

■S3 群 (脳・知能・人間) —脳と神経モデル

---

## 4 章 脳のシステム

■概要■

【本章の構成】

## ■脳と神経モデル—4章

### 4-1 知覚認識システム

(執筆著者：藤田一郎) [2009年3月 受領]

本節では、知覚認識に関わる脳内システムの一例として、視覚システムを概観する。

#### 4-1-1 網膜

網膜では、視細胞が光子をとらえ膜電位応答を起こす。視細胞の応答は、双極細胞を経て、神経節細胞へと伝えられる。網膜の出力は、神経節細胞の軸索（視神経）を伝わる活動電位として脳に送られる。

神経節細胞には異なる種類があり、視覚情報の異なる側面を伝える。まず、受容野の構成からオン中心—オフ周辺型（オン中心型）とオフ中心—オン周辺型（オフ中心型）の2基本型に分けられる。オン中心型では、受容野中心に光を当てると興奮し、周辺部に光を当てると抑制される。オフ中心型では、中心部を照射すると抑制され、周辺部の刺激で興奮が起こる。これらの細胞はそれぞれが更にX細胞とY細胞に大別される。X細胞の反応は持続的で、静止刺激によく応じる。Y細胞の反応は一過性で、動く刺激によく応じる。受容野の位置が同じ場合には、Y細胞はX細胞より大きな受容野をもつ。軸索伝導速度はY細胞の方が速い。X細胞は、網膜上でのその細胞の周辺領域とその外側との明暗の差を抽出するのに適し、Y細胞は明るさの時間変化を抽出するのに適している。網膜の出力は、視細胞が受ける光の強さの絶対値情報ではなく、明暗の空間的・時間的変化の情報である。

サルの網膜では、X細胞は形態学的にミジェット細胞（ $P\alpha$ 細胞：中程度の大きさの細胞体と小さな樹状突起野）に対応し、Y細胞はパラソル細胞（ $P\beta$ 細胞：大きな細胞体と大きな樹状突起野）に対応する。前者では、受容野の中心部と周辺部で光波長に対する感受性が異なる。例えば、中心部は長波長光で興奮し、周辺部は短波長光で抑制される（色対立型反応型）。後者の受容野の中心部と周辺部は広い分光感度を持ち、広帯域型と呼ばれる。

#### 4-1-2 外側膝状体

神経節細胞のほとんどは外側膝状体（LGN）に投射する。LGNは6つの層からなる。腹側の1, 2層は大細胞層（M層）と呼ばれ、 $P\alpha$ 細胞から入力を受け、背側の3~6層は小細胞層（P層）と呼ばれ $P\beta$ 細胞から入力を受ける。LGN細胞もまた同心円状の受容野を持ち、各層の中で、細胞は受容野の位置に対応して並び、視野再現構造を形成する。

P層細胞は、小さな受容野、持続性応答、色対立型の受容野をもつ。M層細胞は、大きな受容野（小細胞の2~3倍）、一過性応答、広帯域型の受容野をもつ。M層細胞の方がP層細胞より明暗コントラストに敏感である。P層の破壊で色、細かい図形パターン、微妙な奥行きを検知が阻害され、M層の破壊では刺激の速い点滅や動きの検出に障害が起きる。

#### 4-1-3 大脳皮質視覚路の構成

マカカ属サルの大脳皮質の半分が視覚に関与しており、その中に30を越える領域がある。視覚経路の後半部は、側頭葉に向かう腹側経路と頭頂葉に向かう背側経路の二つに分かれる。一次視覚野（V1）から側頭葉にいたる腹側経路は「見ているものが何であるか」を知るため

に必要な情報、例えば形や色を主に処理している、一方、V1 から頭頂葉にいたる背側経路は「見ているものの位置や動き」を知るのに重要な情報、例えば対象の動きの方向が処理されている。また背側経路は、自分の手や指を握る対象に合わせて形作るのに必要な情報を担っている。両眼視差の情報は両方の経路で処理されている。

#### 4-1-4 一次視覚野

霊長類の LGN の出力のほとんどは、V1 に到達する。V1 では、LGN で分別されていた情報（右眼由来か左眼由来か、大細胞由来か小細胞由来か、オン中心型細胞かオフ中心型細胞かなど）が統合され、処理後の情報を V2, V3, V4, MT などに分配する。

##### (1) 層構造と視野再現

大脳新皮質は表面に平行な六つの層から成る。LGN の M 層から V1 への入力は 4C $\alpha$ 層に、P 層からの入力は 4C $\beta$ 層に入る。4C $\alpha$ 層は 2, 3 層及び 4B 層に、4C $\beta$ 層は 2, 3 層に出力する。2, 3 層は 5 層や他の皮質へ、4B 層も他の皮質へ出力する。5 層は上丘や視床枕と 6 層へ投射する。6 層は LGN からの直接の入力も少ないが受け、また LGN へ投射する。

V1 細胞は反対側視野に受容野をもつ。受容野の大きさは、視野中心に受容野をもつ 2, 3 層の細胞で、0.2~0.5 度である。V1 の皮質に垂直方向に並ぶ細胞は受容野が重なりあい、横に 1 ミリ離れると重なりがなくなる。つまり、1×1×厚さ 2 ミリ立方の中で視野の中の 1 区画の情報処理がなされている。皮質表面に平行の方向に正確な視野再現が見られる。

##### (2) 方位選択性

4C 層以外の層の V1 細胞の多くは、特定の方位に傾いた細長い光刺激に選択的に反応する（方位選択性）。方位選択性細胞には、明刺激に反応する受容野領域と暗刺激に反応する領域が分離している単純型細胞と、分離しない複雑型細胞がある。

