

## 題目： 快楽が脳を創る

中原裕之

**Hiroyuki Nakahara**

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]

**Laboratory for Integrated Theoretical Neuroscience, Riken Brain Research Institute,  
2-1 Hirosawa, Wako City, Saitama 351-0198, Japan**

心と知能の根源は脳の活動である。それは神経細胞集団の発火活動の時空間パターンから、脳の多様な階層性に支えられ、遺伝子の制御と経験に基づく学習を経て、心と知能が生まれてくる。脳の各階層の数理構造の解明と階層を貫く数理理論を築き、意思決定と運動制御、認知と学習、情動や感情などの脳の情報処理の原理に迫る。

### 1 —— 疑問が脳研究の出発点

#### どんな脳を創りたいですか

この章では「快楽が脳を創る」という言葉と一っしょに、ちょっとした散歩をしたい。脳科学、とりわけ理論神経科学の最前線をかいま見てもらうために。

「快楽が脳を創る」と聞いて、あなたはどんなことを考えるだろうか？ たとえば、鉄棒での逆上がりをほめられた子供が、にこにこして何度も逆上がりを繰り返しているところとか、「快樂スープ」にぷっかりうかんだ脳がバクバク動いているのもいい。もし可能だったら、このページの余白に、あなたのイメージを書きとめてほしい。この章を読み終わったあとで、そのイメージが変わっているかもしれない。

でも、早とちりしないでほしい。快楽が脳を創ると言ったからといって、なんでも楽しいことをしていれば脳が元気になると言いたいわけではない。

「脳を創る」というのは、刺激的なキーワードだ。あなただったらどんな脳を創りたくなるだろうか？ ピカソやダリのような独創的な芸術を生み出す脳？ それとも村上春樹のような小説を生み出す脳？ サッカーの世界カップで、日本を優勝に導ける抜群の身体能力を生み出すような脳がもし創れたら、サッカーファンの私は限りなく嬉しい。残念ながら現在の脳科学はそこまでは進歩していない。それでも、脳を創ろうと考えることは大切なのだ。

でも、そもそも、快楽ってなに？ 脳を創るって具体的にどうするの？ 理論神経科学ってなに？ 疑問が湧いて当然で、その疑問が私たちの出発点となる。

## 体ではなく脳が覚えている

人の活動のおおもとには、脳活動がある。読者の方は何かスポーツをやっているだろうか？ スポーツのトレーニングをするときに「体で覚えるまで繰り返せ」とよく言われる。また、一流のスポーツ選手がその美技を決めた瞬間を振り返って、特に何か考えていたわけではなく体が自然に反応した、とコメントをするのを聞いたことがあるだろう。

これは、実は、ちょっとウソなのだ。なぜか？ 私たちが体で覚えたとか、何も考えずに反応した、と言うのは厳密には、私たちのいわば「意識」の中で考えなくても自然に動いたという意味に過ぎない。正確に言えば、体が覚えているわけではない。脳の一部にその運動パターンが記憶されているのであって、つまり脳が覚えているのだ。何も考えずに反応したのではなく、脳活動が起きて反応したのであり、脳活動すなわち脳の情報処理／計算が起きていたという意味では、確かに「脳は考えていた」のである。

脳活動を考えると、このような意識にのぼらない脳の情報処理も含めて考えることなのだ。一流のスポーツ選手は、手に汗を握る緊迫した場面で平然とその優れたプレーを見せつける。その動きを実現しているのは脳の運動制御であり、その優れた動態視力の実現は脳の知覚／認識機能であり、それらを支える心の平常心や強い精神力も脳の活動の産物だ。

## 人間のすべての活動は、脳活動の産物

スポーツだけに限らない。ダリの絵だって、村上春樹の小説だって、モーツァルトの音楽だって脳活動の産物だ。私たちが赤ちゃんを見て可愛いと思うのも脳活動の産物である。脳活動が社会を作るとも言える。私たちの道徳や礼儀なども脳活動の産物だ。政治制度や経済活動も私たちの脳活動のありかたを反映している。もっと長い時間で見れば、そういった人間のさまざまな活動、つまり脳活動の産物が時間というふるいにかけて文化や社会制度に結実する。

シェークスピアのロミオとジュリエットが、二人の間に生まれた恋に真剣でありながら、同時に彼らの各々の家族との絆と両家の長年の諍いの中で苦しむ。恋をするのも脳活動の産物だし、家族の絆が生まれるのも脳活動の産物だし、長年の諍いを作ってきた歴史ももとをたどれば脳活動の結果だ。このように考えると脳活動を考えることが、どんなに興味深く挑戦しがいのある作業かもわかる。

## 脳は計算高い!?

実は脳活動のほとんどは、私たちの「意識」にはのぼってこない。目の前のコーヒーカップに手を伸ばすときに、おおよその手の動きは意識したとしても、手首や指の関節の角度や各筋肉の強さ自体を、例えば「人差し指の先端の角度一五度、第二関節は二〇度」などと、私たちは事細かに意識しない。意識しなくても、実は脳は計算してくれている。

目の前にあるテレビスクリーンでドラマを見ているときに、ある俳優がどう振る舞うか

に見入っていたとしよう。その時、実際にモニター上にあるのは多数の光の点である。その俳優と、彼が手にしているワイングラスが、それらの光の点の集合からどのように脳内で情報処理され、俳優とそのグラスとして私たちに「知覚」されているのかは意識しない。しかし、これも脳の情報処理の結果である。

テレビドラマのコマーシャルの間にトイレに立って、トイレを済ませて、自分のソファに戻る。コマーシャルが終わり、先ほどの俳優がテレビに現れる。トイレに行くまで我慢することも、どのようにその我慢を実現するかも、私たちの脳が学習して覚えたことである。トイレから出てソファに戻るときには、私たちは脳内の神経細胞の活動で作られる短期記憶を利用している。コマーシャル後に再び現れた俳優が、先ほど見た俳優と同じだと認識するのも脳の計算の結果である（ちなみに、人間の脳は一〇年ぶりに会った友人でもその名前と顔を一致させてしまうが、これはかなりすごい計算なのだ）。こう考えてみると脳の計算はいたるところで行われている。

このように私たちの行動は脳活動の結果なのである。もっと言えば、脳の計算、その情報処理の結果が行動として現れるのだ。脳の計算とその情報処理を理解するのは困難な作業だ。私たちが普段特に意識しないような脳の計算、それらを実現させること自体そんなに容易ではない。いや、むしろ私たちが特に意識しないでやっているような脳の計算ほど難しいとも言える。

「あの人の行動は計算高い」と言うと、あまりいい意味ではない。しかし、実は脳の計算から見ればほぼ全ての行動は計算高いのだ。私たちが何気なく日常を過ごせること自体、無意識のうちに行われる膨大な脳の計算に支えられている。

## 2——脳は計算する

### 喜怒哀楽は脳の情報処理の一部

「脳は計算する」と突然言われても面食らうかもしれない。脳の計算って何、というのは至極当然の質問である。そもそも脳活動が行動を決めると言うときに、私たちは何を想定しているのか？ 脳に入ってきた信号（例えば眼から入ってきた視覚情報）をもとに、脳の中では何かが起きる、つまり脳活動が起きる。そしてそれが、行動（例えば手を伸ばすなどの運動）を引き起こす。端的に言えば、脳はこの入力信号から出力信号への変換、つまり情報処理を行っているのである。

入力信号から出力信号の変換と言われると、味気ないかもしれない。しかし、そんな表面的なもので判断してはいけない。その変換、情報処理は、上述したように私たちが無意識のまま膨大な計算をしてくれている。しかもその情報処理には驚嘆すべき豊かさがある。私たちの感情——泣く・笑う・怒る、これらも脳の情報処理の一部なのだから。

## 脳の計算を再現する

情報処理の実体を理解するには、その数理原理と行われている計算を理解しなければならない。パソコンのCPUのチップを眺めていてもその情報処理の原理自体はわからない。もちろんチップの構造や物質的な特性、そして回路図は有益な情報を与えてはくれる。しかし、そこで行われる計算、そしてその数理原理を把握できなければ、やはりCPUが何をしているかはわからない。

同様に脳の機能を本当に理解したければ、脳の計算を理解する必要がある。「脳を創ろう」と言うことは、脳の計算を実際に再現できるぐらい徹底的に脳を理解しよう、そして、脳の機能を理解したと宣言するには脳の計算を再現できなければだめだ、という決意表明でもある。

脳の計算を再現できるくらいに理解するには、脳の計算の理解を突き詰めていくことが求められる。端的に言えば、その作業には脳の計算、ひいては脳の情報処理の数理原理を解き明かすことが求められる。生半可な理解は許されないのである。

例えば、携帯電話で、電話番号を押して友人と話すことを考えよう。この情報処理のプロセスはどうなっているのだろうか？ 大抵の人は、電話番号を携帯に入力して、発信ボタンを押す。その番号が無線で基地局に送られ、友人の携帯の場所が特定され、そして無線でつながれる、というぐらいは答えられるだろうか。しかし、じゃああなたが携帯電話網を創ってくださいと求められても、この程度の理解でとてもできるものではない。

携帯電話の端末のさまざまなサービスは、さまざまな情報処理を行ったうえで、その画面に表示される。携帯電話から無線で基地局に情報を送るのだから一つの情報処理だ。無線でつながっている膨大な数の携帯電話端末から、そして数多くある無線基地局の中から、友人の電話番号に合致する端末を探し出し、本人の端末と最終的にうまく結びつけるのだから一つの情報処理だ。これらさまざまな情報処理には、各々の数理原理が存在している。その数理を理解して、実際に情報処理として実現可能な方法を見出せなければ、その機能を再現することはできないのである。

## 脳の情報処理と時代の先端技術

脳は計算すると述べた。多くの場合、その時代の情報処理の先端技術は、その時代の数理工学の粋と深い関係がある。面白いことに、各時代で、脳の機能を語るとき、その時々々の情報処理の先端技術の発展が、その語り口に影響を及ぼしている。言いかえれば、その時代の情報処理に関わる先端技術がしばしば脳を語るメタファーとして用いられてきた。それを少し見ていこう。

例えば、電話の発明は当時画期的なものだった。この技術の根幹は、信号の送り手と受け手がいるということである。言いかえると、電話線の端に座っている各々の人は、信号の送り手であり受け手である。二つはなれた場所でお互いに信号をやりとりして、その意味を理解して、逆に何らかの信号を送り返す。つまり、そこには「入力」「反応（出力）」

という連鎖が成立する。この見方は、入力と反応のペアを積み上げていけば、人や動物の知能が理解できるだろうという、心理学などで或る時期隆盛を誇った行動主義の見方に相通じるものがある。

そして、少し時代を下ると、二〇世紀の二つの世界大戦そして戦後の冷戦期にわたって、大陸間弾道ミサイルの技術の発展、その平和的利用で宇宙への人工衛星打ち上げや月面着陸船など、目を見張る技術的発展があった。これらの技術と、フィードバックの概念、サイバネティックスの見方は密接なつながりがある。

普通の生活の身近なところでも、実は同様の技術が使われている。最もわかりやすいのは、エア・コンディショナーの自動運転である。設定温度を決めるとエアコンはその温度を目安にして、設定温度より部屋の温度が低い（高い）ときには動き出し、温度が高け（低）ければその動きを止める。つまり目標の温度に向かって、エアコンから放出する空気の量や温度が自動的に調整される。

