, S., Altoè, A., Vasilkov, V. (2018): Computational modeling of the human auditory periphery: Auditory-nerve responses, evoked potentials and hearing loss, *Hear. Res.*, Vol.360, pp. 55–75; doi:10.1016/J.HEARES.2017.12.018
 - 11) Santoro, R., Moerel, M., De Martino, F., Goebel, R., Ugurbil, K., Yacoub, E., et al. (2014): Encoding of natural sounds at multiple spectral and temporal resolutions in the human auditory cortex, *PLOS Comput. Biol.*, Vol.10, p.e1003412; Available: <https://doi.org/10.1371/journal.pcbi.1003412>
 - 12) Bruce, I.C., Erfani, Y., Zilany, M.S.A. (2018): A phenomenological model of the synapse between the inner hair cell and auditory nerve: Implications of limited neurotransmitter release sites, *Hear. Res.*, Vol.360, pp.40–54; doi:10.1016/J.HEARES.2017.12.016
 - 13) Simoncelli, E.P., Heeger, D.J. (1998): A model of neuronal responses in visual area MT, *Vision Res.*, Vol.38, pp.743–761; doi:10.1016/S0042-6989(97)00183-1
 - 14) Moore, B.C.J., Glasberg, B.R. (1998): Suggested formulae for calculating auditory-filter bandwidths and excitation patterns, *J. Acoust. Soc. Am.*, Vol.74, p.750; doi:10.1121/1.389861
 - 15) Dau, T., Kollmeier, B., Kohlrausch, A. (1997): Modeling auditory processing of amplitude modulation. 1. Detection and masking with narrow-band carriers, *J. Acoust. Soc. Am.*, Vol.102, pp.2892–2905; doi:10.1121/1.420344
 - 16) Schädler, M.R., Meyer, B.T., Kollmeier, B. (2012): Spectro-temporal modulation subspace-spanning filter bank features for robust automatic speech recognition, *J. Acoust. Soc. Am.*, Vol.131, pp.4134–4151; doi:10.1121/1.3699200
 - 17) Olshausen, B.A., Field, D.J. (1996): Emergence of simple-cell receptive field properties by learning a sparse code for natural images, *Nature*, Vol.381, pp.607–609; doi:10.1038/381607a0
 - 18) Lewicki, M.S. (2002): Efficient coding of natural sounds, *Nat. Neurosci.*, Vol.5, pp.356–363; doi:10.1038/nm831
 - 19) Terashima, H., Okada, M. (2012): The topographic unsupervised learning of natural sounds in the auditory cortex, *Advances in Neural Information Processing Systems*, pp.2312–2320.
 - 20) Gaver, W.W. (1993): What in the world do we hear?: An ecological approach to auditory event perception, *Ecol. Psychol.*, Vol.5, pp.1–29; doi:10.1207/s15326969eco0501_1
 - 21) Saeed, A., Grangier, D., Zeghidour, N. (2020): Contrastive learning of general-purpose audio representations, *Institute of Electrical and Electronics Engineers (IEEE)*, pp.3875–3879; Available: <https://arxiv.org/abs/2010.10915v1>
 - 22) Jaiswal, A., Babu, A.R., Zadeh, M.Z., Banerjee, D., Makedon, F. (2020): A survey on contrastive self-supervised learning, *Technologies*, Vol.9, p.2; Available: <https://arxiv.org/abs/2011.00362v3>
 - 23) Kell, A.J.E., Yamins, D.L.K., Shook, E.N., Norman-Haignere, S.V., McDermott, J.H. (2018): A task-optimized neural network replicates human auditory behavior, predicts brain responses, and reveals a cortical processing hierarchy, *Neuron*, Vol.98, pp.630–644.e16; doi:10.1016/J.NEURON.2018.03.044
 - 24) Liu, X., Sawayama, M., Hayashi, R., Ozay, M., Okatani, T., Nishida, S. (2018): Perturbation tolerance of deep neural networks and humans in material recognition, *J. Vis.*, Vol.18,

- pp.756–756; doi:10.1167/18.10.756
- 25) Koumura, T., Terashima, H., Furukawa, S. (2019): Cascaded tuning to amplitude modulation for natural sound recognition, *J. Neurosci.*, Vol.39, pp.5517–5533; doi:10.1523/JNEUROSCI.2914-18.2019
 - 26) Yamins, D.L.K., Hong, H., Cadieu, C.F., Solomon, E.A., Seibert, D., DiCarlo, J.J. (2014): Performance-optimized hierarchical models predict neural responses in higher visual cortex, *Proc. Natl. Acad. Sci. USA*, Vol.111, pp.8619–8624; doi:10.1073/pnas.1403112111
 - 27) Raman, R., Hosoya, H. (2020): Convolutional neural networks explain tuning properties of anterior, but not middle, face-processing areas in macaque inferotemporal cortex, *Commun. Biol.*, Vol.3, pp.1–14; doi:10.1038/s42003-020-0945-x
 - 28) Francl, A., Mcdermott, J.H. (2020): Deep neural network models of sound localization reveal how perception is adapted to real-world environments, *bioRxiv*, 2020.07.21.214486; doi:10.1101/2020.07.21.214486
 - 29) 上村卓也, 寺島裕貴, 古川茂人 (2020): 自然音分類で訓練したニューラルネットワークにみられる心理物理学の変調伝達関数, 2020 年秋季研究発表会, 日本音響学会.
 - 30) Hoshen, Y., Weiss, R.J., Wilson, K.W. (2015): Speech acoustic modeling from raw multi-channel waveforms, *ICASSP, IEEE International Conference on Acoustics, Speech and Signal Processing*, pp. 4624–4628; doi:10.1109/ICASSP.2015.7178847
 - 31) 寺島裕貴, 古川茂人 (2018): 聴神経の計算論再考: 教師なし学習か課題最適化か?, *The 32nd Annual Conference of the Japanese Society for Artificial Intelligence*.
 - 32) Bommasani, R., Hudson, D.A., Adeli, E., Altman, R., Arora, S., von Arx, S., et al. (2021): On the opportunities and risks of foundation models, Available: <https://arxiv.org/abs/2108.07258v2>
 - 33) Montavon, G., Samek, W., Müller, K.-R. (2018): Methods for interpreting and understanding deep neural networks, *Digit. Signal Process.*, Vol.73, pp.1–15; doi:10.1016/J.DSP.2017.10.011