##### (3) コラム構造

個々の LGN 軸索は片眼からの情報を伝え、右眼由来、左眼由来のどちらの情報を伝えるかによって、4C 層の中で幅 0.5 ミリの帯となって互い違いに分布し、4C 層細胞もまた右眼か左眼のどちらかへの刺激にのみ応答を示す。したがって、4C 層細胞もどちらの眼由来の視覚情報を扱うかによって 0.5 ミリ幅をもって分布しわけている。4C 層細胞は、次に自分の上下の 2, 3, 5, 6 層へ軸索を送る。その際、多少横方向にも軸索投射が広がるため、2, 3, 5, 6 層の細胞は、両眼性の入力を受ける。しかし、それでも、これらの細胞はどちらかの眼からの入力をより強く受け、4C 層に見られた眼優位性の帯は皮質の全層を貫く。この構造を眼優位性コラムと呼ぶ。方位選択性細胞も、似た方位に応じる細胞が、皮質の表面から底にいたる方向に並び、方位選択性コラムを形成する。大脳皮質表面に平行な方向に沿って、方位選択性コラムの好む線分の傾きが徐々に変化していく。

Hubel と Wiesel は、V1 の 1 方向には、右眼優位、左眼優位の帯状のコラムが交互に現われ、それに直交する方向に、異なった方位に選択性をもつコラムが順々に並んでいると考えた。実際には、方位選択性コラムは皮質表面から見ると車輪のように放射状に構成されている。また、2, 3 層の中に、方位選択性の弱い細胞が塊となっている構造がある。この塊は、チト

クローム酸化酵素 (CO) を豊富に含み CO プロブと呼ばれる。

#### 4-1-5 V2 野

V2 は、V1 の前方を帯状に広がる。

##### (1) CO 縞構造、大細胞経路、小細胞経路

V2 における CO の分布は縞状になり、CO 縞構造と呼ぶ。「太く濃く染まる縞」と「細く濃く染まる縞」が「淡い縞」を挟みながら、互い違いに現われる。

Livingstone と Hubel (1987) は、霊長類の初期視覚経路には二つの流れがあり、それぞれが視覚の異なった側面を取り扱っていると主張した。一つの経路は、網膜の P $\alpha$ 細胞に始まり、LGN の M 層→V1 の 4C $\alpha$ 層→4B 層→V2 の太い縞へいたる経路 (M 経路) で、もう一方は、網膜の P $\beta$ 細胞→LGN の P 層→V1 の 4C $\beta$ 層→2・3 層→V2 の細い縞にいたる経路 (P 経路) である。この 2 分法は過度に単純化されており、M 経路と P 経路の分離は V1 の 4C 層までは保存されているが、その後、V1 の 2, 3 層で再結合される。そして、その後の情報処理を経て、別セットの並列経路、すなわち腹側経路と背側経路となっていく。腹側経路、背側経路とも、比重は違うにせよ、M, P 両経路からの入力を受ける。

##### (2) V2 細胞の刺激選択性

V2 細胞は、V1 細胞と比べて、受容野が大きく、低めの空間周波数をもつ刺激を好む。

カニッザの三角形とよばれる錯視図形では、視覚刺激としての三角形はないのに三角形があるように見える。三角形を構成する「辺」の部分には物理的に明暗コントラストはないにもかかわらず、輪郭が知覚される。これを主観的輪郭と呼ぶ。V2 の方位選択性細胞の 1/3 が主観的輪郭に応答する。

#### 4-1-6 MT 野

V1, V2 の主な出力先の一つは MT 野である。MT 細胞は反対側視野に受容野をもち、受容野の位置に従い配列し、視野地図を形成している。他の多くの領域と同様、受容野は中心視付近で小さく、周辺ほど大きい。

MT 細胞の 90%以上が、視覚刺激の運動の方向軸と向きに選択性をもっている。大多数の細胞の反応は、視覚刺激のコントラストの方向や色や形には依存せず、運動方向がその細胞の最適方向でさえあれば、強く反応する。MT 細胞は、その最適刺激の運動方向と向きに従いコラム構造をとる。V1 野が視野地図の中に刺激の傾きと眼優位性の地図を埋め込んでいくように、MT 野は、視野地図の中に運動方向の軸と向きの地図を内蔵している。

受容野内に最適運動刺激を与えると同時に、受容野の外に背景となるパターンを与え刺激と同じ方向と速度で動かすと、半数の MT 野細胞において、受容野刺激に対する反応が抑制される。この周辺抑制は、背景パターンの運動方向と速度が受容野内の刺激に近いほど強く、受容野内の刺激からずれるにつれ弱くなる。これらのことは、MT 野において、刺激と背景との相対運動の検出のための情報が取り出されていることを示している。

#### 4-1-7 MST 野

MST 野の投射先である MST 野は、純粋な視覚情報処理のみならず、視覚情報を利用して行う眼球や手指の運動の形成に深くかかわっている。

##### (1) 物体の動きと自分の動き

眼前を鳥が横切ったときは、視野の中のごく一部が動くに過ぎないが、自分の体が動いた時には視野全体が動く。このような広視野の動きを受容野の小さい細胞はコードすることはできない。MST 細胞は、視野全体をカバーするような大きな受容野（平均直径 40 度）を持っており、MST 野背側部の細胞は、広視野パターンのような様々な動き（直線、拡大、縮小・回転）に反応する。これらの細胞は、自分の体の動きを視覚的に知るメカニズムを担う。

視覚刺激の奥行き方向の運動を知ることが、空間視の重要な一側面である。刺激の奥行き方向の運動情報を伝える細胞は、ネコの運動視関連領域であるクレアビショップ野とサル MST 野背側部に存在する。

##### (2) 眼球運動の制御

動く刺激を眼で追いかけるときに（円滑追跡眼球運動中に）、活動を高める細胞が MST 野に多く含まれる。このような部位を破壊すると、円滑追跡眼球運動に障害が起こる。また円滑追跡眼球運動中に、MST 野に電気刺激を施すと、眼球の運動に影響が現われる、これらのことは、MST 野が円滑追跡眼球運動の制御にかかわることを示している。