月に宇宙船を着陸させるには、絶えず宇宙船の軌道を監視しつつ、目標の着陸地点まで誘導しなければならない。似たような技術が使われているのである。この目的に向かう自動調整、そして外乱に対する自動的な学習と調整、これらの視点は、サイバネティックスと呼ばれる数理工学に深く関連する。「入力」と「反応（出力）」のペアを考えると、サイバネティックスでは、自らの目的や内的状態だけでなく、「外的状態」のフィードバック（情報）を取り込んだ上で、情報処理を考えなければいけないという点が強調されるのである。

このサイバネティックスは、例えばシステム理論の学問とあいまって、脳の機能を理解するのに重要な視点を提供した。その見方は現在の脳研究のそこかしこに息づいている。

### 認知主義—プロセスの解明

このサイバネティックスの発展を追うように現れたのが、大型計算機の技術である。大型計算機はのちにパーソナルコンピュータへとつながった。これに呼応するかのように認知科学と呼ばれる研究が大きく立ち上がっていった。これは、認知主義とも呼ばれる。端的に言えば、例えば行動主義の研究と比べると、認知主義では「入力」と「出力」の間にあるプロセスの解明に大きな焦点があてられている。そして、そのプロセスのことを「認知」と呼称し、その「情報処理」を解明しようという考え方で研究が進められた。

つまり「入力」→「情報処理」→「出力」という流れを考え、それが、心理学や脳科学でいうところの「知覚」→「認知」→「運動」という流れに対応しているのである。パーソナルコンピュータには、メモリーという短期的に情報を溜めておく部分と、ハードディスクのように長期的に情報を溜める部分の二つに分かれていることをご存知だろう。認知主義でも、これに対応するかのように脳の中で短期的に記憶を溜めておく処理と長期的に記憶を溜める処理を区別することがしばしば行われた。

この認知主義は、並列情報処理技術の研究と並行するかのように、後に、接続シ

ズムと呼ばれるような心理学的研究へと発展していった。

脳機能の理解はその時々の情報処理の先端技術の発展に刺激を受けてきたのと同時に、脳機能を理解したいという人間の強い欲求が、例えば、並列に情報を処理するとか、適応的に学習するといった、新たな先端技術や数理手法の開発にも刺激を与えてきた。パーセプトロンとか連想記憶は良い例である（第2章参照）。これらの研究が、時代をおいて、コネクショニズムの研究とあたかも旧友に再会するかのように合流して、脳科学の研究へとつながったのだ（このような「再会」は他にもあって、経済学との再会については後に述べよう）。

### 脳の情報処理とインターネット

パソコンから、今や時代はインターネットとなった。インターネットは地球規模でさまざまな情報処理機械をつなげている。今や、パソコンだけが主たる情報処理機械ではなく、ゲーム機器や携帯電話もかなり高度な情報処理機械だ。地球全体がお互いに結び付けられた大きな情報処理機械、そう、「地球脳」みたいになってきている気がしないだろうか？（図11-1）

グーグルという大変人気のあるインターネット検索は、インターネット上にちらばっている情報、つまり記憶を素早く引っ張り出すことをする。これと似て、人間の脳の驚くべき性質の一つは、パソコンのハードディスクなどと比べると素早く記憶を引き出せるところにある。

私たちの脳の中には、グーグル検索と似たようなことができる情報処理機構が既にあるのだ。すると、素朴な疑問が出てくる。グーグルの検索は、脳の中での記憶の読み出しとどれくらい似ているのだろうか？ 実際、グーグルの検索を実現している数理手法は、理論脳科学の研究で使われる数理手法とは、まあ親戚みたいなものなのだ。

数理という視点から見ることで、インターネットのグーグル検索と脳の記憶の想起の方法に関連性が見出せるということは色々な意味で面白い。脳の情報原理を明らかにする数理研究を促進することで新しい先端技術の開発、たとえば検索技術の開発につながるかもしれない。一方、先端技術の発展は、私たちの脳を語る言葉をどのように変えて、そして脳の理解をどのように促進するのだろうか。興味はつきない。

### 脳という「もの」——脳は物質

脳は情報処理をするということを強調してきた。その意味では、たしかに脳は情報処理をする機械と見なせる。私たちの心や知能の源泉となっている、とてつもなく賢い情報処理機械である。しかし、脳は確かに物質でもある。脳は「もの」でもあるのだ。この「もの」が私たちの心や知能を生み出している（心の中でこの言葉を何度か復唱すると、その不思議さがじっくり胸に染みってくる）。

この脳という「もの」は精緻な構造と多様な脳部位をもっている。各々の脳部位は、膨

大な数の神経細胞と、その神経細胞の活動を支えるさまざまなほかの細胞（たとえばグリア細胞）で構成される。もっとマイクロには、多様かつ多数の物質から構成されている。その中でも際立った物質群は、脳活動に直接影響を与える（一例として、ドーパミンという物質については後で詳しく述べる）。つまり脳の情報処理に変化を与え、それは最終的には私たちの心や知能の働きに変化を与える（もちろんその与える影響の大きさにはかなりの程度の違いはある）。私たちの心や知能が、脳という「もの」から生み出されているとき、確かにそこに含まれる物質からの影響を受けつつ、立ち現れてくるのである。

しかも、この脳という「もの」は、実は最初是一个の受精卵という「もの」に過ぎない。受精卵の中に脳という「もの」があるわけではない。受精卵から個体（赤ちゃん）になる。赤ちゃんは脳という「もの」をもってはいるが、そこにはまだ大人の心や知能ほどの情報処理はもっていない。しかし、それらを備えられるように育っていくという意味では、赤ちゃんの脳という「もの」にも、そういった心や知能の「たね」は確かに存在する。

### 脳の学習はシナプスの変化

赤ちゃんは、生まれたときから大人になるまで驚くべき発達を遂げ、さまざまなことを学んでいく。さらに私たちは小学校、中学校、そして高校や大学、あるいは社会人として多くのことを学ぶ。生涯学習という言葉もあるぐらいだ。この「学び」の根本には、脳における学習がある。脳の中にあらかじめこれらの学ばれる事柄が存在しているわけではない。脳があらかじめもっているのはこれらを学べる能力、つまり学習能力をもっているのである。

脳の学習とはなんだろう？ 学習というからには、学習する前と学習した後で脳の中でなんらかの変化が起きていることが必要である。さまざまな変化が脳の中で起こりうるのだが、特に神経細胞間のつながりかたの変化が、もっとも中心的な役割を果たすと考えられている。

神経細胞同士の間をつなぐ主なものとして、シナプスと呼ばれる場所がある。このシナプスのつながりが強くなったり弱くなったりするのである。つながりが強くなれば、それだけシナプス前の神経細胞から、シナプス後の神経細胞へ送られる信号が強くなるし、つながりが弱くなれば送られる信号が弱くなる。このつながりのことを、シナプスの重みと呼ぶ。つながりが強くなる、あるいは弱くなることを、重みが強くなる、あるいは弱くなるというふうに言う。脳で学習が起きているときには、このシナプスの重みが変わるのである。（図 11—2）

実は、ある神経細胞にくっついていてシナプスの数が増えたり減ったりすることでも、情報の流れを変化させることはできる。あるいは個々の神経細胞の活動特性そのものを変化させることでも、その神経細胞の情報処理を変化させることができる。したがって、シナプスの重みの変化だけが学習をつかさどると考えるのは過ちであり、あくまで情報の流れを変化させる中心的役割をつかさどるものとして、シナプスの重みの変化を取り上げて

いるのである。

### 脳の学習——“良い方向”が大切

脳の学習がシナプスの重みの変化で実現されると述べたが、実は、変化するだけでは不十分だ。ただ変わればいいわけではない。学習と言うからには、脳の情報処理を“良い方向”に変えていくものでなければならない。そのためには、何らかの意味で“良い方向”を決めなければならない。

“良い方向”を決めるのはそんなに簡単なことではない。神経細胞は膨大な数がある。多数の神経細胞が同時にさまざまな活動をし、お互いに影響を及ぼすことで脳の情報処理は実現する。脳は並列計算機だといわれる所以である。多数の神経細胞で並列に情報処理が行われている。

脳で学習が行われるとき、シナプスの重みの変化は多数の神経細胞で同時に起きる。つまり学習自体は並列で進む。情報処理の流れも学習も並列だ。学習が起きることで、各神経細胞の情報処理も変化し、ひいては神経細胞の回路網全体に流れる情報も変化する。同時に色々なところで変化が起きるのだから、うまくお互いの調整が図れなければ全体の变化として“良い方向”にいくかどうかは定かではない。

回路網全体の情報処理の変化は行動の変化につながる。行動の変化は、学習そのものにも変化を与える。たとえば今まで取っていた行動が変化すれば、動作も変わるし、注意を向ける対象も変わるし、それによって目に入ってくるものも変わるだろう。それによって脳に入ってくる情報も変化してくる。入ってくる情報が変われば、脳でも異なった新たな情報処理が生まれてくる。

先ほどの情報処理に応じて起こった学習による変化が、この新たな情報処理にどのような変化を与えるかは定かではない。先ほど“良い方向”を求めて行われた学習がこの新たな情報処理に悪影響を与えないか、この新たな情報処理も“良い方向”へガイドしているかを確かめておくのは脳にとって大切である。さまざまな情報処理全体に対して通用するように“良い方向”は最終的に決められていく。

## 3——脳の隠されたデザイン

### 良い方向に向かう

脳の情報処理には隠されたデザインがあると考えられる。脳にあらかじめ備わっている学習能力こそ、脳の情報処理の隠されたデザインである。そのデザインの根幹に、この“良い方向”の決め方がある。

神経科学、とりわけ理論神経科学では、この“良い方向”の決め方に大きな関心をもっている。この“良い方向”は、多数の神経細胞の活動に関わっている。また神経細胞同士

がお互いにどのように結合しているかに依存する。そしてどの情報処理のときに、どの神経細胞の集団が一緒に活動を高めるかにも依存する。このように多数の因子に関係してくるので、本質的に高次元の空間の中で、“良い方向”が決められることになる。もっと言えば、線形の空間になっていれば数学的に比較的簡単になるのだが、神経細胞の活動は非線形性をもつので、それも考慮する必要もある。

神経細胞の活動は互いに影響を及ぼしあうので、その相互作用も考慮されなければならない。相互作用と言われると普通は二体間の相互作用を考えるかもしれない。しかし、厳密に言えばこれは二次の相互作用で、多数の因子（今の場合は神経細胞）が関係するときには、三次、四次の相互作用、一般には  $n$  次の相互作用が存在しうる。さらに神経細胞の活動や、その情報処理の特性から、その空間は私たちが普通に考える空間とは異なっていて、曲がった空間になっているし、特異点があったりする。“良い方向”をきちんと考えるには、しっかりとした数理的視点をもつことが必要不可欠なのだ。

