

## 題目：社会知性の脳計算モデル —実験と数理の融合脳科学—

中原裕之（なかはら ひろゆき）

理化学研究所 脳科学総合研究センター 理論統合脳科学研究チーム〔〒351-0198 埼玉県和光市広沢 2-1〕

### はじめに

私たちの日常生活の大半は何らかの社会行動である。つまり、「ヒトである」ことの多くはその社会性にあるといえる。人々が互いの心やふるまいを推断するときに働かせる社会知性は、脳の産物である。より精緻に述べれば、それは脳(神経回路)の情報処理すなわち脳計算の産物だ。ここで、科学者としては次のような疑問が生まれる。社会知性の脳機能解明、あるいは心の生物学はあり得るのだろうか。脳の生物学的実体に即しながら社会知性を理解することは可能なのだろうか。その脳計算の理解はどうか。筆者はこういったことが可能だと信じ、社会的意思決定の研究を通じてその探究を進めている。駆け足にはなるが、その構想をここに述べてみたい。

### 社会知性の脳機能：「自己システム」＋「他者システム」で大胆に切り込む

どうすれば社会知性の脳機能の本質に迫れるのか。筆者は、社会知性の脳計算の基本構成要素を解明することが最も重要だと考える。社会行動や社会知性というと、その複雑さと奥の深さにめまいを覚えるかもしれない。確かに複雑で多様だ。それらをきちんと記述しようとする、脳という実体との関係づけが難しすぎると感じられる向きもあるだろう。しかし、いまだに摩擦力の数理的記述には不明な点があっても、ニュートンの力学がもたらす恩恵は計り知れない。また、いまやエピジェネティクスが盛んでラマルクも一部は正しいのかもしれないが、やはりワトソンとクリックの DNA 構造の発見が学問を大きく進展させたのである。科学の歴史は、複雑すぎると思われたことを本質的な問いに際限まで昇華し、それに答えることの積み重ねだったのではないだろうか。一見複雑な社会行動と社会知性についても同様のことが言える。その行動と知性（すなわち脳機能あるいは脳活動）の対応をつけるときに、カギとなるのが脳計算である。この行動・脳活動・脳計算の三位一体を読み解くには、実験と理論の融合研究が重要だ。以下でみるように、行動を十分に説明し、その脳活動との対応を明確にすることが、脳計算と言える。つまり、三位一体の実現とは脳計算の解明であり、社会知性の脳計算の基本構成要素を明快、かつ定量的、数理的に明らかにすればよい。

では、脳の社会知性の本質とは何か。社会とは、究極的には（直接的であれ、間接的であれ）他者の存在を意味する。よって、“脳の中に他者がいる”ことがその本質であろう。これを理解するには、他者がいない状況、つまり非社会的状況と対比して、社会的状況を考えてみればよい。たとえば、行動を選択する意思決定を考えてみよう。非社会的状況であれば他者のことを考えないで意思決定をする脳内プロセスが起きるだろうが、社会的意思決定ではその脳内プロセスの中に他者への考慮が入ってくる。つまり、脳内過程に他者が存在するのである。

この脳内他者の存在を実現する脳計算の基本構成要素を徹底的に明らかにするのが、社会知性の脳機能解明の根幹になる。意思決定で考えると、自己の意思決定を行う「自己システム」と他者の意思決定を自己の脳内でシミュレーションする「他者システム」があることになる（図1）。つまり、社会的意思決定は、「自己システム」+「他者システム」の統合と捉えることができる。「意思決定」を「社会的」という言葉で修飾する、つまり「社会的意思決定」とするためには、そもそも「社会的」という言葉で修飾する前の「意思決定」、つまり「非社会的意思決定」の脳機能についての確固とした理解が必要なことは自明だろう。この（非社会的）意思決定の脳計算は「自己システム」の土台になる。社会知性の脳計算の解明の第一歩を進めるためにも、これには価値意思決定の脳計算がうってつけだと筆者は考えている[1, 2]。この約20年の間、報酬や利得などの「価値」に基づく意思決定の脳情報処理の解明には大きな進展があり、脳計算モデルが最も活躍している分野である。ちなみに、この価値意思決定の数理的な基礎は、マルコフ過程における逐次選択のオンライン学習理論（強化学習）であり、強化学習は動的計画法や機械学習、あるいは経済学の効用理論と関連が深い。この価値意思決定の脳計算を「自己システム」の土台として、「自己システム」+「他者システム」として社会的意思決定の脳計算の解明を進めるのである。