#### 4-1-8 頭頂連合野

頭頂間溝と上側頭溝に挟まれた下頭頂小葉には、視覚関連野として、7a, LIP, VIP がある。これら頭頂葉の損傷は、空間視や視覚依存性行動の阻害をもたらす。

##### (1) 網膜部位再現座標から空間座標へ

眼を動かしても、知覚する世界は安定して動かず、物体は空間の中でじっとしているように知覚される。もしも視覚系がすべて、網膜部位再現的な構造しかもっていなければ、物体像の網膜での位置変化が、自分の動きにより引き起こされたのか、物体の動きにより引き起こされたかを区別できないので、このような知覚は得られない。視覚系のどこかで、網膜上のどの位置が興奮したかを示す座標系から自分自身の頭や体を基準とした空間座標系への変換がなされているはずである。このような座標変換のための方法の一つは、眼球位置に関する情報を視覚系に導くことである。7a や LIP 及びその入力供給野である V3A, MST には、視線がどこを向いているかによって、視覚刺激に対する反応の大きさが変化する細胞がある。これらの細胞は通常の視覚細胞と同様、注視点に対して一定の相対位置に受容野をもつが、最大応答を示すのは、眼球が特定方向を向き、そこから一定の相対位置にある受容野に視覚刺激が入ったときである、つまり、視覚刺激は空間内の特定位置になくはならず、これらの細胞は空間内の特定位置に選択的に反応することになる。もう一つの方法は、眼の位置に左右されることなく、自分の頭または体との相対位置が不変であるような受容野をもつことである。すなわち、視線がどこに動こうと空間内の一定の場所に受容野をもつような細胞を作り出すことである。V6 がそのような細胞を含む。

## (2) 視覚誘導性行動のための情報処理

後頭頂葉に損傷を受けた患者は、目の前の物体が何であるかが分かるにもかかわらず、その物体をうまく握ることができない。すなわち、視覚情報を使って、その物体を握るのに適したように自分の手や指の形を合わせるができない。サルの後頭頂溝には、自身の手の形と動きのパターンに選択的な細胞の一群がある。これらの細胞の活動は、手の運動を制御するために必要な視覚情報を担っていると考えられている。

### 4-1-9 V4野

V4は視覚経路の中の多くの重要な領域を結ぶ交差点のような位置に存在する。V4細胞の受容野はV1やV2より大きく、中心視付近に受容野をもつ細胞で直径5度である。視野地図は維持されている。初期の研究において、V4が色の情報処理専用の領域であるという主張がなされた。V4が色情報の処理にかかわっているのは確かだが、V4が形の処理に大きくかかわっている下側頭葉皮質への重要な入力供給源であること、この領域が形や図形パターン(例えば渦巻き模様)に選択性を持つ細胞を多く含むこと、また一部の出力は背側経路へ向かうことから、この領域が色処理のみにかかわっているとは言えない。

### 4-1-10 下側頭葉皮質

下側頭葉皮質(IT)を破壊するとサルは、視覚対象物を弁別したり記憶しなければならないような行動課題の遂行が著しく困難になる。このとき、サルは視力障害や視野欠損を伴うことはなく、また物体の位置関係を知るような課題は行うことができる。ITは図形弁別や物体認知に深くかかわっている。

#### (1) IT野の解剖学

ITは、後半部のTEOと前半部のTEに分けられる。この二つの領域の破壊の効果は異なり、細胞の生理学的性質もTEOとTEでは顕著に異なる。ITへの主な視覚情報入力、V4からTEO及びTE後半部への投射によりもたらされ、出力先の主なものは、嗅周皮質(35, 36野)、扁桃体、大脳基底核、前頭葉皮質(8, 46野)である。

#### (2) 受容野と刺激選択性

TEO細胞の受容野の大きさは、中心視において半径5度前後だが、TE細胞は15度から40度に及ぶ。TE細胞の受容野は中心視を含み、反対側の視野を広くカバーし、さらに同側視野に及ぶことも多い。TEOでは粗い視野再現があり、TEでは視野再現は失われる。

TEOとTEでは反応を引き起こす有効刺激図形の複雑さが異なる。TEOには線分やスポットにも十分強く応じる細胞と、より複雑な図形に応じる細胞が混在するが、TEでは、ほとんどの細胞が、複雑な図形にのみ反応する。しかし、個々の図形特徴は一つの物体を特定できるほどの情報を含んでいない。複数の物体がこれらの特徴の一つを共有している場合には、その特徴に選択性をもつTE細胞がそれぞれの物体に反応する。また、これらの細胞の反応選択性はかなり鋭いが絶対的ではない。これらのことは、物体視経路の最終段であるTEにいたっても、特定物体が視野の中にあることを一つの細胞の活動が表わしているというようなことは起きていないことを意味している。

### (3) 図形特徴コラム

TE も機能的なコラム構造をもつ。幅 0.4~0.5 ミリのコラムの中に、似た図形に反応する細胞が集まっている。コラムの数は TE 全体で約 2000 と見積られている。ある特定の図形特徴に反応するコラムは複数あり、TE のコラム群が取り扱っている図形特徴の種類は数百と推定される。個々の TE 細胞が反応する図形特徴は物体の部分特徴にすぎず、一つの物体を特定できるほどには複雑でない。サルがある物体を見たときには、その物体のもつ様々な部分特徴が、それぞれに選択性をもつコラム群に反応を引き起こす。

## ■脳と神経モデル-4章

---

### 4-2 認知認識モデル

※執筆中



## ■脳と神経モデル—4章

### 4-3 記憶と学習システム

(執筆者：塚田 稔) [2008年5月 受領]

脳の記憶は保持できる時間に応じて、短期記憶（現在の時点から数時間、数日まで保存する現在の記憶で、海馬に存在すると考えている）と長期記憶（長い年月保存できる過去の記憶で、連合野に存在すると考えられている）に分類されている。長期記憶は陳述記憶と非陳述記憶に分けられ、陳述記憶は様々な事実や出来事に関する記憶（エピソード記憶）と個々の出来事の記憶を一般化、抽象化された意味記憶（概念構成もこの記憶に含まれる）に分けられている。言語は抽象化したシンボルとして、この記憶内で働いていると考えることができる。更には、目的的行动をするために、予想、計画などの記憶が存在する。この記憶は以前経験した出来事の記憶を基盤とし、将来の入力や未だ経験のない出来事を予想する記憶（未来の記憶）であり、前頭葉に存在すると考えられる。非陳述記憶は運動技能を獲得する手続き記憶やパブロフの犬で知られる古典的条件付の記憶などがある。人間の記憶システムは過去、現在、未来の記憶が三位一体となって一つの文脈を形成し、ある目的を実行するために働いている（図4・1）。