さらに、ひとくちで脳と言っても、脳にはさまざまな部位、たとえば大脳皮質、小脳、大脳基底核、海馬などがある（ここに挙げた各々の脳部位は更にいくつかの部位に下位分類される）。これらの異なる脳部位は、それぞれ特徴的な情報処理の機構をもち、また人間や動物の心や知能の異なる側面に関係する。情報処理機構が異なれば、その「隠されたデザイン」も異なってくる。当然“良い方向”の決め方も異なってくる。

この各情報処理機構の“良い方向”の決め方が、脳の隠されたデザインの核である。今までの研究で、たとえば、小脳の“良い方向”の決め方は、パーセプトロンという情報処理機構での“良い方向”の決め方に近いと考えられているし、海馬では連想記憶の“良い方向”の決め方、大脳基底核では強化学習理論での“良い方向”の決め方（これについては後に詳しく述べる）に近いと考えられている。

ただし、各々の脳部位に予め存在するデザインだけで“良い方向”が決まるわけではない。その部位の神経細胞活動、つまりどんな情報が流れるかも“良い方向”の決め方に影響を与える。

そして、もちろん“良い方向”に学習がおきることで何の情報も流れるかも変化していく。お互いに影響を及ぼしあう、或る脳の部位の情報処理の変化は、そこから他の脳の部位に送られる情報の変化につながる。その送られる情報の変化は、またその情報を送られた脳の部位の情報処理に変化を与え、それはその部位での学習を促す。各神経細胞の小さな変化は、このようにして脳全体の情報処理の変化とつながっている。

## 脳の快樂

脳の各部位の各神経細胞が“良い方向”に導かれ、その活動を変化させる。この“良い方向”こそ、あたかもその神経細胞に与えられた「快樂」のようではないか。各神経細胞がその快樂を求めてその情報処理を変化させていく（図 11-3）。

快樂を決めているのは、各脳の部位に隠された脳のデザインなのである。デザインが異

なれば快樂の現れ方は異なる。各々の快樂が、私たちが日常の行動レベルで考える快樂と同じである必要はない。その脳部位の情報処理機能を促進するように決められていればよい。もちろん脳全体の情報処理の調整は別の問題として存在する。いずれにしても個々の神経細胞がその快樂に従って変化していくなかで、その脳部位での情報処理、ひいては脳全体としてその情報処理が促進される。

一方、行動レベルの快樂とある程度合致するデザインの脳の快樂があっても不思議はない。そもそも私たちの行動は脳の計算の結果として起こる。だから、私たちの行動にとって重要な要素が、脳のデザインの中で重要な要素になっていても不思議はない。日常レベルの快樂に対応するかのようにデザインされた脳の快樂、例えば生存に必要な食べ物や飲み物を得るといったような行動が促進されるようなデザインが脳にないほうが不思議だろう。後述するように、ドーパミン神経細胞はそういうデザインに深く関わっている。

脳の計算が行動を決める一方で、行動も脳の計算を形作る。選ばれた行動によって、脳が得る外界からの入力も変わる。脳で処理される情報も変わるので、“良い方向”、つまり快樂も変わり、脳の学習も変わる。快樂は、脳の隠されたデザインと私たちの行動の両方に支えられている。私たちの学び、あるいは教育とは、この両方の相互作用を利用しながら行われなければならないのは明白だ。脳科学と教育の関係は、今後、より深いものになっていくだろう。

### 脳は先読みをする

脳は先読みする能力は私たち人間にとって本質的な能力だと思わないではいられない。

私たち生き物と、そして脳は、たしかに時間の中を生きている。脳はたえまなく動いている。スポーツ選手の体が勝手に動くときにも、「ぼーっとして何にも考えてなかった」というときにでも脳は膨大な計算を行って情報処理をしている。

外界から絶えず入ってくる情報を、ある程度までは脳は自動的に処理している（テレビの前に座って画面をぼーっと眺めているときだって脳は情報処理をしている）。この自動的にとにかく入力を処理するのは、いわば受動的な情報処理と言える。

能動的な情報処理だってある。私たち生き物はいつもただ黙って与えられることを待つわけではなく、自ら必要なものを求めて行動することもできる。これは能動的な行動である。これと同様に、能動的な情報処理とは、ただ入ってくる情報を受け身で処理するのではなく、あらかじめ入力情報を予測し、もっと一般的には外界の状況あるいはこれから起きる事柄を予測し、その情報処理および行動を準備することを指している。

この外界の状況の変化を予測するには、脳の情報処理機構の中に、外界の状況をシミュレートできる能力が備わっていなければならない。この考え方は神経科学ではたとえば内部モデルなどと呼ばれている。

### 未来を作り出す能力

能動的な情報処理とは、脳が先読みをするということである。これから起きるだろうことを脳の中であらかじめ予測している。あるいは、起きてほしいことを予測、つまり望むこともあれば、起きて欲しくないことを予測することもあるだろう。言いかえれば、脳の先読みする情報処理は、脳の中でさまざまな未来を作り出す。この能力があるからこそ、私たちは未来を考え、希望を抱き、または将来を思い悩むことになる。

能動的な情報処理は受動的な情報処理にも役立つはずだ。受け身でただ入ってくる情報をひたすら情報処理する場合に比べて、能動的に欲しい情報を入力の中から探すほうが処理する量は少なくてすむ。処理される情報が変わればそれは行動にも変化を与える。その行動の変化は次の情報処理をも変化させる。ここでも、さきほど脳の快樂について述べたときと同じような循環が成立する。しかも、この能動的な情報処理自体、脳に仕組みられたさまざまな快樂を通じて、学習し、変化していく。

能動的な情報処理の重要性は、実は私たちが普段耳にする仕事のノウハウでも強調されていたりする。たとえば、ビジネスの世界で仮説検証型のアプローチの重要性がよく言われる。これも脳の能動的な情報処理の効率性を反映していると考えればもっともなことに思えてくる。私たちが日ごろ耳にする格言の裏には、脳の情報処理のありかたが気づかれずに反映されているのかもしれない。

### 想像力と創造力が脳を生み出す

ちょっと話が飛躍するが、アメリカ合衆国の首都であるワシントンD. C. にはスパイ博物館がある。その展示物を見ていると、現実のスパイのための技術開発が、ハリウッドの映画に影響されていることがわかる。映画を作るときに、未来の技術が映画の中で実現されていたほうが面白い。そこで、想像力を働かせ、現実にはまだ実現されていない技術が、映画の中では実現される。そして逆に、その映画が、未来の技術開発の創造力に刺激を与えたのだ。

想像力と創造力、この二つの力がお互いを刺激し、先導したのだ。想像が新しい未来を示し、創造が新しい未来を生むのである。

脳の本質的な能力である先読み能力の中でも、この想像力と創造力は、特に高度な先読み能力だ（繰り返すが、これらも全て脳の計算だ！）。脳の先読み能力は、行動に影響を与え、行動は脳の快樂に影響を与え、そして先読み能力自体をも変える。想像が予測を変え、それは行動を変え、快樂を変え。創造で行動が変わり、行動が変われば予測も変わる。この先読み能力が人間の脳で発達しているのは、私たちの脳に隠された快樂がそれを誘導しているからではないだろうか。想像力あるいは創造力豊かな人間は、その脳の快樂と行動のサイクルをうまく循環させているのだろう。この視点はきっと教育にもあてはまる。

実は、想像力と創造力の相互作用は、脳研究という科学の現場にも存在する。各時代の先端の情報処理技術や数理が、脳機能の理解に役立てられてきたことは既に述べた。私たちの脳のグーグル検索はどうなっているのか、そんな素朴な疑問でも、想像力が創造力と

出会う端緒となる。

### 思いがけない快樂とドーパミン神経細胞

脳の先読みの最も原初的な形態は、今この行動をしたらあれが得られるかもしれないという報酬の予測だろう。ここで報酬と言ってもお金とか会社員のボーナスのことではない。とりあえずはもっと本能的に必要なもの、例えば食べ物とか飲み水を想定しよう。これらの報酬を得る“良い方向”に学習が進むようなデザインが脳にはきっとあるはずだろう。これらの報酬なしには生き物、つまり脳は生きていけないのだから。

ちょっと待てという人がいるかもしれない。食べ物や飲み水を報酬と考えるなら、お金は報酬になるのか、好きな人に喜んでもらう、愛する人の笑顔は報酬になるのか、という質問も出るかもしれない。すばらしい質問だ。そういう率直な問いかけが科学の新しい発展につながるのだ。実際、そういった質問が今の脳科学に新しい展開をもたらしつつある。そのことは後で話すことにして、今は報酬としてまず食べ物や飲み水を想定しよう。

報酬の予測が、最終的にその報酬の獲得につながるには何が必要だろうか。この問いかけに対して、適切な抽象化をしつつ本質的な状況をえぐりだした学習アルゴリズムが「強化学習」と呼ばれるものである。このアルゴリズムは、計算科学、機械学習、人工知能の分野などで研究されてきた。実験と理論の見事な融合的研究から、どうやら脳にはこのようなアルゴリズムが実際に存在するらしいことがわかってきた。その主役がドーパミン神経細胞である。

ドーパミン神経細胞は脳の奥深くの、比較的小さな領域に存在している。その領域は、黒質（正確には黒質緻密部）と呼ばれる部位や、腹側被蓋野と呼ばれる部位だ。中でも、黒質のドーパミン細胞は、大脳皮質の奥にある大脳基底核と呼ばれる神経回路に対して多くの出力（基底核にとっては入力）を送っている。（図 11-4）

大脳基底核の中でも線条体という部位は大脳皮質からの入力を豊富に受けていて、そこでドーパミン神経細胞からの入力も豊富に受ける。線条体は、さらに尾状核、被殻、そして側坐核に分かれる。尾状核と被殻は主に黒質からドーパミン神経細胞の投射を受けていて、尾状核は意思決定、そして被殻は運動出力に関連が深いと考えられている。一方、側坐核、そして側坐核の出力先である腹側淡蒼球と呼ばれる部位は、主に腹側被蓋野からドーパミン神経細胞の投射を受けている。

側坐核は、扁桃体や視床下部などとの関連が相対的に強く、意思決定や判断に関連すると同時に、情動や感情の面にも関連がより深いと考えられている。ドーパミン神経細胞は基底核回路以外にも、大脳皮質を含む脳全体のかなりの部位に投射している。ドーパミン神経細胞がひとたび活動をはじめると、その投射先であるこれらの脳の部位に、その活動が伝えられ、その影響が脳全体に及ぼされることになる。

### ドーパミンの学習効果

この基底核回路で、ドーパミン神経細胞の活動を利用した学習プロセスの研究が大きく進展しつつある。以前から、動物や人間の実験から、ドーパミン神経細胞は「快」の入力が与えられたときに強く反応することが知られていた。「快」の入力というのはわかりづらい学術用語だが、ここでは単純に「うれしいこと」あるいは「報酬」と考えておけばいいだろう（ふつう動物実験では食べ物や飲み物を動物が得ることをさす）。

近年の研究でさらにドーパミン神経細胞は「快」の入力よりも、むしろ「予測されていた快」と「実際に得られた快」の差に強く反応していることがわかってきた。そのプロセスを理解するために、できるかぎり簡単に定式化してみよう。

実際に得られた快、これを「実際の報酬」、予測していた快を、「予測された報酬」と呼ぶことにする。ドーパミン神経細胞は、「実際の報酬」－「予測された報酬」に主として反応するというのが近年の研究の成果である。この二つの量の差を取ることから、しばしば、ドーパミン神経細胞の反応は、「報酬予測誤差」（＝「実際の報酬」－「予測された報酬」）を表すと称される。

