

コンピューターは心を持つのだろうか？

田森 佳秀

■研究のねらい

「コンピューターは心を持つのだろうか？」という、使い古された問いかけから始めた研究ですが、これが科学の対象として本気で取り扱おうとすると、ちっとも古くないどころか、全く始まってすらいらないとも言える問題なのです。まず、こういった問題を考えるときに、抽象的な言葉の定義をどうするか、ここではすなわち「心」の定義をどうするかということが最初の問題となります。そして、「心」の定義をどうするかという議論を始めたら、それだけでどんどん哲学の世界に入り込んで戻れなくなってしまいます。

「心」の定義をはっきりさせること、それはすなわち、「心」と「心でないもの(物質あるいは身体)」との間に境界線を引くことですから、「心」の定義をはっきりしてしまえば、表題の問いかけに対する答¹は自明なこととなります。人間の場合は、この境界線は人間(脳)の中にあり、恐らく現在のコンピューターの場合は、外にあることは間違い無いでしょう。従いまして、本研究の目的は、このような定義が終った後の簡単な問題に答えるところにあるのではなく、この問いかけを借りて、むしろ心身問題(心脳問題)²と言われるデカルト以来の大問題に対して科学の言葉で答(あるいは、できるかぎり答に近い科学的結果)を与えようというところにあります。以下では、心身問題に対する科学の厳しい基準³を満足するような回答に、脳の物理学がどれだけせまれるかを見てゆくことにします。

1. 「機械」を特徴づけるもの

コンピューターのキーボードで、「A」というキーを押したら、画面に「A」と表示される状況を考えてみます。この「A」という文字が表示されるまでには、様々な「A」の表現が存在していることが分かります。まずキートップに書かれた「A」という文字、キーボードからコンピューターに向かうケーブルの中を走る電気信号、コンピューターのメモリの中の「A」を意味する状態、ディスプレイ上のビットマップ等です。このような「対応づけ」のことは、「コーディング(符号化、或は符号化されたもの)」と呼ばれます。ここでは、「表現」、「モデル」等も、コーディングとほぼ同義語と考えても構いません。このような対応づけはコンピューターの設計者が仕組んだものであって、当然、コンピューターがこの対応づけを決めたわけではありません。「心」の「心」らしさは、対応づけの意味を理解しているところであって、機械の機械らしさは、コーディングしかないところに

¹ つまり、その境界線が、コンピューターの中にあれば、コンピューターは心を持つということ

² 心と体(或は脳)との境界がどこにあるのかをはっきりさせようという哲学上の問題。

³ 哲学では、言葉の定義を与えることに労力の多くを割きますが、理論科学においては、(1)定義を(可能なら)単純な仮定から導き出すこと；(2)たとえ定義の解釈が困難でも導かれた結果が有用であるかどうか(予言が可能かどうか)；が問われるという非常に厳しい基準があります。例えば、物理学において「物質」の定義は、「質点」から「波動関数」へと進歩したわけですが、これらの定義そのものよりも、これらの定義を用いた運動方程式(シュレディンガー方程式)から、どのような予言が導かれるかが重要であったわけです。「波動関数」においては、導き出された定義を解釈することがむしろ後の問題となっていたりするわけです。

あるとも言えましょう。この定義のもとでは、「A」の意味を感じている心が、このコンピューターの中に無いことだけは分かります。

2. 「緑」のクオリア

意味とか理解という、「A」という文字一つをとっても、様々な因子を考慮せねばならず、なかなか問題がはっきりしてきません。特に、意味や理解は外界の情報との対応づけを持たずとも存在するので、「感覚情報と独立して存在する」という事を「心」の条件としたくなってきてしまいます。しかし、実は感覚入力が必要な「心」の状態もあります。それば例えば「緑」を見たときの感じがそれです。これを専門用語で「クオリア」と呼びます。例えば、「緑の緑らしさ(緑の感じそのもの)」は「緑のクオリア」と呼ばれます。クオリアは、確実に感覚入力と対応づいていて、しかも「心」無くしては存在できないものです。

3. 「心」を持つ脳のモデルの条件

ここまでで、コーディングそのものには「心」は存在せず、逆に「心」はコーディングを取り決めている場所にあると「定義」してきました。すでに、「定義する」という罪を犯していますので、脳のモデルの中で、個々の「心」の状態であるクオリアのコーディングの定義を、モデル構築者である私が行なうことだけは避けねばなりません。そのためには、コーディング(クオリアに対応する脳内記述)を見ただけで、どのようなクオリアを「心」が感じているか答えられねばなりません。クオリアを言い当てられないまでも、コーディング(ニューロンの発火パターン)だけから、せめて、どのような刺激が入力されているかぐらいは言い当てられなければなりません。私がコーディングを定義していいのなら簡単なことですが、モデル構築者に許されているのは、せいぜいコーディングが満たすべき条件を見出すことぐらいなので、大変なことなのです。