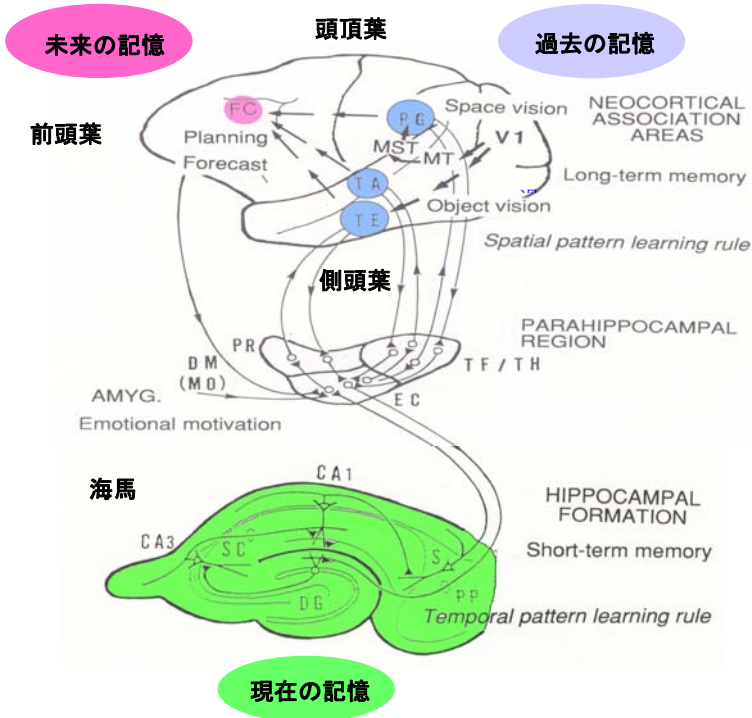


図4・1 過去・現在・未来記憶の三位一体

### 4-3-1 海馬-皮質の記憶システム

記憶の書き込みは次の二つの過程からなっていると考えられる。第一の過程は出来事などの情報を時空間の文脈として蓄える短期記憶の過程であり、一時的に文脈の保持が重要であり、海馬にその機能が存在すると考えられている。第2の過程は海馬での一時的な短期記憶を必要に応じて安定に蓄える長期記憶の過程であり、皮質(連合野)に存在すると考えられている。そこには文脈の類似性や関連性に基づいた記憶構造が存在している。したがって、海馬-皮質系の記憶書き込み過程では2段階の学習課程が存在し、海馬と皮質ではその機能に適した学習則が存在していると考えられる(図4-1参照)。

### 4-3-2 短期記憶と海馬

海馬の長軸方向の断面を図4-2に示す。この断面内には、一巡する神経回路網が三つのシナプスを介して形成されている。海馬への入力としての空間信号は、内臭野(EC)から歯状回(DG)の顆粒細胞へ入力され、その出力はCA3錐体細胞の樹状突起のシナプスを介して処理され、最終的にCA1錐体細胞の樹状突起のシナプ스에情報を記憶する。CA1錐体細胞の軸索は海馬台に至り海馬の出力となる。また、内臭野から直接CA1に接続するルートも存在する。この回路はちょうど金太郎飴と同じように断面はすべて同じ構造をもつ。海馬のCA3にはフィードバックをもった回帰性の結合をもつ特徴的な神経回路網が存在する。この回路網は、過去の履歴現象を現在に重ね合わせ、時間関係の情報を処理する。利根川(マサチューセッツ工科大学)らのグループは、遺伝子技術によってフィードバック神経回路網を破壊された欠陥モルモットは一つの行動目的を達成するのに非常に多くの手掛りが必要になることを報告している<sup>1)</sup>。

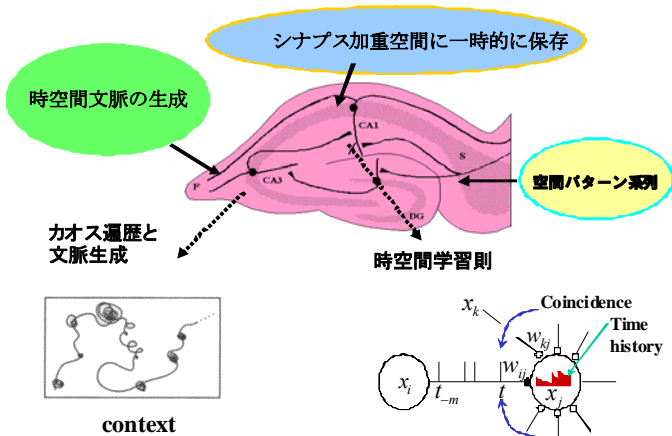


図4-2 短期記憶と海馬神経回路

このことによって、海馬の CA3 神経回路網は、時空間の文脈を形成するのに重要な役割をしていることが分かる。津田は海馬 CA3 の時空間文脈の形成にはカオスアトラクタ (カオスの遍歴) モデルを提唱している<sup>2)</sup>。海馬 CA1 回路網ではその文脈を一時的に記憶するために時空間学習則が有効に働いている<sup>3)~5)</sup>。この記憶の書き込みでは自己相似形表現 (カントールコード) が使われている可能性が理論と実験によって示されている<sup>6),7)</sup>。以上図 4・2 参照。

#### 4-3-3 海馬の学習則

##### (1) ヘブ学習と非ヘブ学習 (時空間学習則—STLR) の共存

海馬 CA1 錐体細胞におけるヘブ学習則と時空間学習則の共存することが明らかにされた (2 章 2-1-1 と 2-1-3 参照)<sup>5)</sup>。STLR では、シナプス結合強度は入力パターンとそれを受け取るニューロン側のシナプス荷重分布の相関によって決定される (Bottom-up)。一方、ヘブ学習則においては入力パターンと出力細胞の発火 (Top-down) の相関によっている。出力細胞を発火させるには入力の樹状突起の局所電位の時空間加重が細胞体 (厳密にはイニシャルセグメント) の閾値を超えるか、あるいは、別ルートからのアテンションや意志のようなトップダウン情報による場合が考えられる。海馬 CA1 錐体細胞では両学習則が共存する。両学習則の相互作用によるシナプス可塑性の存在は、記憶学習システムの次のような可能性メカニズムを推測することができる。