このドーパミン神経細胞の反応を利用して学習するとしたなら、どんな学習をすれば、生物にとって役立つ学習になるだろうか？ まずは思考実験をしてみよう。

ある状況で右か左かの二つの選択を迫られて、とりあえず右を選んだときドーパミン神経細胞が強く反応したとしよう。これは何を意味するのか？ ドーパミン神経細胞が強く反応したということは、「報酬予測誤差」が正の大きな値をとっていることになる。すなわち、右という行動選択によって実際に得られた報酬が、予め想定していた報酬よりもずっと大きかったことを示す。

右を選ぶことで、思っていたよりも得をしたことを、ドーパミン神経細胞の反応は示している。選択した行動は思っていたよりも良い選択だったということだ。ならば、再び同じ状況になって右を選択するとき、予測する報酬をもう少し多く期待することになるだろう。

また、そもそも右を選択するとき、前回よりは、もう少し自信を持って右を選んでいだろう。よって、ドーパミン神経細胞の反応を利用した脳の学習としては、その状況での報酬予測のレベルをあげ、そのときに選択された行動をより頻繁に選択するように脳活動を変化させていくのが望ましいということになる。

実際の脳ではどうなのだろうか？ 実は、この推測はあたっているようなのだ。ドーパミン神経細胞の活動を利用した脳の学習でも、同様のことが起きていることがわかってきている。ドーパミン神経細胞活動が表す報酬予測誤差を利用して、より精度の高い報酬予測を実現できるようにしつつ、同時により大きな報酬を獲得できる行動を選択するように脳は学習を行っているらしいのだ。（図 11—5）

ドーパミン神経細胞の活動が脳活動全体に影響を与えることは知られていた。徐々にはっきりしてきたのは、神経細胞間の結線の変化（シナプスの重みの変化）、つまり活動が脳の学習を引き起こすことがわかってきた。ドーパミン神経細胞の活動が強まったときに選

択された行動が、のちにも好まれて選択される傾向があることも示された。このシナプスの変化を通じて、精度の高い報酬予測とより大きな報酬が得られる行動選択が実現できるように、脳の学習を促進するらしいことが明らかになりつつある。さらには、極端な例としてドラッグなどの薬物依存をやめられなくなる性向と、ドーパミン神経細胞の異常な活動の関係が既に指摘されている。

## 4——脳の強化学習とは

### 報酬予測誤差が学習を強化

上述した仮説は「報酬予測誤差」仮説と呼ばれる。この仮説はその提案後、大変な注目を浴びた。ドーパミン神経細胞の報酬予測誤差の振る舞いが、「強化学習」(reinforcement learning)と呼ばれる学習アルゴリズムで利用される学習信号と非常に似ていたからである。強化学習アルゴリズムでは、その学習信号にまさしく報酬予測誤差を使うのである。つまりドーパミン神経細胞の活動の振る舞いと生物の行動レベルの変化が、強化学習という学習アルゴリズムのもとで統一的に理解できる可能性が示されたのである。

「強化学習」というのは、例えば機械学習・人工知能・ロボット学習などで盛んに研究されている学習アルゴリズムのことで、このアルゴリズムの問題設定は後述するように非常に単純である(設定が単純だからと言って、その問題を解くのが簡単というわけではない)。

この学習アルゴリズムはなかなか強力である。例えば計算機(パソコン)に、この学習アルゴリズムを実装して、ゲームで優勢になると報酬を与えるようにしてやると、バックギャモンというゲーム(西洋双六のようなゲーム)で人間の世界チャンピオンと同等に戦えるぐらいの学習能力を示すようになったのである。オセロゲームなど他のゲームでもかなりの能力が示されているし、ロボットなどへの実装も試みられている。

「強化学習」で最も本質的なのは、報酬の獲得そして報酬の予測という概念である。動物や人間が行動を決める際には、報酬の予測は最も本質的だし、行動選択は報酬予測と密接に関連しているからだ。

最も単純な行動選択とは、二つの選択のうち的一方を選ぶということだろう。右を選ぶか、左を選ぶか、というのも一例だろう。道に迷って目的のレストランに向かうのに右の道を選ぶか、左を選ぶか、サッカーのゲームでドリブルをしていて、目の前に迫ってくる相手ディフェンダーの右を狙うか、左を狙うか、目の前に差し出された右と左にある二つのケーキのどちらかを選ぶか、ライオンに追いかけて右のヤブに逃げ込むか左の穴に逃げ込むかという生死に関わる選択だってある。

これら行動選択の目安になるのは、将来の報酬予測である。右か左か、各々の選択を取った後で何が起きるか、その起きる事柄が自分にとってどちらが好ましいのかが、その選択の基準となる。

一般的な形でこの問題を定式化してみよう。まずはそのおかれた環境や状況がある。これらを総称して「状態」と呼ぶことにしよう。その「状態」の中で、人間や動物などが行動選択をするわけだが、その行動を選択するものを「主体」と呼ぼう。さて「主体」はある「状態」のなかで主体的にある「行動」を選択する。するとその選択された行動によって、今までの状態は次の状態へと移り変わっていく。(図 11—6)

つまり「行動」は「ある状態」から「次の状態」への遷移を引き起こす(数理的に言うと、マルコフ過程を想定していることに該当する)。そしてこの状態の遷移時に「報酬」が「主体」に与えられると考えるのである(もちろん報酬がまったく与えられない状態遷移もありうる。そのような「無報酬」のときも含めて「報酬」と呼んでいる)。これで強化学習の基本的な役者、「状態」「行動」「報酬」の三点セットが出揃った。

### 強化学習アルゴリズム

まずは一番簡単な右と左の選択を考えよう。図 11—7 上のように三つの状態  $S_0$ 、 $S_1$ 、 $S_2$  を考える。右と左の選択を各々  $A_1$ 、 $A_2$  と書き、右と左の各々の報酬を  $R_1$ 、 $R_2$  とする。話を簡単にするために今、 $R_1=100$ 、 $R_2=0$  という報酬になっているとしよう。

さて、今私たちは  $S_0$  の状態にいる。行動選択をどのようにしたらよいだろうか。つまり右と左どちらを選んだらよいだろうか。もちろん  $A_1$  の行動選択、つまり右を選ぶのが正解なのは明白だろう。ここで注意しておいて欲しいのは、 $A_1$  を選ぶためには、 $S_0$  の状態にあるときに既に  $S_1$ 、 $S_2$  各々の状態に移ったときに得られる報酬を予測できていなければならないことである。こんな単純な状況でも報酬予測ができるということは、適切な行動選択ができるために必要な能力なのである。

それでは図 11—7 下の状況ではどうだろう。 $S_0$  の状態にあるときに  $A_1$  と  $A_2$  どちらを選んだらよいだろうか? もし  $S_1$  と  $S_2$  だけを考えていたら、この場合、 $S_2$  への行動選択つまり  $A_2$  を選んだほうがよいことになる。しかし、 $S_1$  の状態のあとさらに  $A_1$  を選んだ場合、より大きな報酬が得られるはずだ。それなのに  $S_2$  を選ぶのが本当によいのだろうか。

今の例では行動選択を二回すればよいのだが、より一般的な場合には何度も行動選択をしたあとでやっと報酬が得られることも多い。そんな場合に目の前の利益だけ考えて行動選択をすることは決して賢い方法とは言えない。このような状況を、報酬を獲得するまでに時間遅れがある、と言う。

時間遅れの報酬に対処するためには、目先の報酬と遠い先の報酬の両方を考慮に入れた上で行動選択をしなければならない。また現在の状態がどれだけ良い状態なのかを判断するには、この状態から将来どれだけの報酬が得られそうなのかを含めて判断しなければならない。強化学習ではこの現在の状態の良さを表す指標を「(その状態の) 価値関数」と呼んでいる。

一般的に考えてみよう。各時点  $t$  で行動  $A(t)$  を選び、次の時点  $t+1$  で行動  $A(t+1)$  などと選んでいったとしよう。するとそれに対応して状態  $S(t)$ 、 $S(t+1)$ 、 $S(t+2)$ 、...

と主体は状態の遷移を経験しつつ、それに対応した報酬の系列  $R(t)$ ,  $R(t+1)$ ,  $R(t+2)$ , ... を得ることになる。今状態  $S(0)$  に主体がいるとすれば、目の前の報酬は  $R(1)$  となるし、その先の報酬は一般には  $R(t)$  で表される。時刻  $t$  での価値関数を  $V(t)$  と表せば、状態  $S(0)$  に対応する価値関数は  $V(0)$  になる。

目先の報酬だけでなく、その先の報酬も含めて価値関数で表現するには、まずは単純に価値関数を報酬の和で表す。つまり状態  $S(0)$  に対しては、 $V(0) = R(1) + R(2) + R(3) + \dots$  とすればよい（一般的には、 $V(t) = R(t+1) + R(t+2) + R(t+3) + \dots$  と書ける）。こうしておけば現在の状態の良さ、あるいは価値（関数）を、目先の報酬だけでなく将来の報酬も含めて評価できる。

この式では、実は目先の報酬と将来の報酬を同じものと考えている。例えば  $V(0) = R(1) + R(2) + \dots + R(30)$  の場合、 $R(1) = 100$ ,  $R(30) = 20$ （ほかの  $R(t)$  は全てゼロ）なら  $V(0) = 120$  だが、実は、 $R(1) = 20$ ,  $R(30) = 100$ （ほかの  $R(t)$  は全てゼロ）でも同じ  $V(0) = 120$  になってしまう。

しかし、やはり同じ量の報酬なら、目先の報酬のほうが価値は高いと思えないだろうか。とんでもなく腹が減っているときに、目の前のステーキのほうが一〇日後にもらえるステーキより価値は高いだろう。強化学習ではこの点を考慮して、割引率というのを導入する。この割引率を  $k$  で表そう。この  $k$  はゼロよりは大きく 1 より小さいとする。このとき状態  $S(0)$  に対する価値関数を、 $V(0) = R(1) + k \times R(2) + k \times k \times R(3) + \dots$  のように定義する。こうしておけば割引率によって、同じ量の報酬でも、将来の報酬よりは目の前の報酬のほうが価値があることを反映できる。（一方で  $k=1$  としてやれば、もとの式に戻ることもできる）

この割引率を含めた価値関数を一般的に書くと、

$$V(t) = R(t+1) + kR(t+2) + k^2R(t+3) + \dots = \sum_{n=1}^{\infty} k^{n-1}R(t+n)$$

と書くことができる。この割引率の方法は、指数的割引率と呼ばれる。実は、この割引率で本当によいのかという議論は、現在の研究の最前線ではさまざまにあるのだが、ここではともかく先に進むことにしよう。