以上で構想を簡単にまとめたが、この考え方に基づいて行った直近の研究の成果を紹介したい[3]。これはヒト非侵襲脳活動イメージング実験（磁気共鳴機能画像法、**functional magnetic resonance imaging**；以下、ヒト fMRI）を用いた研究である。社会認知や精神疾患の分野では、多くの議論の基礎に、いわゆる「心のシミュレーション理論」という考え方が根強くある。他者の心を理解あるいは予測するには、他者の心的過程（つまり、他者の脳で起きている脳計算）を、同様の脳計算をする自己の脳部位を使ってシミュレーションすればよいという考え方である。この研究のポイントは、「心の理論」に取り組み、心のシミュレーションに本気で挑んだことにある（図2）。その脳計算の基本構成要素を重点的に抜き出すことを可能にする実験パラダイムを開発し、脳数理モデルを構築し、それをヒトの実際の行動とそのときの脳活動に照らして定量的検証を行った。さらに、基本構成要素の徹底的解明のために、心のシミュレーションを可能にする学習（シミュレーション学習）の脳計算に焦点を当てた。実験パラダイムは、コントロール課題では自己の報酬選択課題を行わせ、メイン課題は他者行動予測課題で、他者が被験者と同じ報酬選択課題をしているときに、被験者にその他者の選択を予測させるということを行った。価値意思決定の知

見に基づき、この他者の行動予測をする脳数理モデルを構築し、概念的に議論されることが多い「心のシミュレーション」を定量的かつ数理的に検証した。その結果の詳細は紙面の都合上ここでは省くが、端的に述べれば、実際の行動に一致する脳活動の実体（脳計算）と対応する、新たな心のシミュレーション学習のモデルを提示したと考えている。興味を持たれた読者は、概説や原著論文を参照されたい[3, 4]。

上記の研究でも用いたモデル化解析手法に、ここで少しだけ触れておきたい（図 3）。この手法を用いることで、ヒト fMRI 実験と脳数理モデルの融合研究が可能になる。脳計算の主要変数の定量的評価を、行動データと脳活動データの両方を参照して行うのである。したがって、行動を十分に説明することを確認しつつ、仮説（脳数理モデル）の情報処理が脳活動として実現されていることを同時に検証する。また、この手法に依拠した実験課題の作成自体が、融合研究に不可欠な“より数理的に透徹し、より実際データに対応した理論構築”と“より理論の明確化につながる実験”を意識して行われるので、理論と実験のあいだで相互へのフィードバックが形成される。なお、脳数理モデルの検証に耐える実験課題の作成には、理論と実験の両方にまたがる優れたセンスが必要になるが、それについてはまた別の機会に触れたいと思う。

## さいごに

現代の脳科学では、視覚や運動あるいは認知機能について、徐々にではあるが、その神経回路や脳計算についての理解が進みつつある。振り返れば、初期視覚野の方向選択性神経細胞活動を発見し、のちにノーベル賞を受賞したヒューベルとウィーゼル(Hubel, D. H. & Wiesel, T. N.)の発見は、視覚の脳計算の基本構成要素についての発見であった。その発見は、現在の空港などでの自動顔識別システムなどにつながっている。社会知性の脳機能を視覚・運動・認知機能と同様に理解するには、その脳計算の基本構成要素をまずは徹底的に解明しなければならない。我々はそれを目ざしており、この挑戦しがたいある課題にとともに取り組む仲間が増えることを歓迎したい。新たな人間総合科学の幕開けを期待したい。脳の社会知性あるいは心の生物学の新しい教科書が書かれる日もそれほど遠くないのではなかろうか。

## 引用文献

1. 中原裕之: 生体の科学 **64**(4): 323-328, 2013
2. Nakahara, H.: Current Opinion in Neurobiology **25**: 123-129, 2014
3. Suzuki, S., N. Harasawa, K. Ueno, et al.: Neuron **74**: 1125-1137, 2012
4. 中原裕之 and 鈴木真介: BRAIN and NERVE **65**(8): 973-982, 2013

Figures

Fig 1

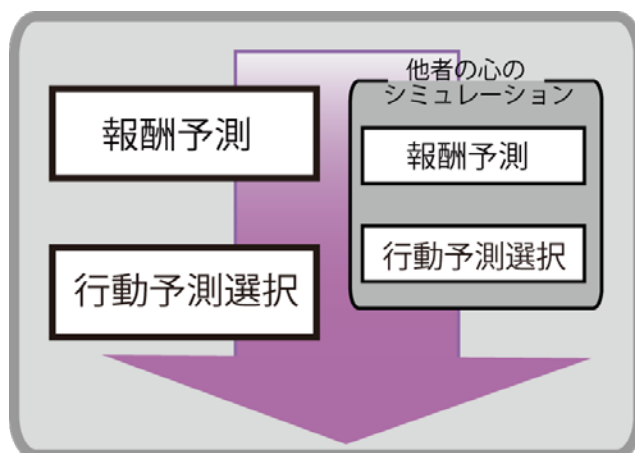


図 1 社会的意思決定の脳数理

「自己システム」 + 「(脳内で行う他者の心のシミュレーションである) 他者システム」の統合として、社会的意思決定の脳数理モデルを考えることができる (図 2 も参照のこと)。