我々はある状況において、無意識にその状況と過去の経験とを比較する。次に発生可能性事象を予測し、予測報酬が最大となるような行動を計画する。この過程において我々の過去、現在、そして予測は統合して働き、我々の行動を生成する。もし行動が適切ではなかった場合、新たな作業仮説が創られ、または推論し、以前の仮説を訂正する。この脳内の時間経過過程は外部環境と適合するまで反復される内部モデルの細胞レベルの修正過程を表現するかもしれない。また強化学習の枠組みにおいては、海馬 CA1 錐体細胞の樹状突起-細胞体の相互作用は戦略、報酬、価値といった文脈の細胞レベルや回路網レベルのメカニズムに重要であろう。

##### (2) ヘブの学習則<sup>8)</sup> (HEBB) と時空間学習則<sup>4)</sup> (STLR) の特徴

時空間系列を処理する単純な一層神経回路網に二つの学習則を適用し、時空間分離能力を比較すると、ヘブの学習則 (拡張学習則も含む) よりも時空間学習則が時空間パターン系列の識別能力が高い (2-1-4 参照)<sup>9)</sup>。一方、ヘブ型学習則は、類似な入力系列パターンを同一出力パターンに引き込む性質 (Pattern Completion) の特徴がある<sup>9)</sup>。このことから短期記憶のようにパターン分離機能が要求される海馬では時空間学習則が適している。一方、長期記憶のようにパターンの類似性の構造が要求される連合野ではヘブの学習則が適している。

#### 4-3-4 海馬-皮質系の記憶システム

皮質では、記憶は神経細胞の興奮パターンによって表現され、細胞間のシナプス結合を変化させる学習則によって神経回路網に安定なパターン (アトラクタと呼ぶ) として記録され、その数によって情報容量が決まると考えられている。

出力から入力へのフィードバック結合をもつニューラルネットワーク (再帰的ネットワーク) は、与えられた入力に対し出力を計算し、その結果をフィードバックして入力を修正し、

その入力に対して出力を再計算する．このような処理を何度も繰り返すと安定なネットワークでは出力値の変化が少なくなり一定の出力値（平衡点）をとるようになる．力学系では、平衡点をアトラクタ、アトラクタに達する状態の集合を引込み領域と呼ぶ．このようにアトラクタ構造を創ることによって皮質では長期記憶を実現していると考えられる．アトラクタには固定パターン、リミットサイクル、カオスパターンなどが存在する．人間の記憶は連想的である．すなわち、記憶の一部から関係の深いより多くのことがらを想起できる．例えば、ある曲の数小節を聞いただけで、情景、音、においなどの感覚的な経験を完全に思い出すことができる．これに対して、普通のコンピュータの記憶は、番地指定によって取り出される．再帰的なネットワークは連想メモリを構成する．人間の記憶のように、想起してほしいデータの一部を与えると、完全な“記憶内容”を返してくる． $n$  個のニューロンからなる再帰的なネットワークを用いて連想記憶を実現するには、 $2n$  個の状態に対応する  $n$  次元超立方体の“記憶すべき頂点”でのエネルギーが最小になるように重みを選べばよい．最も簡単なものは、 $m$  個の記憶させたいパターンに対して、細胞間の結合荷重を次の式によって与える方法で、自己相関モデルと呼ばれるものである (Nakano, 1972<sup>10</sup>, Amari, 1972<sup>11</sup>, Hopfield, 1982<sup>12</sup>).

$$w_{ij} = \sum_{d=1}^m (OUT_{i,d} OUT_{j,d})$$

ただし、 $m$  = 記憶させたい出力ベクトル数、 $d$  = ある記憶させたい出力ベクトルの番号、 $OUT_{i,d}$  = 記憶させたい出力ベクトルの  $i$  番目の成分．

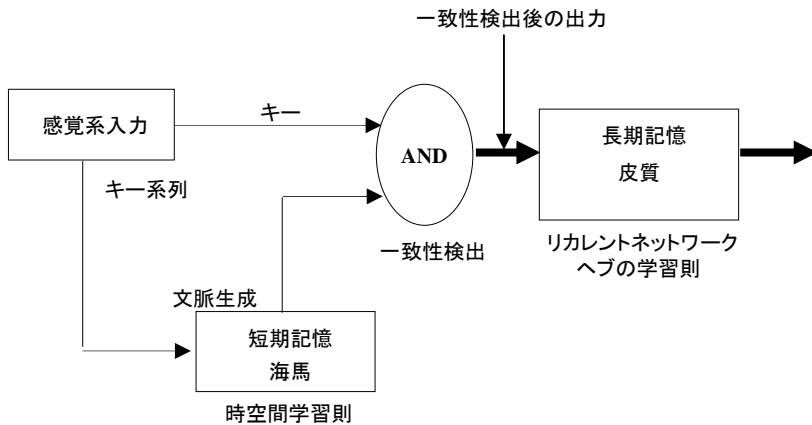


図 4・3 海馬—皮質システムのブロック図

塚田らは生理実験に基づいた海馬—皮質系のモデル<sup>8)</sup>を提案し、文脈に基づく記憶書き込み、及び読み出しのモデルを実現している (図 4・3)<sup>8)</sup>。このモデルは次の 2 段階の機能からなる。

- 1 段階：塚田らの提案した時空間学習則<sup>8)</sup>が有効に働き海馬神経回路網に時空間の出来事の文脈を作る。

2 段階：ヘブの学習則により連合野に出来事の類似性の構造（距離空間）をもつ強固なアトラクタ構造を作る．このモデルでは海馬の文脈情報と感覚入力 of キーパターン情報が長期記憶の入力となる．実験結果に基づいて，感覚情報と海馬情報の間にタイミングに依存したタイミング窓を一致性検出回路として導入している．