### 強化学習の報酬予測誤差

上の式は実は大変都合がよくできている。学校の授業で指数級数の和の公式を習ったことのある読者も多いだろう。その公式と同じようなトリックが使えるのである。それを使うと、となりあう状態同士（つまり一般的には時刻  $t$  と時刻  $t+1$  の状態同士）の間の価値関数について、

$$V(t) = R(t+1) + kV(t+1)$$

という関係式が成立する。この関係式はとなりあう状態同士であれば成立する。割引率を先ほどのように決めておいたことの恩恵である。

さてこの式の左辺は現在の状態の価値を表す。右辺は、行動選択によって得られる目先の報酬  $R(t+1)$  と、将来にわたって得られる報酬の和である。

何か思い出さないだろうか？ ドーパミン神経細胞の活動は報酬予測誤差を表し、その誤差は、ごく簡単に言うと、「予測された報酬」-「実際の報酬」の差に相当すると述べた。実は、このドーパミンの報酬予測誤差は、上の式の両辺の差、つまり、

$$R(t+1) + kV(t+1) - V(t)$$

を表していると考えられるのである。より正確に言うと、

「目先の報酬とその次の状態で見込んでいる報酬の和」-「現在の状態で見込んでいる報酬」の差を取っていると考えられている。

ここで、もし  $V(t) = R(t+1) + kV(t+1)$  の式が成立していたら、 $R(t+1) + kV(t+1) - V(t) = 0$  じゃないかと異議をとねえたくなるだろう。そのとおりである。たしかに、もし価値関数が脳の中で十分学習されていたら、報酬予測誤差はゼロになるのである。そして報酬予測誤差がゼロになっていないということは、まだ価値関数の学習が十分に進んでいないことを示している。そのような場合には学習を進めましょうということなのだ。報酬予測誤差が正の値のときには、現在の状態の価値をもっと大きくしてよい（期待してよい）ということだから、現在の状態の価値  $V(t)$  の値を大きくするように学習することになる。

この価値関数は、各状態でどのような行動を選択するかで、その値が変わってくる。私たちとしては、価値関数を最大化したい。よって、行動選択の学習も同時に行う。価値関数を最大化できるように行動選択の学習も同時に行うことになる。このときにも報酬予測誤差が正に大きければ、よりその行動選択を選ぶように学習するし、負になっていけば、次回はその行動を選択しないように学習していくことになる。このように強化学習では、価値関数と行動選択の学習を同時に進めることになるのだが、その両方に報酬予測誤差がうまく使われているのである。

### 単純はすばらしい

要約すれば、「強化学習」のアルゴリズムを考えていくのに、「状態」「行動」「報酬」の三点セットがあればよい。シンプル極まりない。相対性理論を提唱したことで有名なアルバート・アインシュタインは、かつて、どのような理論を考えるのが良いかと質問されたとき、**Make everything as simple as possible, but not simpler.**（「できる限り単純な理論を作りなさい。でも単純過ぎないように」）と答えたと言われる。

「状態」「行動」「報酬」の三点セットは、人間や動物の行動選択と報酬獲得を考えるのに最も単純化された枠組みだろう。アインシュタインの言う意味でも、強化学習の枠組みから、脳の報酬予測や報酬獲得の機能、そして行動選択や意思決定を考えていくのは実りが

多い。

単純にはじめよ、というのは数理手法としてもうなずける。強化学習で使われる数理を丹念に見ると面白いことがわかる。今は数学的定式化を仔細に議論することはできないが、この強化学習の数理は思いのほか、他分野とその基盤を共有している。

例えば、制御工学、組み合わせ最適化の数理、そして動的計画法がある。あるいはサイバネティクスなどの研究で活躍した数学者リチャード・ベルマンの名前を冠せられた「ベルマンの最適化原理の方程式」とも関係する。後で触れるが、ゲーム理論など経済学で使われる数理とも近い関係にある。強化学習の比較的簡素な問題の定式化が、まさしくそれゆえに、実は深く広くさまざまな数理と関連づけられるというのは良い知らせである。別の言い方をすれば、本質的な要素を数理的に簡素に定式化したからこそ、さまざまな数理と関連づけられるのである。

### 強化学習とロボット

強化学習の単純なところからはじめようという戦略は、けっこう強力だ。賢いロボットを作るときにも一つの羅針盤になる。ソニーのロボット犬「アイボ」のことを多くの読者はご存知だろう。残念なことに、アイボの生産中止が決まったと聞いている（二〇〇七年四月現在）。

人間の呼びかけに反応して可愛い仕草をするアイボは、人気がある。また一部には熱狂的マニアもいる。私はこのアイボの詳細な設計デザインのことは良く知らないので、詳しいことは言えないが、基本は、「入力」→「情報処理」→「出力」である。そして「出力」が環境に変化を引き起こす。環境が新しい状態になるのである。

たとえば、アイボが頭をかく仕草をするというのは「頭をかく」という出力をだしているのである。そしてそれを見た人がその可愛らしさに触発され、アイボの頭をなでたり、あるいは抱き寄せたりする。つまり、その出力が「環境の変化」を引き起こしたことになる。アイボは、その新しい環境の状態を、視覚センサーとか触覚センサーなどを用いて、新たな「入力」として受け、そこから再び情報処理を行い、新たな出力、つまり新しい動作を行うのである。

実は、ブライテンベルクというドイツの神経科学者が一九八〇年代の半ばに、「乗り物」(Vehicles) というタイトルの英語の本を出版している。この本の中核になる部分は、一九六〇年代中ごろには既に書かれていたそうだ！ この Vehicles という本の副題は Experiments in Synthetic Psychology とある。「人工心理学の実験」とでも訳せばよいのだろうか？ 彼がその本に書いているのは、紙の上で行う思考実験である。子供が遊ぶような四つの車輪のある単純な台車をまず最初に用意する。そして、どんなセンサーをつけたら、あるいはどんな出力機構をつけたら、そして最も肝心などんな情報処理をつけたら、その台車が生き物のように、賢く振る舞えるようになるかということ考えた。

ここでも、強化学習的視点、何が得か、報酬か、という見方が大活躍する。面白いのは、

センサーや出力機構や情報処理は最初は非常に単純なところからスタートする。そして賢さを求めるなかで、それらの機構を徐々に複雑にしていくにしたがって台車が少しずつ賢くなっていくところである。賢さの設計、その根底には、何が報酬かあるいは何が心地よいか、何を求める生き物にするかというデザインがある。そしてその求めるもの、報酬に対応させつつ、機構を構築していくことになる。その過程のなかで、強化学習のように、単純なところからはじめる戦略が、読者を見事に誘い込むのだ。

## 5——強化学習による脳研究の進展

### 大脳皮質—大脳基底核回路

報酬予測誤差仮説の提案は、報酬の予測や獲得に関する脳研究に多くの進展をもたらした。その流れの中にある私たちの研究を一つだけ簡単に紹介しておこう。

一九九〇年代後半の話になる。ちょうど報酬予測誤差仮説が提案されたころだ。大脳基底核の回路が、運動制御において重要な役割を果たすことは脳損傷の患者の知見などから良く知られていたが、その情報処理そのものについては、もう一つの重要な回路である小脳にくらべ解明が遅れていた。

このような背景のもと、報酬予測誤差仮説に触発されながら、私たちは大脳基底核回路が運動系列の学習においてどのような情報処理を行うか研究を行った。運動系列の学習というと堅苦しく聞こえるかもしれない。たとえばピアノの初心者が、練習を繰り返すことで或る曲を暗譜するような過程を思い浮かべてもらえばよい。スキルがどのように形成されるかの情報処理研究とも言えよう。

私たちが着目したのは、大脳基底核の解剖学的特徴だった。大脳基底核は、大脳皮質のかなりの領域から入力を受けている。面白いことにその皮質からの入力はいくつかの束に分かれている。この束に分かれていることは脳の解剖学的研究から明らかになってはいたのだが、その情報処理における役割はよくわかっていなかった。

実験研究と強化学習理論を用いた理論研究を併用することで私たちがたどりついた結論は、このいくつかの束になっている大脳皮質—大脳基底核回路では、それぞれの束で学習速度が異なるということだった。速く学ぶ束があれば、遅く学ぶ束がある。どの束も強化学習に基づいて学習する。ここで、速いほうの束は、その時々で新しい系列を学習し、制御に貢献する。一方、遅いほうは、なかなか貢献できない。ところが、学習が遅いぶん、ゆっくりとではあるが確実に学んでいく。ひとたび十分な学習をすると、その記憶はなかなか消えない。遅くてもその特性の生かし方があるのである。そんなことがわかってきた。この私たちの研究結果は、その後のほかの研究によっても概ね支持を受けている。

私たちは他にもこの基底核回路と強化学習、情動がどのように行動に現れてくるかについて研究を進めてきた。残念ながら、今はこれ以上詳しく述べることができない。ただ確

実に言えるのは、強化学習の理論は、私たちの研究も含め、他の多くの研究に刺激を与え続けてきた。今後も大いに必要とされ、さらなる脳研究の発展に貢献するだろう。

### 強化学習のその先は？

先に述べたように強化学習の数理的枠組みの単純さは大きな魅力である。その単純さゆえに、実験結果を対比、検証するのに大変見通しが良い。

ここまで聞くとめでたい話ばかりのように聞こえるが、話はそんな簡単には終わらない。アインシュタインも言っていたように、めでたしめでたしばかりではいけなくて、強化学習の数理の定式化が単純すぎないかという考察も必要である。

結論から言えば、実はちょっと単純すぎるところがあるのだ。いやむしろ正確に言うと、強化学習の枠組みは複雑な脳の計算に分け入るための「入り口」としては十分有用な枠組みなのである。しかし、その枠組みだけで「出口」までたどり着くのは困難そうなのだ。

なぜだろうか？ その手がかりは、「人間」に、そして、強化学習の強みそのものにある。強化学習の理論は、生物にとって最も原初的な報酬の予測や獲得という機能に関する理論という点に大きな強みがある。原初的な機能に関する理論ゆえに、もっと広い意味での報酬の予測や獲得を含めた理論へ発展させるための基礎になりうる。

「広い意味での報酬」を考えるには、「人間」について考えると良い。

人間とはまさしく「人」の「間」と書いて「人間」と読む。「人間」という個人を考えるとき、私たちはほとんど無意識に「他者である人々」も考える。同様に、個人の脳の機能を考えるときにも、ほとんど無意識のうちに他者との関わりにおける脳の機能も考えている。そして、他者との関わりの中で意味をもってくるような報酬も実はけっこうあるのだ。

## 6——人文科学への脳科学の広がり

### 脳と数理と多分野との交流

一般的に言えば、脳科学は、脳の機能解明を目的とした科学である。脳は生物という個体の一つの器官であり、神経細胞は多種多様な細胞群の一部に過ぎない。従って生物学や医学の研究とさまざまな側面で交流があるのは言うまでもない。

同様に心の働きや知能の形成、そして無意識に行われる膨大な計算を理解しようとするとき、脳科学は、人間や動物の行動レベルでその働きを理解しようとする、心理学や行動科学の分野とも絶えず交流してきたし、脳の数理を理解しようとするとき、数学、物理学、情報統計科学などの多様な数理科学の分野とも交流してきた。

人間の行動を理解しようとする欲求は何もこれらの学問に限ったことではない。いわゆる人文社会の学問、例えば哲学、文学、政治学、社会学なども似たような興味を持つ。特に人々のお互いへの関わりに関する行動を理解しようとする欲求は、これらの学問のほう

