

脳は愛を計算せずには いられない

中原裕之

なかはら ひろゆき

独立行政法人理化学研究所・脳科学総合研究センター・理論統合脳科学研究チーム

恋人たちが愛を囁きあうとき、何が起きているのか？ なぜ私たちは愛する人を喜ばせたいのか？ 幼いわが子を抱きしめるときに心に満ち溢れる愛情は何なのか？ 愛は、私たちの人生に不可欠である。そして、愛も脳の情報処理である。脳は愛を計算せずにはいられない。愛を計算する脳の物質原理と情報原理を理解することで、私たちの愛への理解が深まり、いずれ新たな人間観・社会観へとつながるだろう。

愛は語りつくせない？

私は、「愛の研究者」ではないし、以下は「愛の科学的事実の報告」ではない。意思決定や動機づけの脳機能、その情報原理／数理原理の解明を目指して日夜悪戦苦闘する中で育まれた私論、つぶやきほどにしかすぎない。それでも、筆者としては、ごくささやかでも読者の読む楽しみにつながることを望まずにはいられない。

私たちは、「愛」を知っている、そして知らない。遠い昔からあまたの形態——神話や叙事詩、小説や詩、歌劇や映画——で様々な愛、そして恋

について語られてきた。日々の生活の中で、親と過ごした時間の中で、青春の真っ只中、あるいはその甘い思い出、子や孫との時間に、私たちは何らかの愛を経験している。それらは、しばしば個人的なものである。愛のあり方は、個々人において多様でもある。また、私たちは、人類愛、親の愛、恋愛など、様々な愛を語る。愛情は同じ種のヒトにだけでなく、ペットや他の生き物に向けられることもある。愛に関する豊かな言語表現は、愛に対する私たちの深い関心を、そして、その捕まえきれぬ豊穡さを表している¹。この豊穡さの一部は、様々な隠喩に支えられている²。たとえば、「愛は泉」「愛は旅」などの物理的・具体的経験の隠喩に、または、「愛は狂気」「愛は魔法」「愛は戦争」などのより抽象的な概念構造を隠喩でつなぐことで、愛は語られる。

しかし、ひとたび、愛について“科学的”に（「科学的」が何を意味するかはとりあえずさしおいて）考えようとするとき、その豊かさが、私たちを迷宮に誘いこむように思える。それは、私たちの心のどこか奥底に潜む願望——愛は語りつくせぬものであってほしい——なのかもしれない。この言語表現の豊かさ自体が科学的に考えることを遠まわしに拒否するかに見える。これらは、愛に限ったことではない。たとえば、多くの人の関心をひく“意識”や“感情”についても同様のことが言える。なぜワタシはそんなに関心をもつのか、と問いかけてみると、一つの大きな理由は、言語表現の豊かさにあることに気づく。この言語表現の豊かさと、その運用に深く分け入っていく試みは連綿と行わ

Love as neural computation: From a perspective of value-based decision making

Hiroyuki NAKAHARA

Email: hn@brain.riken.jp, HP: <http://www.itn.brain.riken.jp>. 理論脳科学が専門。脳の情報処理のメカニズムとその原理の探求。特に大脳基底核関連回路を研究。意思決定や情動的／社会的行動に関心。『脳研究の最前線(下)』(講談社ブルーバックス、第11章)、『脳の計算論』(第5章、東京大学出版会、シリーズ脳科学第1巻)などに執筆。

れている(だろう)が、私はそれがどこに行き着くのかは知らない。ただ、たとえば、一つの見方は、かつてヴィトゲンシュタインが指摘した「言語ゲーム」としての世界だろう³。言語表現の豊かさにもとづく私たちの関心は、もしかすると、豊かな表現が生き生きと使われる個々の場面の中だけでのみ、十分に理解されるものなのかもしれない^{4,5}。それは、社会的な使用、社会的な理解にもとづく構成の中でのみ理解できるのかもしれない。その意味では、私たちの日常の言語表現の中の“愛”は、語りつくせぬものであって当然とも言える。

しかし、たじろぐことはない。語りつくせぬものなどいくらでもある。語りつくせぬとも、語れることを語ることで、語れることを明確にすることで、私たちの理解は大いに進む。ニュートン力学は、厳密な意味では「摩擦」を捉えきれていないが、たしかに、それは「摩擦」を含めて、私たちの力学的理解を大いに深めた。「実数」の厳密な定義を知らずとも、その操作性を把握するだけで、私たちの多くは多大なる恩恵を受けている。