このモデルの計算機シミュレーション結果によれば，海馬の文脈情報の導入により引き込み領域が拡大するとともに，書き込み及び読み出し時間が著しく改善され，記憶パターンの正解率も向上した<sup>13)</sup>．

#### 4-3-5 意識と無意識と記憶

事故で脳に重度の損傷を受け，人間として生きる重要な機能である意識を失った状態を植物状態と言っている．植物状態では自己認識や意識の兆候がなく，高度な思考を処理することができないようだ．人間の注意の脳活動範囲を fMRI で調べた結果によれば広く脳の高次機能部位が活動することが知られている．したがって，意識は脳の解剖学的境界を越えて機能し，脳の高次機能をダイナミックに情報操作している．意識とは脳に内在し，ダイナミックに細胞集合体を操作する機能と仮定するのが妥当であろう．

推論による思考は論理的でかつ因果律に基づいた意識的活動であるのに対し，想像や妄想や夢は直感的で連想的な無意識の要素が強い．いわゆる感性と呼ばれている世界は推論による論理的世界と対比して使われている．物事の新しい文脈を創るときは極めて意識的・論理的思考が働くが，訓練や慣れによってその文脈は無意識的な反射や連想的思考に姿を変えてしまう．その無意識の世界に落とされた記憶情報はお呼びがかかれば再び意識上に登場し，再び修正し，新たに記憶情報を作り出す．したがってあたかも論理構想のない反射的な感性と称する記憶世界にも論理構造が当然含まれていると考えるべきであろう．このように，意識，無意識の記憶情報をダイナミックに捉えることが重要である（図 4・4）．

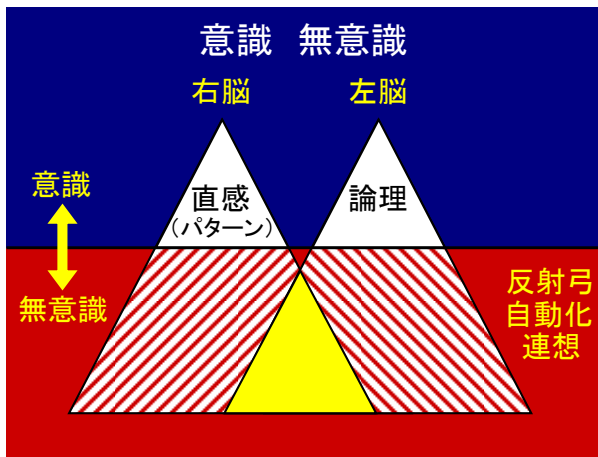


図 4・4 意識・無意識の動的思考過程

**■参考文献**

- 1) K. Nakazawa, M. C. Quirk, R. A. Chitwood, M. Watanabe, M. F. Yeckel, L. D. Sun, A. Kato, C. A. Carr, D. Johnston, M. A. Wilson and S. Tonegawa “Requirement for hippocampal CA3 NMDA receptors in associative memory recall.,” *Science*, 297, pp.211-218, 2002.
- 2) I. Tsuda, “Dynamic link of memory-chaotic memory map in nonequilibrium neural networks,” *Neural Networks*, 5, pp.313-326, 1992.
- 3) M. Tsukada, T. Aihara, M. Mizuno, H. Kato and K. Ito, “Temporal pattern sensitivity of long-term potentiation in hippocampal CA1 neurons,” *Biological Cybernetics*, 70, pp.495-503, 1994.
- 4) M. Tsukada, T. Aihara, H. Saito and H. Kato, “Hippocampal LTP depends on spatial and temporal correlation of inputs,” *Neural Networks*, 9, pp.1357-1365, 1996.
- 5) M. Tsukada, Y. Yamazaki and H. Kojima, “Interaction between the Spatio-Temporal Learning Rule (STLR) and Hebb type (HEBB) in single pyramidal cells in the hippocampal CA1 Area,” *Cogn Neurodyn*, 1, pp.157-167, 2007.
- 6) I. Tsuda and S. Kuroda, “Cantor coding in the hippocampus,” *Japan Journal of Industrial and Applied Mathematics*, 18, pp.249-258, 2001.
- 7) Y. Fukushima, M. Tsukada, I. Tsuda, Y. Yamaguchi and S. Kuroda, “Spatial clustering property and its self-similarity in membrane potentials of hippocampal CA1 pyramidal neurons for a spatio-temporal input sequence,” *Cogn Neurodyn*, 1, pp.305-316, 2007.
- 8) D. O. Hebb, “The Organization of Behavior,” New York: John Wiley, 1949.
- 9) M. Tsukada and X. Pan, “The spatiotemporal learning rule and its efficiency in separating spatiotemporal patterns,” *Biological Cybern*, 92, pp.139-146, 2005.
- 10) K. Nakano, “Associatron-a model of associative memory,” *IEEE Trans., SMC-2*, pp.381-388, 1972.
- 11) S. Amari, “Learning patterns and pattern sequences by self-organizing nets of threshold elements,” *IEEE Trans. Computers*, C-21, 11, pp.1197-1206, 1972.
- 12) J. J. Hopfield, “Neural networks and physical systems with emergent collective computational abilities,” *Proceedings of the National Academy of Sciences*, 79, pp.2554-2558, 1982.
- 13) X. Pan and M. Tsukada, “A model of the hippocampal-cortical memory system,” *Biological Cybernetics* 95, pp.159-167, 2006.