がより強かったりする。そして忘れてはいけないのは経済学である。

いわゆる人文社会の学問の中で、特に経済学は人間の意思決定に強い関心を持つ学問である。相対的に言えば、経済学は心理学に比べると人間同士の相互作用により強い関心があったといえよう。市場などでお互いに売り買いをするような経済活動を考えれば、それも納得がいく。しかし、その経済活動を行っているのは人間であり、脳活動なのだ。

### 脳科学と経済学が結婚すると神経経済学

経済学と脳科学はその発想と関心を実はかなり共有している。実際、近年、脳科学と経済学が結婚をしたかのように、神経経済学という研究分野が勃興しつつある。もっと広く言えば、経済学を含め、社会的機能を問いかける学問が脳科学と合流しつつあり、それは社会脳科学と総称されている。

神経経済学の研究分野が現れるまでには、脳科学と経済学のそれぞれで育まれた伏線があった。脳科学の側からの伏線の一つは、fMRI（機能的核磁気共鳴イメージング）などに代表される人間の脳活動全体を計測する技術の発展がある。このような計測技術によって、人間が意思決定や行動選択をしているときに、脳のどの部分の活動が特に大きいかを観測することが可能になった。以前であれば、心理学や経済学で実験をしてもデータとして取れるのは行動に表れる結果だけだった。

しかし、現在は、その行動選択の脳の活動の変化、つまりその「思考」過程を観測することが可能になったのである。これは大きな進歩である（ただし、これは格段の技術の進歩ではあるが限界もある。例えば現時点のfMRIの時間分解能や空間分解能は、情報処理の詳細な過程を観測できるレベルではなく、あくまで目安程度に観測できるに過ぎない）。

経済学からの伏線の一つは、行動経済学、実験経済学などと呼ばれる研究の蓄積である。行動選択や意思決定の研究で、経済学は市場などの集団的過程での意思決定やそこでの合理性により相対的に関心が強い（あくまで相対的な話ではある）。その意味では、個人の情報処理過程に重きをおいた研究は、少なくとも今までは、経済学の中では比較的少数派と言ってよい。そんな経済学の中で、行動経済学の研究は、人間が実際の場面でどのように意思決定するかをつぶさに検討する研究を推進してきた。

その種の研究に関して記憶に新しいのは、ダニエル・カーネマンが二〇〇二年にノーベル経済学賞を受けたことである。カーネマンはエイモス・タバスキと並んで行動経済学の開拓者である。

もっと時間を遡るとハーバード・サイモンという人物に行き当たる。彼は、一九七八年にノーベル経済学賞を受けた。実は、サイモンは、一九五〇年代半ばから本格的にはじまった認知科学そして人工知能の両方の分野で、その初期から活躍した研究者でもある。たとえば、数学の定理を証明する思考過程を計算機上に実現させることを目指した研究や、チェスを行っているときの人間の眼球運動の動きを計測することで、思考過程が視覚情報をどのように生かしているかなどの研究も行った。

サイモンの視点からは、もともと経済学と脳科学は同じ研究領域と言えるのだろう。しかし、少なくとも過去数十年、それほど交流が盛んだったとは言えない。しかし、この神経経済学という言葉に代表されるように、今、私たち研究者の意識のなかではこの両分野は急速に接近しつつある。

### 経済学の快樂という視点

脳科学と経済学を接近させているもう一つの大きな要因は、数理的視点から脳機能を理解しようという気運の高まりだ。その点、経済学と数理神経科学は相性が良い。経済学はもともと数理的に人間の判断や意思決定を評価しようとする気運を強く持っていた（ちなみに、私は文系・理系という区別は好きではない。数理神経科学は、脳の計算を理解するために数理を使って行う哲学であり心理学と考えても良く、文系・理系で区別できるわけではない）。特に強化学習に代表される報酬の予測と獲得という概念は、経済的利得最大化の概念とも親近性がある。

経済学でも、上述したような「快樂」や「報酬」が大活躍する。もう少し正確に言おう。経済活動における人間の行動選択や意思決定を定量的に評価／検証することは、経済学の大きな目的である。この目的のもと、「経済人」という概念がしばしば用いられる。これは、人間は経済的利得を最大化するように振る舞う、その意味で人間の行動は合理的であるという考え方に基づいている。そのように振る舞う人間を「経済人」と呼ぶ。

経済的利得の最大化、合理的行動、枝葉末節を切り捨ててしまえば、その根源的な意味は、数理的に人間の行動選択を定式化したとき、ある変数の最大化を行っている、そしてその定式化に基づいた振る舞いをするを指している。ひらたく言えば、どれだけ自分の利益を最大化できるかを考えて、最も得をする行動を選択するということになる。

つまり、経済学もそこで用いられる定式化の中で「快樂」や「報酬」を考えているのである。利益の最大化を考えるための指標として、先ほど述べた強化学習で使われる「価値関数」という概念と非常に似たような関数を考えるのである（正確にはこれらは同じではないのだが、今は細かな違いは無視しよう）。経済学の中でも実はいくつかの異なる定式化があつて、それぞれで「快樂」に対応するような変数がある。定式化が異なれば、その「快樂」が果たす役割は異なってくる（これは脳科学における「快樂」を考えるときにも同様に注意しなければいけない）。しかし、いずれにしても「快樂」という発想自体は、経済学でも脳科学でも共通している。

### 快樂、報酬の視点のさらなる広がり

広い意味での報酬、他者とのかかわりの中で考えられる報酬というのは何があるのだろうか？ 本能的欲求としては、例えば、一定程度の性行動もその一部と考えられる。それではそれら性行動あるいは性衝動が満たされるのは報酬獲得と同じだろうか。本能的欲求以外でもさまざまなものが報酬になりうる。簡単なものとしては、金銭的報酬あるいは、

経済的報酬があげられる。

人間は社会をつくる。その意味で社会的動物である。社会で暮らす中で、主従関係や、相手に対して強いあるいは弱い立場という社会的関係はしばしば起こる。権力欲という言葉があるが、相手に対して優位にたつことは報酬か、権力の行使は報酬としての機能があるのか。

社会の中では、人はお金や権力のみで動くわけではない。たとえば仲間からの信頼とか身近な関係も大切だし、社会的貢献に生きがいを見出すこともある。社会的名誉、社会的信用を得るなどは報酬獲得と同じだろうか。道徳（モラル）は、たとえば老人をいたわるとか幼き子供たちを守るとか社会の成員たちに共通に抱かれることで、人々の行動選択の規範となる。この行動選択に影響を与える道徳は、報酬とどのようにつながるのか。道徳的に振る舞うことは報酬たりうるのか。公平に振る舞うことや正義を貫くことは、報酬獲得と同じだろうか。

人類の歴史の中で無数の詩や小説や歌が、人々の愛について語ってきた。人を愛することは、報酬になりうるのだろうか？ それとも愛されることが報酬になるのか。与える喜びと与えられる喜びとは何が違うのか、それともこの二つの喜びは報酬としては同じことなのか。

このように何が報酬かとちょっと考えてみても、疑問がぞくぞくと湧いてくる。愛と報酬が同じものかなんて問うのは、罰当たりすぎるだろうか。私はそうは思わない。愛と報酬で何が同じで、何が違うのか、それを問うことで私たちの心の理解がより深まると思うし、脳機能の理解、脳の計算を明らかにできるようになると考えている。

### 情報処理という視点から

何が同じか違うかという疑問は、情報処理という観点からは、それぞれの情報処理のプロセスの中でどの部分が共通で、どの部分が違うかという形で整理される。数理的な式で書き下せればその比較はより明確になる。

数理あるいは数式は科学的研究において大変強固なコミュニケーションの武器だ。定式化を通じて、経済学であれ、脳科学であれ、心理学であれ、お互いにコミュニケーションが可能になる。名誉や愛とかが報酬獲得とどのように情報処理として似ているか、何が違うのか問いかけるのは自然なのである。食べ物、お金、名誉、愛と一見まったく異なって見える事柄も報酬という視点から統一的に考えると同時に、これら異なる報酬の情報処理の過程の中でどの変数が共通に利用されるか、どの変数が共通でないか、そして、どの情報処理の数式が共通に使われ、どの数式が共通でないか、という形で議論を整理することが可能になる。

このような脳の計算を解明するためのヒントを与えつつあるのが、神経経済学であり、また近年急速に発展している情動や社会性に関する脳研究である（しばしば社会脳科学と総称される）。

## 愛情も脳の計算

愛のある人生は、愛のない人生と何が違うのか？ ロマンティックな愛を抱いている人は、世界が違って見えるのか？ 愛する人を見ると、彼や彼女の脳活動はどうなっているのだろうか。

fMRIを使って、ロマンスの中にある人が、その愛する人の写真を見たとき、他人の写真を見たときに比べて脳のどの部位で活動が顕著に見られるかを調べた実験がある。これによると大脳基底核回路の線条体の一部（尾状核）や、ドーパミン神経細胞が豊富にある腹側被蓋野での活動が特に高くなっていたのである。

この結果は、ロマンスあるいは愛しているといういわばとても情緒的な感情も、情報処理としては、報酬予測や報酬獲得などの機能の情報処理と少なくとも部分的には同じプロセスを利用していることを示す。

さらに面白いことがある。同じ愛するといっても、燃え上がるような愛と、もっと長期的に確かに根付きながら抱き続ける愛、この二つの愛のありようは同じだろうか。私たちの日常の経験から考えると、この二つの愛は似ているところがあるにしても、まったく同じものには思えない。

同様な実験から、実は、この長期的な愛の場合、上述した尾状核と腹側被蓋野に加えて、大脳皮質の中で島皮質と呼ばれる部位と前頭葉内壁にある帯状回前部が活動していたのだ。脳活動から眺めると、どうやらこの二つの愛には違いがあるらしいのだ。

この島皮質は今までの脳研究から、身体の状態に関する脳活動があることが知られている。身体の状態の中でも、特に痛みや空腹や臭いなどのネガティブな身体的反応に関連が深いことが知られている。また、嫌悪の感情に関して、大脳基底核回路と共に島皮質が関係しているようなのだ。ちなみに嫌悪に関しては、恐怖の感情に関連の深い扁桃体も関係する。この扁桃体も基底核回路とは関連が深い。

一方、帯状回も興味深い脳部位で、自らの行動の結果を判断することやそれに伴う行動の選択の変更などに関係しており、帯状回の機能もドーパミン神経細胞の報酬予測誤差の反応と関連が深い。

私たちが日常レベルでやや違ったように感じる二つの愛の形態は、やはり脳活動として違った活動をともなっている。この違った愛の形態では、報酬獲得や報酬予測に関連の深い脳部位の活動が違った形で組み合わせられている。

この二つの愛の形態では、お互いのパートナーの振る舞いについて「報酬」（うれしいという気持ちに相当するのだろうか）と覚えることが違ってくるのかもしれない。そして相手を喜ばせようという行為や相手への期待、つまり報酬を与える行為や報酬の予測も違ってくるのかもしれない。その違いは、異なる脳部位の組み合わせで情報処理が行われていることから生まれてくるはずだ。

また、二つの愛は、一方があれば他方がないという排他的関係というよりも、共存しな

がら徐々に変化していくような関係にある。そのような「気分」や「感情」の変化は、脳内の情報処理の変化に基づくはずである。

愛情も脳の計算なのである。愛は私たちの気づかないところで計算高いのである。脳の機能を理解したい、脳の報酬獲得の機能を理解したい、脳を創りたいというのは、脳の愛情の計算をも理解したいということなのだ。