愛と脳の物質原理

そもそも、愛に対する私たちの関心は言語表現だけに起因するわけではない。愛するという行動、その経験にも起因する。愛だけでなく、あらゆる感情、われわれの思考、意識、そして心は、脳と身体という物質から生まれる。したがって、物質原理に支配されているし、脳の身体性は確実に存在する。

たとえば、愛は、抽象的な(心)の概念でもあるが、重層的な身体的反応としても現れる。心拍数の増加、発汗作用、あるいは性的反応かもしれない。これらの身体反応は愛そのものではないにしろ、愛には身体反応として計測可能な部分がある。脳と身体は、自律神経系や内分泌系を通じて、互いに作用しあう一つのシステムを成している。

物質としての脳は、精緻な神経回路網を成して、遺伝子・分子による巧妙な制御のもと、様々な化

学物質(たとえば神経栄養因子・神経伝達物質・神経修飾物質など)が作用して機能する。脳の物質原理は愛にも作用する。一例を挙げれば、神経細胞が産出する神経ペプチド(オキソトシンやバソプレッシン)は、ハタネズミ属において、単婚あるいは乱婚のどちらを好むかに多大なる影響を与えることが示されている⁶。また、性差や性行動に関わる脳回路や遺伝子制御の違いも研究が進みつつある。ショウジョウバエで見つかった、性志向性に直接関与する *fru* 遺伝子などはその典型例であろう^{7,8}。

愛は脳の情報処理

脳では、物質原理と情報原理が交錯する。端的に言えば、物質原理に支えられた脳における多数の神経細胞集団の活動が、脳機能を生み出している。この神経細胞集団活動のパターンと相互作用こそが、知覚・認知・運動であり、愛・感情・思考・意識・心なのである。

脳は“考える”だけではない。日常用語でいう“考える”とは、意識にのぼっていることを指す。脳で行われている「計算」は、その“考える”ことはむろんのこと、それには直接立ち現れない計算も含む。「ワタシが考えている」こと、つまり意識にあることよりも、「ワタシの脳の計算」、つまり意識・無意識の両方を含めた、脳で行われている情報処理のほうが膨大なのである。その意味では、無意識の計算こそが、脳の計算としては主流である。

そして、愛も確かに脳の計算である。これは、誰かに愛を語るときに、その語る効果を意識して、「ワタシが計算高く愛について考えている」という意味ではない。「愛している」と囁くヒトの脳では、本人が意識しない、無意識の愛の計算が行われている。

愛をはじめとする脳の豊穡な諸機能の実現には、膨大な情報処理が必要である。脳を理解するとは、物質原理と情報原理の交錯の中で、豊穡かつ複雑な諸機能を解きほぐし、神経細胞集団活動のパターンと相互作用によって諸機能がいかにか実現され

るのかを理解すること、つまりは、脳の計算を理解することに他ならない。脳の情報処理の解明に研究の力点をおく脳科学(神経科学)のアプローチは、理論脳科学(計算論的脳科学、数理脳科学とも呼ばれる)として発展してきた⁹⁻¹²。情報幾何などを含む広義の情報統計科学や数学/物理学に支えられ、脳の情報原理と数理原理を十全に記述しうる新たな数学体系の構築(数理脳科学)と、各脳回路の具体的な計算理論とその表現とアルゴリズムの解明(計算論的神経科学)が、進められつつある。

さて、愛には感情的側面(feeling)と情動的側面(emotion)がある。意識・無意識の区別とも関連するが、「愛」を“感じる”という自覚(感情認知)と、その認知にかかわらず、愛の情動的側面が行動に影響をあたえることは、関連はあっても、同じものではない。相対的に言えば、愛の認知はより意識に近く、情動的側面は必ずしも近くない。愛の情報原理を考えるには、まずは、この無意識に近く、行動に影響を与える脳の計算に目を向けるのがわかりやすい。

愛は(報酬)たりえるか

「報酬」は、主体にとって獲得が望ましい「対象」をさすために使われる。たとえば、食べ物や飲み物といった原初的報酬(primary reward)の獲得は、生物の生存と直接的に結びついている。では、愛は報酬たりうるのだろうか？

日常ではそれほど意識されていないが、実は、報酬というコトバには、少なくとも二つの意味——「対象」と「強化」——がある¹³。主体の行動という視点から見れば、「報酬」は、主体がその獲得のための行動を選ぶ、つまり主体の行動選択をある方向に影響を与える。このように、ある行動を選択的に学習することを「強化」という。