## ■脳と神経モデル—4章

### 4-4 記憶と学習のモデル

(執筆: 石川眞澄) [2009年4月 受領]

「4-3 記憶と学習システム」は生物やヒトの記憶と学習の仕組みを述べたものであるのに対し、本節は生物やヒトから着想を得たモデルの立場から記憶と学習について述べる。学習のモデルは多種多様であるが、教師情報の有無が分類の一つの基準となる。入力データに対応する正答を教師情報として与える教師あり学習 (Supervised Learning) と教師情報のない自己組織化学習 (Self-organized Learning) は、その両極端に位置づけられる。その中間には、モデルの振る舞いの良さを報酬あるいは罰という形で与える強化学習 (Reinforcement Learning) がある。記憶のモデルに関しては、相関学習及びホップフィールドネットワーク (Hopfield network) に基づく連想記憶 (Associative Memory) モデルがよく知られている。

1950年代末にローゼンブラット (Rosenblatt) は脳を模した学習機械であるパーセプトロン (Perceptron) を提案し<sup>1)</sup>、これがきっかけとなり1960年代にはパーセプトロンに関する多くの研究がなされた。1969年にミンスキー (Minsky) らはパーセプトロンの理論的限界を数学的に解明し<sup>2)</sup>、これが一因で1970年代にこの分野の研究者人口が減少したとも言われている。

1980年代前半には、ホップフィールドネットワーク<sup>3)</sup>、ボルツマン機械 (Boltzmann Machine)<sup>4)</sup>、誤差逆伝播学習 (Error Backpropagation Learning)<sup>5)</sup>などの重要な提案が相次いでなされた。このような提案に加え、英語の発音規則の誤差逆伝播学習による学習やホップフィールドネットワークによる巡回セールスマン問題の求解などの魅力的な応用例が注目を集めた。このため1980年代後半にはニューラルネットワーク研究が再び盛んになった。

熱狂的なブームは1980年代で終わったが、その後多くの周辺理論分野と結びつき、新分野として着実に発展した。工学应用に関して、当初はニューラルネットワークでもできるという牧歌的時代であったが、現在はニューラルネットワークでしかできない研究例や、従来型技術の性能を凌駕する研究例が求められている。

#### 4-4-1 教師あり学習

教師あり学習は、入力と対応する出力目標のペア群を訓練データとして与え、入力から出力への写像をできるだけ再現する結合重み (Connection Weight) を求めることを目的とする。その典型例として、パーセプトロン学習と誤差逆伝播学習について述べる。

##### (1) パーセプトロン

ローゼンブラットは、感覚ユニット (Sensory Unit)、連合ユニット (Association Unit) 及び反応ユニット (Response Unit) から構成され、可変の結合重みを有するパーセプトロンを提案した。ここで感覚ユニットは、視覚の場合には網膜の視細胞に相当し、外界からの物理的刺激に反応して出力を出す。連合ユニットは、感覚ユニットから得た情報の前処理を行い、感覚層に関する断片的情報を表現する。反応ユニットは、連合ユニットの情報を統合して、ネットワーク全体としての判断を与える。

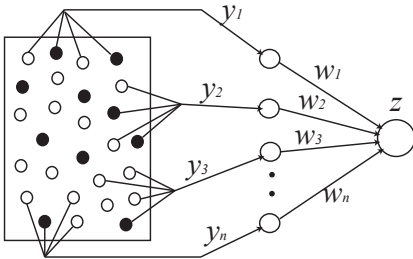
ローゼンブラットは様々なパーセプトロンを提案し、これらの特性を解明した。そのうち

最も簡単なものが図4・5の単純パーセプトロン (Simple Perceptron) である。単純パーセプトロンの中でも、連合ユニット、反応ユニットの出力関数が線形閾値関数 (Linear Threshold Function) であるものを基本パーセプトロン (Elementary Perceptron) と呼ぶ。感覚層空間で線形分離可能 (Linearly Separable) でなくても、感覚・連合層間のランダムな固定結合により非線形に変換され、連合層空間では線形分離可能になり得る。ただ、複雑な課題を線形分離可能問題に変換するには、一般に膨大な数の連合ユニットを必要とする。

図4・6は1個のニューロンの特性を抽象化した数学的モデルを表す。ここで入力の重み付き和  $x_i$  (膜電位) は  $x_i = \sum_k w_k y_k - \theta$  で表される。ただし、 $y_k$  は  $k$  番目の入力、 $w_k$  はその結合重み、 $\theta$  は閾値である。また出力  $z$  は次の線形閾値関数で与えられる。

$$z = l(x) = \begin{cases} 0 & (x \leq 0) \\ 1 & (x > 0) \end{cases} \quad (4 \cdot 1)$$

ここで  $\mathbf{W} = (w_1, w_2, \dots, w_n, -\theta)$ 、 $\mathbf{Y} = (y_1, y_2, \dots, y_n, 1)$  とおくと、 $z = l(\mathbf{W}^T \mathbf{Y})$  と表される。このときパーセプトロン学習則は  $\Delta \mathbf{W} = -\varepsilon (z - \hat{z}) \mathbf{Y}$  で与えられる。連合層上で線形分離可能な課題は、有限回の繰り返しで全ての入力サンプルに対して正答を出力するパーセプトロンが得られることがパーセプトロン収束定理により保証される。



左から順に感覚層、連合層、反応層

図4・5 単純パーセプトロンの構造

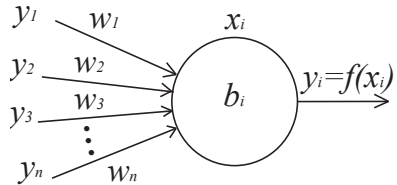


図4・6 ニューロンの数学的モデル

## (2) 誤差逆伝播学習

誤差逆伝播学習の基本的考え方は、出力自乗誤差和を最小にする結合重みを、繰り返し計算により求めることである。出力自乗誤差和の結合重みに関する偏微分値に比例して結合重みを修正するが、その際、最小の計算量で偏微分値を計算できるよう再帰的計算を工夫している。

図4・7の階層的ニューラルネットワーク構造で、出力自乗誤差和は次式で与えられる。

$$E = \sum_k E_k = \frac{1}{2} \sum_k \sum_i (y_i^{(k)} - d_i^{(k)})^2 \quad (4 \cdot 2)$$

ただし  $E_k$  は  $k$  番目の訓練パターンに対応する出力自乗誤差、 $y_i^{(k)}$  は  $k$  番目の訓練入力パターンに対する  $i$  番目の出力ユニットの出力、 $d_i^{(k)}$  はその出力目標である。 $k$  番目の訓練パターンに



対する出力自乗誤差の結合重みに関する偏微分は、

$$\frac{\partial E_k}{\partial w_{ij}} = \frac{\partial E_k}{\partial x_i^{(k)}} \frac{\partial x_i^{(k)}}{\partial w_{ij}} \quad (4 \cdot 3)$$