## 7——社会的判断と社会的人格

### 囚人のジレンマ

経済学では人間同士の相互作用や社会的環境が、本人の意思決定にどのような影響を与えるかについてさまざまな議論が行われてきた。例えば、囚人のジレンマというゲームは良く知られた話だろう。それ以外にも多くの実験に使われるゲームがある。例えば最終提案ゲームと呼ばれるゲームでは、社会的判断のもとでは、単なる自分の利得だけでなく、公平性、配慮、感情が含まれて意思決定が行われることが示されたりする。

ここでは、更に別のゲームで、信用ゲームと呼ばれるゲームを取り上げよう。最初に二人（AとBと呼ぶことにする）はいくらか渡される（ここでは四〇〇円としよう）。まずAは、自分の持ち分四〇〇円からBに贈与する金額を宣言する（X円としよう）。このときAの金額は $(400-X)$ 円となる。このX円はBに渡される前に、何倍かされる（三倍としよう）ことになっていて、Bの金額は、 $(400+3X)$ 円となる。最後に、Bはこの自分の手持ちのお金からAに贈る金額を決める（Y円としよう）。これで1回分のゲームが終わり、最終的にAは $(400-X+Y)$ 円、Bは $(400+3X-Y)$ 円になる。

このゲームの妙味は、Aが最初に贈与する金額が多ければ多いほど、AとB二人合わせた手持ちのお金は増えることになる一方で、最後にBが十分にAに贈り返しをしなければ、Aは最初の贈与額を多くしても何も得をしないというところにある。

あなたがAならどのように振る舞うだろうか？ 自分の利得を最大にするにはどうしたらいいのだろうか？ もしBのことを何も知らなければ、実際にこのゲームを繰り返しやりながら、Bがどのように振る舞うかを見て、どれくらいの金額を贈与すればいいか調整していくしかないだろう。

この実験でAの脳活動を見てみると、はじめにBの振る舞いを判断している段階では線条体（特に尾状核）の活動が高いことが示されている。対人関係における信用を含めた報酬予測でも、線条体が関連していることがわかる。さらに、一般的に強化学習の理論ではこの線条体（そして側坐核）が報酬予測や自らの報酬予測の正しさを評価する主な部位とされているので、この結果は理論とよく合致している。Bの振る舞いを見ながら信用の度合いとそれに基づいた報酬の予測に関する学習に関与しているのだろう。

ところで、だ。日常の生活を考えると、そもそも相手（ここではB）が信頼できる人か

どうか知っているし、あるいは先入観があったりするし、あるいは周囲からその評判を聞いていたりする。そんなときにはどうするだろうか、脳活動はどうなっているのだろうか？

実験では、Bとして三通りの人物像が準備された。特に情報が与えられていない中立のパートナーつまり「普通の人」、評判の「良い人」、そして「悪い人」である。面白いことに、この「良い人」と「悪い人」と同じ課題を行ったときには尾状核の活動はそれほど高くならなかったのである。報酬予測そして報酬予測の評価の主要部位の一つである尾状核の活動が、予め与えられた相手の評判によって活動が変化したのである。

そもそも報酬の予測やその評価をする部位の活動は、その活動を通じて、後々に行動選択をより優れたもの、すなわちより大きな報酬を得られるように脳の情報処理を変化させていく。つまり脳における学習を進める役割がある。尾状核の活動が評判によってその活動を変化させるということは、報酬予測の調整（学習）や行動選択の学習においても、相手の評判を考慮にいれてその学習を進めることを示唆している。また、相手の評判の程度によっては、脳の他の部位での活動や学習がより重要に効いてくるかもしれない。

他人に対する評価や信頼という社会的機能も、もちろん脳の計算に基づいている。そして、その脳の計算においては、原初的な報酬獲得の部位が深く関与しているのである。

### 神経修飾物質の働き

報酬予測について述べるとき、今までドーパミンにハイライトを当ててきた。実は、神経細胞間の情報のやりとりには、さまざまな物質が使われている。

たとえばグルタミン酸と言われる物質が、シナプス前にある神経細胞から放出されると、シナプス後の神経細胞の活動を興奮させるように働く。一方、GABAと呼ばれる抑制性の神経伝達物質を放出する神経細胞もある。この場合には、シナプス後の神経細胞の活動を抑制するように働く。このようにシナプス間の伝達に使われる物質の中で、神経細胞間の情報のやりとりに直接働くような物質は、神経伝達物質と呼ばれている。

シナプス間でやりとりされる物質は他にも多数ある。ドーパミンもそのひとつである。ドーパミン神経細胞はドーパミンを放出することで、情報をやりとりするシナプスの結線（重み）を変化させるように働くことは先に述べた。このドーパミンのように、直接情報のやりとりをするよりもむしろ情報のやりとりを修飾するように働く物質は、神経修飾物質と呼ばれる。

この神経修飾物質には他にも多くの種類があり、これらの物質のうちいくつかは心の働きとの関係が議論されている。

セロトニンと呼ばれる物質は、神経修飾物質の一種であって、うつ病との関連が議論されたりしている。ドーパミンは報酬予測や報酬獲得と関連が深いことを述べたが、実はドーパミンもいくつかの精神疾患、たとえば注意欠陥多動性障害（ADHD）の関連が議論されたりしている。さらには、ドーパミンとセロトニンはまったく別々に脳活動に影響を与えるわけではなく、お互いに関連しながら脳活動に影響を与えることもある。

## オキシトシンは人を信用させる？

オキシトシンというホルモンがある。このホルモンは女性が赤ちゃんを産むとき、その分娩時に子宮の収縮を誘発する、あるいはお母さんのお乳を出すのを助けることで知られている。ホルモンとは、主に血中への分泌が行われる物質のことである。

実はこのオキシトシンは脳内にもあって、神経修飾物質の一種としても働く。オキシトシン神経細胞は、脳内ではたとえば視床下部と呼ばれる部位に存在する。視床下部のオキシトシン神経細胞は、その出力をたとえば側坐核に送っている。視床下部は、大脳基底核やドーパミンが存在する黒質や腹側被蓋野と大変密接な関係にある。

面白いことに、先ほど述べた信用ゲームで、このオキシトシンを人間に嗅がせると相手を信用しやすくなるという実験結果がある。あのゲームで、Aの役割をする人がオキシトシンを嗅いだ後ではBのことをより信用して、より大きなお金をBに贈ろうとすることが示されている。

実は、このオキシトシンの脳における働きは、ドーパミンの機能の働きと関連が深いこともわかっている。オキシトシンとドーパミンが多く出力を送っている脳の部位（たとえば側坐核）は重なるし、またお互いの機能に影響を及ぼすことがわかっている。

しかも、話は飛ぶが、ハタネズミを使った面白い研究がある。ハタネズミにも色々種類があるのだが、或る種では、一夫一婦制が普通なのに、他の別の種では、一夫多妻制が普通だったりする。そして、一夫一婦制を好むハタネズミのメスでは、どうやら脳内で、特に大脳基底核（特に線条体と側坐核）で、オキシトシンの影響力が強いらしいのだ。実際、オキシトシンを脳内注入してやると、自分のパートナーとより長い時間を過ごす、つまり、より一夫一婦制を好むメスの数が増えるのである。

それに対し、一夫多妻制を好むハタネズミの種のメスでは、オキシトシンの影響力が弱いようなのだ。つまりオキシトシンの脳内での影響力の強弱と、一夫一婦制を好むかあるいは一夫多妻制を好むかについては強い相関がある。

このように私たちの「身体」をかけめぐり身体の機能を調節するように働く物質（ホルモン）が、脳内の報酬予測や報酬獲得に関わる機能と密接に関連している。その物質、つまり「もの」は、私たちの心の働きに影響を与え、他人を信用するという社会的行動に関連するし、少なくともハタネズミではパートナーの選び方やその絆の形成にも直接効いてくる。

一方、私たちの行動や振る舞いは、たとえばどの報酬を獲得するかという違いを通じてドーパミンやオキシトシンの分泌の仕方を変えることができる。行動は、これらの「もの」の量を変えることで心の働き、つまり脳の情報処理を変えうる。落ち込んだときに元気に振る舞うことで精神的に立ち直れるんだよ、といったアドバイスも、実はこのような脳の情報処理の機構の上に基づいているのかもしれない。

脳の中に存在する「もの」が心の働き、脳の情報処理を変える。情報処理の視点からは、

オキシトシンがドーパミンとの相互作用の中からどのような変数に作用するのか、それがどのように学習を、つまり“良い方向”を作り上げるのに効いてくるのかが問われることになる。

## 8——これから進む道

### 数理の眼を養う

愛情や信頼などの社会的な機能が、食べ物などの原初的な報酬の予測や獲得の機能と、脳活動としては実はかなりの重複があることを見てきた。面白いことに、これらの機能ではドーパミンという「もの」が大きな役割を果たしている。強化学習は、社会的機能も含め広い意味での報酬の獲得や予測の情報処理の理解を深めていくことに役立つ。そしてその情報処理の理解は、私たちの心や精神や知能の物質的あるいは生物的基盤を明らかにすることにつながっていくのである。

脳の数理研究をますます発展させることが大切だ。実は、この章で述べた強化学習の方式は、学習理論としては一つの定式化にすぎない。広い意味での報酬獲得の情報処理を考えるには、二つの作業が必要不可欠なのだ。一つには、その機能が脳でどのように実現されるのかを知るための実験的研究である。もう一つは、それらを統合的に表現できる数理そのものを開発することである。「数理の眼」が大切なのだ。

これは、報酬獲得や強化学習だけではない。他の脳の隠れたデザイン、その快樂の作りかたを理解していくのにも「数理の眼」が重要だ。多数の神経細胞からなる個々の特徴をもった神経回路の計算能力や学習能力を明らかにしていくことが大切である。

これらの研究はいずれもとてもチャレンジングだ。だからこそ面白い。何せ脳を創る研究なのだから。いつか脳が行っている愛の計算を明らかにする数理理論も作れるだろう。夢は壮大に、そして足元では着実な一歩をたゆまなく進めていくのが大切だ。

### 楽しみはこれからだ

私たちの生活や認識も脳科学の理解が深まると共に変化していくことだろう。たとえば、私たちひとりひとりの人間の多様な個性も、根本的には、脳の情報処理における違いだろう。人間の脳としての基本的な情報処理は同じでも、各処理でのアクセントのような違いが、個性の違いになるのだろう。血液型の性格分類よりは、脳活動の個性に基づいて、もっと確かなことが占えるときがくるかもしれない。いつか、人生のアドバイスも脳の言葉に書き換わるのだろうか？ 「孔子の教え」のいくつかは脳科学の言葉におきかえられるのかもしれない。脳科学の興奮や楽しみはこれからさらに花開く。

科学は限りなく面白いから人類は喜んでやっている。脳科学もその一つだ。「科学だから」と気難しい顔をしていてもその面白さはなかなか伝わらない。この章では、難しい顔はや