愛は、獲得が望ましい対象である。「愛するがゆえの行い」、愛は行動選択に現れる。すなわち、愛は強化でもある。愛はたしかに報酬の二つの意味を満たしている。

生物の本能的欲求に関わる原初的報酬ではない

が、その原初的報酬との近接性あるいはその派生物として報酬(「対象」かつ「強化」)として機能するとき、二次的報酬(secondary reward)と呼ばれるものがある。金銭はその代表例である。愛は、生得的あるいは獲得的なのか、または、愛は利己的それとも利他的か、これらの問いは、原初的あるいは二次的報酬なのかと問うことに通底する。

価値意思決定をする脳

この20年ほどの間に、報酬獲得——もっと広くは価値判断にもとづく意思決定と行動選択(以下、価値意思決定)——に関する脳機能の理解は大きく進展してきた。この脳機能において、大きな役割を果たすと考えられているのが大脳基底核回路である。大脳基底核回路は大脳皮質に包まれるようにして存在する。ほかにも、大脳皮質の領野の一部や別の皮質下回路も深く関連しており、これらをまとめて大脳基底核関連回路とよぶことにする。

大脳基底核は、大脳皮質の広い領野から投射(信号の入力)を受け一方、脳幹運動中枢に出力を送ると同時に、視床を介して大脳皮質にも投射を戻している。大脳基底核の疾患としてパーキンソン病やハンチントン病があるが、これらの疾患では、それぞれの病気に特徴的な運動障害を示す。このことから、大脳基底核が運動制御や行動選択に重要な役割を果たすことがわかる。

さらに、大脳基底核回路にはドーパミン神経細胞が豊富に存在する。ドーパミン神経細胞とは、神経伝達物質としてドーパミンを蓄えている細胞のことであり、報酬に関連して活動する脳の回路(報酬系)の中心的な存在である。この回路が報酬と関連することは、脳内自己刺激(ICSS: intracranial self-stimulation)の実験から鮮明に示されている¹⁴。この実験では、動物が自らある行動をとる(たとえばバーを押す)ことにより、脳内に埋め込まれた電極から、特定の部位を刺激できるようになっている。そして、ドーパミン神経細胞群を刺激する位置に電極が埋め込まれたとき、動物のその行動は著しく強化されるのである。また、ドーパミン神経細

胞群が存在する脳部位には、たとえば自律神経系、内分泌、本能行動で重要な役割を果たす視床下部や、あるいは情動表出や表情認知で重要な役割を果たす扁桃体などからの投射がある。

これらの知見は、大脳基底核関連回路が、ドーパミン神経細胞群から得られる報酬関連情報を、行動選択に反映させうることを示唆する。

愛とドーパミン神経細胞

愛を行動に影響を与えるものとしてみたとき、報酬そして価値意思決定との類似性は高い。実は、脳機能画像法(fMRI)を使った研究により、恋愛中の人が愛する人の写真を見たときに、大脳基底核の一部(線条体)や、ドーパミン神経細胞群が豊富な部位が、特に活性化することが示されている¹⁵。

時に、愛は人を狂わせ、すべきではないことをしてしまったり、愛する人以外のことが目に入らなくさせる。“愛は狂気(madness)”, “愛は耽溺(adiction)”などといわれる所以である。ドーパミン神経細胞は、ある種の依存癖(addiction; 薬物依存など)に深く関与する。実は、価値意思決定の学習がある方向にずれてしまうことが、依存癖を起こす原因である可能性が指摘されてもいる¹⁶。これは、報酬が、対象と強化の二面性を持つこととも関係する。どれくらいそれを強く求めるために行動が影響されるか(「強化」にあたる)と、実際にそれが得られたときにどれくらい「快」(hedonic)な経験であるか(「対象」にあたる)は、分離しうるのである。愛の場合と同様に。

ちなみに、ドーパミンの機能は、先述のオキシトシンと相互作用する。実は、人間の実験で、オキシトシンが他者をより信用しやすい傾向を生み出すことが示唆されてもいる¹⁷。ドーパミンとオキシトシンの相互作用は、まさに物質原理と情報原理の交錯の中で、脳の社会的な価値意思決定の計算につながるのである。