目の中の網膜には、緑錐体と呼ばれる「緑」に最もよく反応する細胞が存在します(同様に、赤錐体や青錐体がある)。確かに、網膜の緑錐体を同定し、これが発火しているかどうかを調べることによって、「緑」が入力されていることを予測できそうです。しかしながら、これは赤青緑の三色のうちどれか一色が入力されたときに限る話で、「黄色」に対しては、緑と赤の錐体が両方とも発火します。「黄色」の場合は組み合わせでコードしているとすればよいというわけにはいきません。何故なら、赤錐体が発火しているにもかかわらず、「黄色」をいくら目を凝らしてみても、その中に「赤」は見えないからです。この為、色をコードしている細胞が大腦皮質視覚野の中に存在するという人もいます。しかし、これも随分変な話です。何故なら、視覚野から、「赤」に反応していたニューロンと、「緑」に反応していたニューロンの両方を取り出してきて顕微鏡で見たときに、どちらのニューロンだったのかを言うことができないからです。⁴

⁴ 少なくともこれができる神経解剖学者は存在しません。

従って、網膜神経節細胞を含みうる⁵視覚野の \underline{Y} underline{ニューロン集団の活動パターンの時間的変化} (活動の時空間パターンと呼びます)が、色をコードしていると考えることが自然でしょう。しかし、ここでも問題となるのは、この活動の時空間パターンが、どのような条件を満たしているのかということです。また、特定の時空間パターンが「緑」の入力に対応していることを、モデル構築者が定義できないのは言うまでもありません。

■研究の成果

本研究の成果は、現在のところ、

「神経系の活動の時空間パターンが、ある条件を満たす時に特定のクオリアをコードしている」

ということにすると、色々なことがともうまくいくということを示せるということです。その条件とは、ニューロンの活動の時空の座標系が満たすべき条件です。まず、ニューロンの時空を、そのニューロンが経験したスパイク (神経発火) の数で測ることにして、茂木が提唱している相互作用の同時性原理に基き、複数のニューロン活動によって張られる時空の中での距離を書いておきます。この距離を繋げていった時にできあがる空間は、多次元複素空間の中での回転や反転に対して不変な軌道 (射影空間、あるいは部分多様体) を持ち得ます。この軌道が、それぞれ個々のクオリアと対応しているのだと予想したのです。

1. クオリアに対応する軌道

クオリアに対応する軌道を、私が定義するわけにはいきませんから、緑を見た時の網膜神経節細胞群の発火パターンを、よく知られている網膜内回路網から予測します。すると、最低で3個のニューロンのON-OFFの時間変化が生じることがわかります (図1)。⁶

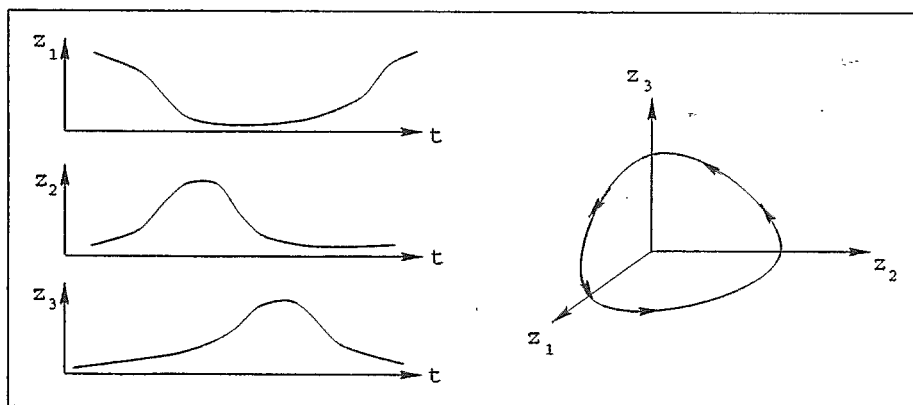


図1 網膜神経回路網から予想された網膜神経節細胞のON-OFFパターン

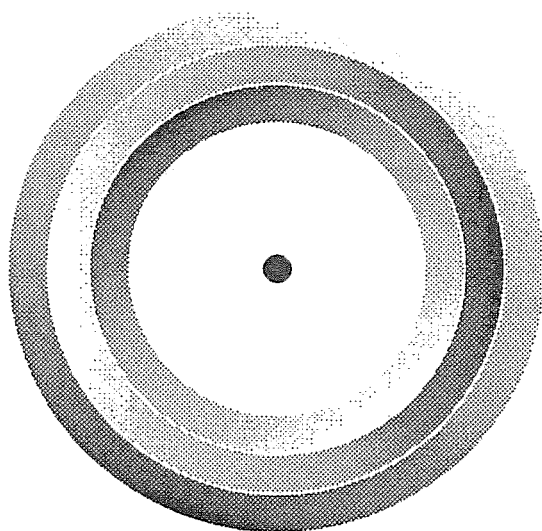
⁵ 大脳皮質ニューロンと区別すべき積極的な理由が無いので「含みうる」とします。

⁶ 但し、この予想は非常に大雑把なものなので、妥当性は座標変換の普遍性を確認することで行なわなければならない。

これが、本当に緑のクオリアに固有のものであるためには、この活動パターンを座標変換しても色が変わってはいけません。また、このニューロンの間の実空間における距離は何の意味も持ちませんから、3つのニューロンがそれぞれ網膜上で離れたところにあっても緑が見えなくてはなりません。更に ON-OFF の時間間隔が同じ比率で間延びしても、同じ色でなくてはなりません。本当にこのようなことが起るのでしょうか？