めて、散策気分で語ることで、研究の面白さが伝わるようにと努めた。

そろそろこの章も終わりだ。理論神経科学のトレジャーハンティング、楽しんでもらえたなら嬉しい。つまらなかったら、すみません。駆け足で通り抜けたので、少し物足りなかった人もいるかもしれない。いつかもう少し長旅をするような著作を書いてみたい。

さあ、この本を閉じて、目をつぶって深呼吸してください。「快樂が脳を創る」と心の中でつぶやいてください。そのイメージは豊かに広がってくれていますか？

Figures



図 11-1 インターネット・地球・脳

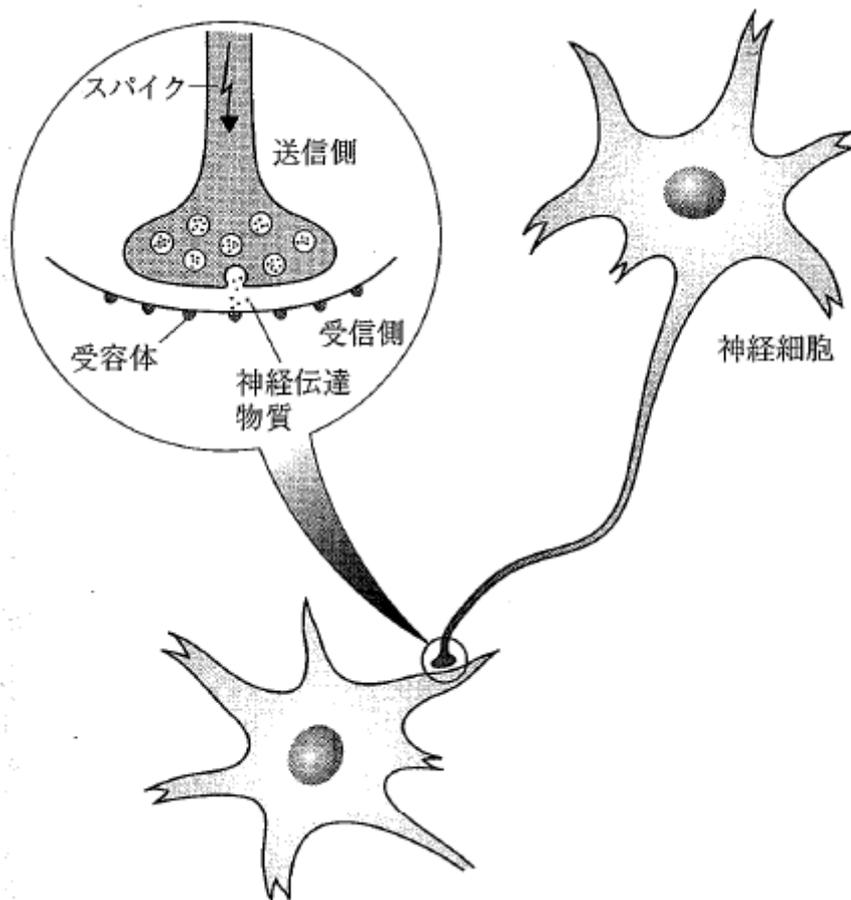


図 11-2 シナプス

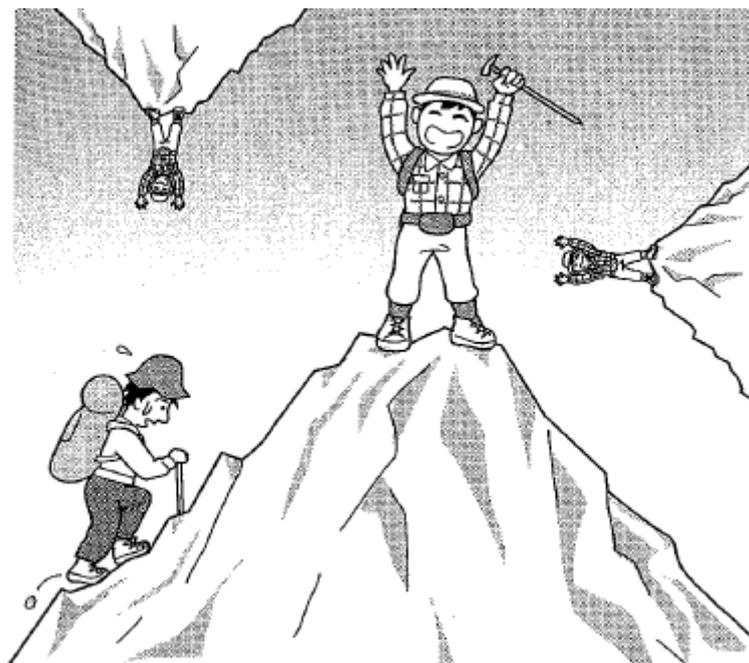


図 11-3 脳は「良い方向」を好む

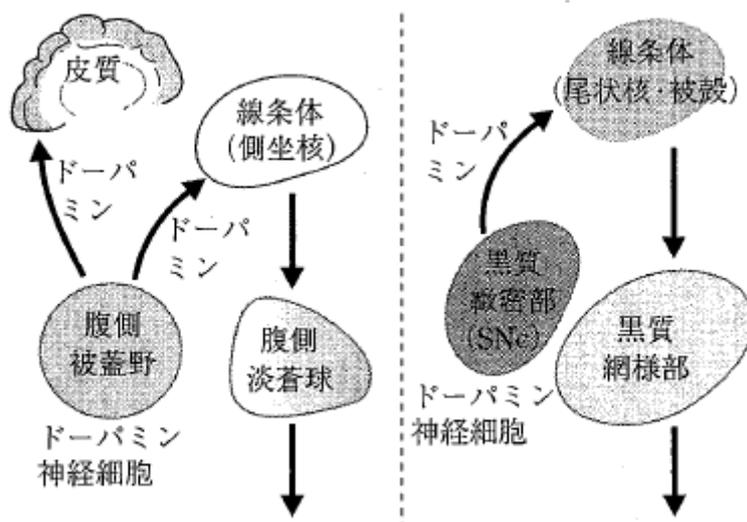


図 11-4 快樂とドーパミン細胞

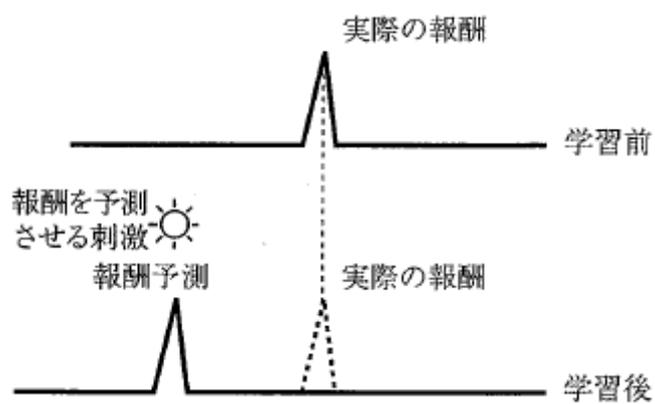


図 11-5 報酬とドーパミン神経細胞の反応

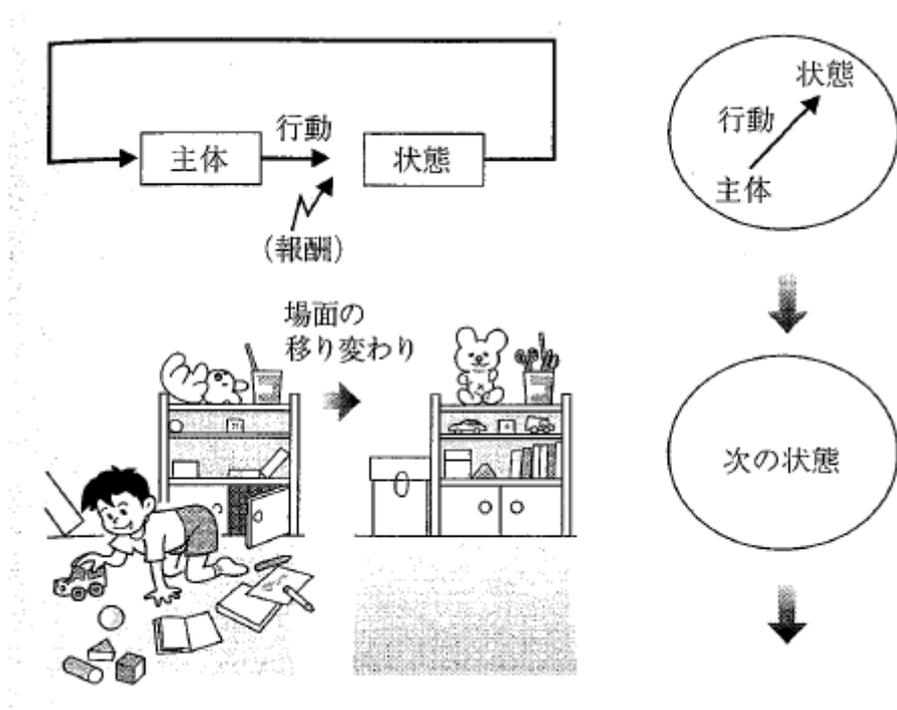


図 11-6 状態・報酬・行動の関係

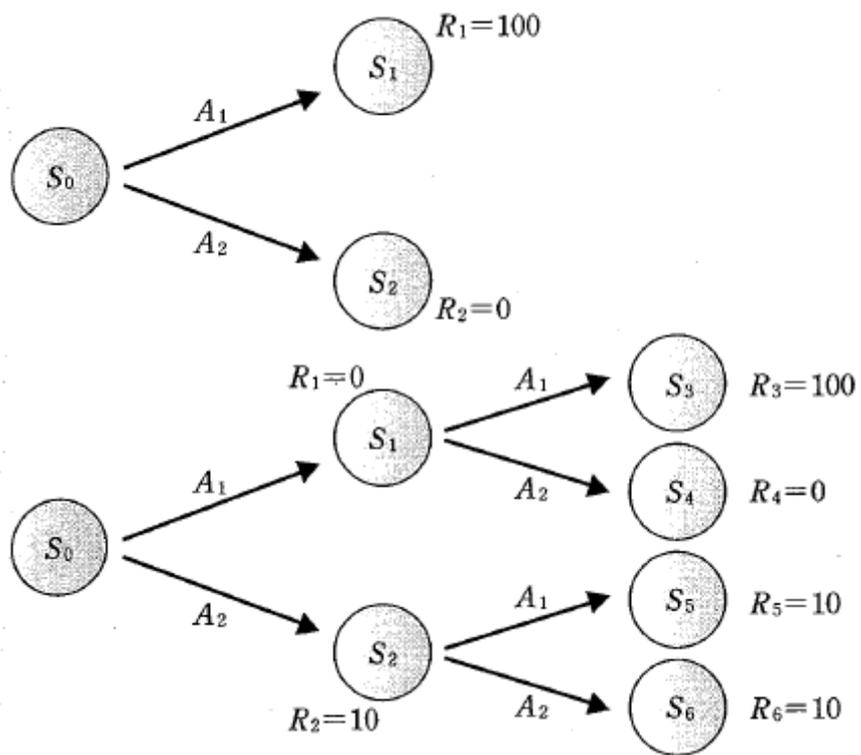


図 11-7 強化学習アルゴリズム