脳の数理から多様な愛へ

近年、大脳基底核関連回路の脳活動とその機能が、強化学習と呼ばれる数理的な学習理論と広く対応づけられることが様々な実験結果の蓄積から明らかにされてきた¹⁸。たとえば、前述のドーパミン神経細胞の活動は、この学習理論のある方式で使われる学習信号(強化信号)に対応づけられる。その信号は、報酬予測の誤差、すなわち「予測した報酬」と「実際に得られた報酬」の差、を利用すると考えられている。理論が新たな実験を求め、新たな実験結果が理論の更新を求めるという理想的な循環が生まれて、価値意思決定の脳機能理解が進展している。

誌面の都合から、強化学習理論を詳述しないが、数理的観点から端的には、マルコフ決定過程の実時間の近似学習といえる。その骨子は、現時点の評価、および評価を最大化するための行動選択にある。基本的構成としては、たとえば、現在の「状態」(入力)を評価し、その価値(状態価値)を推定する。ここで状態価値は、その状態で直ちに得られる報酬と、将来の状態の状態価値の和に対応する。したがって、現在の価値評価は、将来の先読みを含んでいる。主体は二つの学習——状態価値関数の学習と行動選択関数の学習——を必要とする。状態価値の学習と同時に、状態価値を最大化するための行動選択の学習を行う。粗く言えば、これらはそれぞれ、報酬の「対象」と「強化」、そして愛の「快」と愛を求める行動、に対応すると見なすことができる。

強化学習の方式には、バリエーションがいくつかある。上述の学習では、状態価値関数と行動選択関数を別々に学習するシナリオだが、状態と行動選択のペアについて、行動価値関数を形成して学習する方式もある。また、状態と行動のペアに対して次の状態がどうなるかという状態遷移の関数そのものも直接学習して、それを利用する方式もある。いわば、外界の状態変化を主体内部でシミュレーションできるようにするのである。外界

の部分的観測のみから価値関数を形成する、つまり状態推定と価値形成が共に必要な場合もある。生きていくためには、限られた情報から判断を迫られるからである。

実は、脳の価値意思決定では、これらの複数の情報処理が並列に存在するようだ。もしかすると、ヒトや生物のしなやかな知能の源泉は、そんなところにあるのかもしれない。たとえば、習慣的な行動(habitual behavior)やピアノを弾くなどの習熟した技能(skill)は一見すると、価値を最大化する行動とは無関係に見えるかもしれない。しかし、それらも、上述の複数の価値評価とその最大化の情報処理が並行に機能することから生み出されることが指摘されている^{19,20}。無意識の膨大な計算に、習慣や技能も支えられているのである。

愛は多様である。たとえば、ロマンスのように燃え上がる愛、長年連れ添った夫婦の愛、あるいは親から子への愛がある。じつは、こうした愛情の種類によって、異なる脳部位の組合せの活動が強まる、あるいは抑えられることが示されている^{15,21}。多様な愛の現れは、複数の価値意思決定の情報処理の異なる組合せとして実現されているのかもしれない。

愛する〈他者〉、その脳計算の深みと広がり

上述した外界の状態変化を主体内部でシミュレーションをする強化学習の方式は、理論的には、他者のシミュレーションにも適用可能である。そして、実際、私たちが行ったfMRI実験でも、他者の価値意思決定を勘案して価値意思決定する課題では、その他者を脳内でシミュレーションするような活動が観測できる。価値意思決定の学習では、報酬予測の誤差を学習信号にすることに特徴があるが、実はこの誤差信号の形成の仕方には、理論的な汎用性がある。最近の実験により、脳は、他者の意図の学習にも同様な学習方式を利用していることを示す結果も得られている²²。

ある場面で特定の他者に向けられる愛、その行動選択を考えよう。そこには、他者を含んだ状態

という入力、その他者の内部状態の推定およびシミュレーション、そして、その愛する他者に向けた特別な行動選択、などが含まれるのではないか。これらはまさしくこれまで述べてきた価値意思決定と共通の課題である。

脳の価値意思決定の研究に関連して、行動経済学やゲーム理論で盛んに行われていた実験パラダイムを取り込みながら、神経経済学あるいは社会脳科学という分野が興りつつある。これらの研究との交流が、価値意思決定の脳機能研究を深化させている。それは、計算論的精神病学(computational psychiatry)や、また社会的意思決定や社会的脳機能の脳計算論という新たな研究領域(computational social neuroscience)へと広がりを見せている²²。上述の他者のシミュレーションはその一例である。