Fig 2

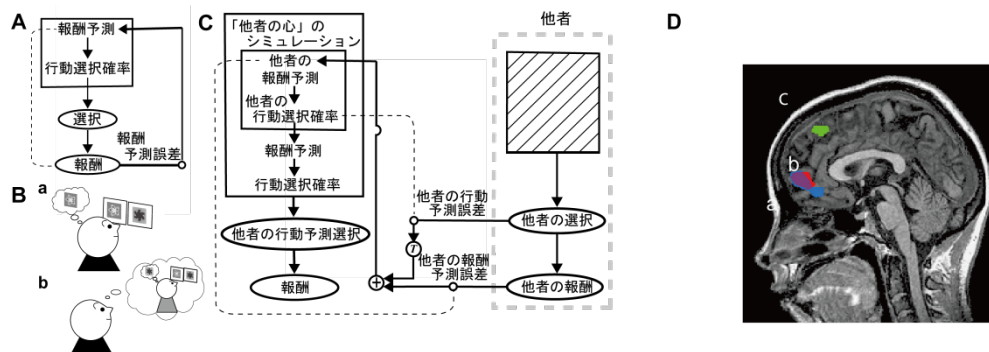


図2 本構想の背景となった成果 (Suzuki et al., 2012)

A. 価値意思決定の脳内過程: 入力からの直接連想による報酬予測 (価値) に基づいて確率的に行動選択する。その結果得られた報酬から算出した報酬予測誤差 (実際の報酬量と予測報酬量の差) により報酬予測を学習 (更新)。四角の枠は個体の内部プロセスを表す。B. (a) 価値意思決定の実験課題。(b) 他者の価値意思決定予測課題。C. B の (b) の課題で他者の意思決定をシミュレーション学習する脳内過程: 2つの学習信号である「他者の報酬予測誤差」(他者が実際に得た報酬量とシミュレーションによる他者の予測報酬量の差) と「他者の行動予測誤差」(他者の実際の行動と予測された行動の差) の存在と脳機能局在を示した。D. 主な脳活動: 価値意思決定課題を遂行中の脳活動部位 (a) と他者予測課題を遂行中の脳活動部位 (b, c)。a: 被験者自身の「報酬予測誤差」(前頭葉腹内側部: 画像の青色部分)。b: 「他者報酬予測誤差」(前頭葉腹内側部: 画像の赤色部分)。c: 「他者行動予測誤差」(前頭葉背内側部: 画像の緑色部分)。

Fig 3

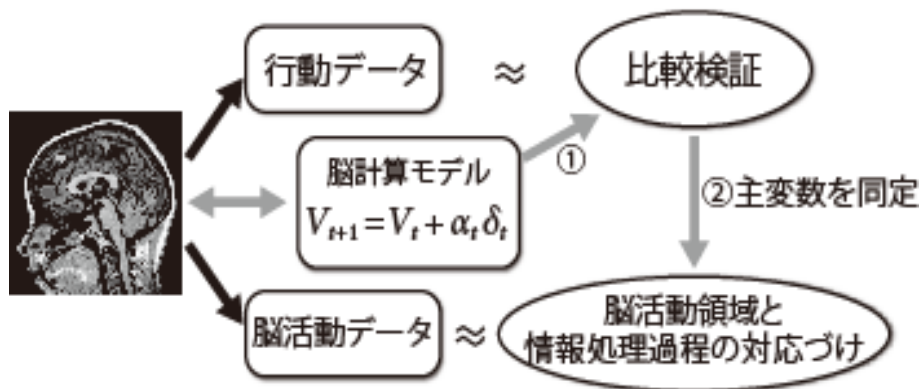


図3 モデル化解析手法

一つの仮説は一つの脳数理モデルとして実現される。そのモデルのふるまいを「行動」と比較検証する (①) ことで、行動の十分な説明能力の可否を評価し、行動生成を担う脳情報処理の主変数を同定する (②)。この主変数を脳活動データ解析に適用することで、その情報処理がどの脳活動で表現されるかを検証する。これにより、行動、脳活動そして脳計算の3つの対応づけが可能になる。