2. ベンハムのコマ

図2が上記の ON-OFF のパターンを網膜上に作るために描いたコマの模様です。この



コマを回転させると、白と黒しか無いのに、緑色が見えます。これは100年以上も前から Benham のコマとして知られるのと同じ現象です(但し、この模様の方が古典的な Benham のコマよりもきれいな緑になります。) この模様のコマは、大きさを拡大しても、回転数を速くしても遅くしても(時間的座標変換に相当)、同じ色が見えます。このコマと同じパターンで網膜や大脳皮質を直接刺激できたら、もっときれいな緑が見えることでしょう。

図2 導かれた緑に見えるベンハムコマ

■まとめと展望

この理論に基づくならば、「緑」を感じることをできない(緑錐体が機能していない)人にも、緑色を見せることが可能なはず⁷。

本研究の結果によると、網膜神経節細胞が関わって色がコードされていることとなります。このことは、視覚野で色がコードされているとする、現代神経科学の常識とは矛盾しないのでしょうか？ 私は、決して矛盾しないと考えています。むしろ、この結果は、色のコーディングにも様々な段階があることを教えてくれています。我々は、目をつぶって「緑」を想像することは可能ですが、クオリアのレベルでは、目をつぶった時の「緑」は、本当に「緑」を見ている時のものとは異なっています(実際にやって比較してみましょう)。この意味で、緑のクオリアは、感覚情報と対応づいていて、別々に考えることはできないと言えます。視覚野でコードされた色の情報は、目をつぶった場合の色とも相関のあるも

⁷ Benham ディスクでは、未だ決定的な結果は出ていない

のなのでしょう。

本研究によって、ニューロンの時空間発火パターンだけの情報から、どのような入力によるものかを予想することのできる、理論が始めて生まれました。個々の入力とそのコーディングとの対応関係を、モデル構築者が定義せずに構成されているところが、これまでの理論と大きく異なる点です。この理論の解釈を認めるならば、外界の入力に対応するコーディングは、脳のモデルを作ろうとする限り、自然が一意に決めているのであり、モデル構築者が自由に定義することは許されないということになります。すなわち、自然は「表すもの」と「表されるもの」との自由な対応関係を許していないということを主張しているのです。

最後に、表題の「コンピューターは心を持つのだろうか？」という問いかけをもう一度考えてみましょう。まず、「コンピューターがコーディングを生成しないうちは、心を持つことは無い」ということは最初から分かっていました。しかし、本研究では、今までとは異なり、コンピューターにコーディングを生成させる時に、そのコーディングが満たすべき条件が得られています。すなわち、多次元複素空間の対称な座標変換に対して不変な軌道を作れるようなコンピューター(このコンピューターは必然的にアナログです)が出現すれば、それが「心」を持つこともある、というのが私の答です⁸。

対称な軌道というのは、純粹に数学的な存在ですから、それが脳の中でコーディングとして出現する前から実在しています。言い換えれば、脳は、数学的実在(対称な軌道)である「心」の状態をサーチする機械であると言うことができます。

■発表論文リスト

- ・ 田森佳秀 (1995) 「脳の情報処理はいかにして機械の情報処理と区別されるのか」、生体の科学 46 : 60-66
- ・ Kuroiwa J., Miyake S., Inawashiro S., Aso H., Tamori Y. (1996), Self-organization in a formal model and self-consistent Monte-Carlo simulation. Proc. of WCNNM' 96:543-546
- ・ Inawashiro S., Tamori Y., Miyake S., Kuroiwa J. (1996) Hamiltonian formalism for self-organization of formal neurons J. of Phys. A. :7389 - 7399
- ・ 田森佳秀 (1997) 「樹状突起の形態を考慮したニューロンモデル」 —平成7年度日本神経回路学会奨励賞の研究概要—、日本神経回路学会 4 : 94-95
- ・ Mogi, K. and Tamori, Y. (1997) Creatiy and the neural basis of qualia : Proceedings of the Mind II conference in press

⁸ 仮に、コーディングを生成できる機械が作れたとしても、クオリアが何なのか依然として分からないことに注意(コーディング生成機械の中に、クオリアとコーディングの間に引く境界線があることは分かるが、その境界線の中に残るものが物理的に何なのか分からない。