ほかにも、たとえば、他者の報酬(対象と強化)と自己の報酬(対象と強化)の関係がある。他者の報酬は、己の報酬たりえるのか？ この二つの報酬の関係は、すべての他者で共通か、それともお互いの人間関係により変化するのだろうか？ 共感を感じる相手とそうでない相手に対して、その各々が報酬を受け取るのを見るとき、共感する相手の場合には特に線条体の脳活動が高いという知見もある²³。愛は与えられるものか、それとも与えるものか？ 与えられることが、それとも与えることが報酬なのか？ 相手の喜び(報酬)が、ワタシの報酬になるのか？ 報酬の対象または強化として、あるいは両方として機能するのか？ まさしく利己と利他の境界がどこにあるかを問いかけてくる²⁴。

私たちが社会的な動物であることに間違いはないが、他者から見られているというワタシは、意思決定に影響を与えるのだろうか？ 実は、直接的な報酬により活性化される脳部位の多くが、他者の好意的な評価により活性化されることが示されている²⁵。それでは、愛する他者がワタシのことをどう考えるか、それはワタシの報酬たりうることなのか？ その人のことを想像(シミュレーション)することで、相手が喜ぶことを、あるいは、相手よりワタシに好意をもつことを想像するだけで、

報酬になりうるのだろうか。まさしく〈他者〉の存在のもとで行われる社会的な価値意思決定、その脳情報処理は、〈愛〉の脳機能の秘密を語りかけるのである。

* *

愛は脳の情報処理である。愛と報酬は同義であるというのは過度の単純化だと思うが、愛と報酬の関係から、そして価値意思決定という視座から考えると、愛の情報処理の姿が垣間見えてくる。語りつくせぬ愛に憧憬と敬意を抱きつつも、物質原理と神経回路網、そしてその情報原理と数理原理が明らかにされていくことで、いずれ、豊穡な愛に関わる行動と機能は、豊かな情報処理として解きほぐされていくことと思う。もちろん、まだまだ道半ばではあるが。愛が情報処理だなんて味気ないだろうか。私はそうは思わない。むしろ、愛の豊穡さのより説得的な理解、その豊穡さへのより確かな想いにつながるように思える。それは新たな利己と利他の境界、社会観、心の理解へとつながるだろう。

文献

- 1—ロラン・バルト: 恋愛のディスクール・断章, みすず書房(1980)
- 2—G. レイコフ, M. ジョンソン: レトリックと人生, 大修館書店(1986)
- 3—L. ヴィトゲンシュタイン: 「哲学探究」抄 論理哲学論考, 法政大学出版局(1968)
- 4—イアン・ハッキング: 言語はなぜ哲学の問題になるのか, 勁草書房(1989)
- 5—ジェフ・クルター: 心の社会的構成, 新曜社(1998)
- 6—L. Young & Z. Wang: Nature Neuroscience, 7, 1048(2004)
- 7—K.-I. Kimura et al.: Nature, 438, 229(2005)
- 8—E. Demir & B. J. Dickson: Cell, 121, 785(2005)
- 9—甘利俊一: 神経回路網の数理—脳の情報処理様式, 産業図書(1978)
- 10—川人光男: 脳の計算理論, 産業図書(1996)
- 11—銅谷賢治: 計算神経科学への招待, サイエンス社(2007)
- 12—甘利俊一・深井朋樹編: 脳の計算論, シリーズ脳科学 1, 東京大学出版会(2009)
- 13—R. Wise: Nature Reviews Neuroscience, 5, 483(2004)
- 14—J. Olds & P. Milner: Journal of Comparative and Physiological Psychology, 47, 419(1954)
- 15—A. Aron et al.: Journal of Neurophysiology, 94, 327(2005)
- 16—A. Redish: Science, 306, 1944(2004)
- 17—M. Kosfeld et al.: Nature, 435, 673(2005)

- 18—W. Schultz et al.: Science, 275, 1593(1997)
- 19—H. Nakahara et al.: Journal of Cognitive Neuroscience, 13, 626(2001)
- 20—N. Daw et al.: Nature Neuroscience, 8, 1704(2005)
- 21—A. Bartels & S. Zeki: NeuroImage, 21, 1155(2004)
- 22—T. Behrens et al.: Science, 324, 1160(2009)
- 23—D. Mobbs et al.: Science, 324, 900(2009)
- 24—E. Fehr: in 'Neuroeconomics: Decision Making and the Brain', P. Glimcher et al. eds., Academic Press(2009)pp. 215~232
- 25—K. Izuma et al.: Neuron, 58, 284(2008)