ただし  $w_{ij}$  は  $j$  番目のユニットから一つ上層の  $i$  番目のユニットへの結合重み、 $x_i^{(k)}$  は  $k$  番目の訓練入力パターンに対する  $i$  番目のユニットの膜電位である。このユニットの誤差信号  $\delta_i^{(k)}$  は、

$$\delta_i^{(k)} = -\frac{\partial E_k}{\partial x_i^{(k)}} \quad (4 \cdot 4)$$

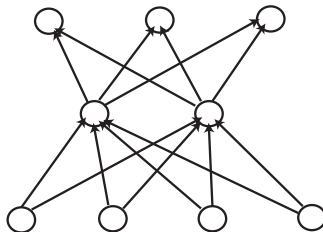
ここで微分の連鎖規則を用いると、結合重みの修正量  $\Delta w_{ij}$  は次式で与えられる。

$$\Delta w_{ij} = \eta \delta_i^{(k)} y_j^{(k)} \quad (4 \cdot 5)$$

まず出力層の誤差信号を求め、次にこれを利用して一つ下層の誤差信号を求めるという再帰的計算を順次入力層に向かって行うというのが誤差逆伝播学習と名付けられた理由である。

パーセプトロン学習と誤差逆伝播学習の主要な違いは、①前者は出力関数として線形閾値関数を用いるのに対し、後者はシグモイド関数 (Sigmoid Function) を用いる。②前者は連合層・反応層間の結合重みだけしか学習できないのに対し、後者はすべての結合重みを学習可能。ここでシグモイド関数とは、単調増加な連続関数であり、 $\pm\infty$  で有限な値に収束する関数群である。

パーセプトロンは線形閾値関数を用いるので、入力の微小な変化に対して出力が一般に変化しない。これが上記②にあるように、連合層・反応層間の結合重みだけしか学習できない理由である。これとは対照的に誤差逆伝播学習では出力自乗誤差評価の任意の結合重みに関する勾配を計算できるので、すべての重み結合の学習が可能となる。これにより、パーセプトロンでは必要になるかも知れない連合ユニットをあらかじめすべて用意しておく必要があるのに対し、誤差逆伝播学習では必要な隠れユニットをデータに基づいて学習途中で生成できるという利点が生じる。



下から順に入力層、隠れ層 (中間層)、出力層

図 4・7 階層的ニューラルネットワーク構

誤差逆伝播学習は全ての結合重みを学習できるという点でパーセプトロン学習よりも強力である。また次元の呪いから逃れられるという利点もある。これらの代償として、誤差曲面

が一般に多峰性となり、出力自乗誤差和の極小解に相当する結合重みしか求まらないという致命的な欠陥がある。実用上は、結合重みの初期値を多数用意することによりこの欠陥を解消している。

誤差逆伝播学習には以下の問題点がある。

- (1) 階層数や隠れユニット数の事前情報がない場合、ネットワーク構造に関する試行錯誤的探索が必要となる。何故ならば、ネットワークが小さすぎれば学習が不十分となり、逆に大きすぎれば過学習のため汎化能力が低下するので、適切な大きさのネットワークで学習することが不可欠だからである。
- (2) 学習後のネットワークがブラックボックスとなり、隠れユニットの意味付けが困難である。
- (3) 解が極小値に陥る危険性があり、ネットワーク規模の増大とともに深刻化する。
- (4) ネットワークの規模が増大すると、学習のための計算量が急増する。

(1), (2)に対する解決策として構造学習がある。ここで学習可能性と解釈可能性について触れておきたい。誤差逆伝播学習で求まるニューラルネットワークは分散表現 (Distributed Representation) となり、学習結果の解釈が困難である。他方、固定結合ネットワークで活性化度伝播に基づくダイナミクスに着目するモデルもあり、これは局所表現 (Local Representation) をとるので解釈が容易である反面、結合重みを人が与えるので説得力に欠ける。この両者の矛盾を止揚しようとする方法が、学習を行い、かつ学習結果が局所表現となる構造学習である。言い換えれば、構造学習の学習結果は、学習可能性と解釈可能性を兼ね備えている。

(3), (4)に対する解決策としてモジュール構造ネットワークがある。何故ならば各モジュールは小さいため、計算量も少なく、極小値問題も深刻ではないからである。

ここまでは階層的ニューラルネットワークに対する誤差逆伝播学習について述べてきたが、リカレントネットワーク (Recurrent Network) に対する学習アルゴリズムも提案されている。

#### 4-4-2 自己組織化学習

ある入力に特徴的に反応する特徴抽出細胞が教師信号なしで生成され、これら特徴抽出細胞が位相保存的なマップを形成するという自己組織化は、視覚野、聴覚野、体性感覚野、運動野などで見られる現象である。第一次視覚野でコラム構造が発見され、そのモデルがマルスブルク (Malsburg)<sup>6)</sup>、ウィルショウ (Willshaw) ら<sup>7)</sup>によって提案された。位相保存的なマップ (Topology Preserving Map) 形成にとって、競合学習 (Competitive Learning) による勝者ニューロンの決定及び近傍ニューロン群による結合重みの同時的学習が本質的である。これらモデルは一次視覚野の位相保存的なマップ形成の優れたモデルである反面、データ次元が二次元に限定されている、計算量の大きな連立微分方程式を多数回解く必要があるなど、工学応用の観点からは適していない。

コホネン (Kohonen) は、任意の次元の入力が扱え、連立微分方程式の代わりに線形演算により勝者ニューロンを決め、近傍ニューロン群の同時的学習を実現する自己組織化マップ (Self-Organizing Map : SOM) を提案した<sup>8)</sup>。工学応用に適しているため、提案以来広く使われてきた。

### 4-4-3 強化学習<sup>9)</sup>

強化学習は、例えば子供の遊びに典型的に見られるように、環境との相互作用に基づく学習であり、明示的な教師は存在しない。ある状態である行動をすると、状態が変化し、同時に環境からの報酬 (Reward) が与えられ、将来の累積報酬を最大化するよう最適方策 (Policy) を学習する。この累積報酬を価値関数 (Value Function) あるいは行動価値関数 (Action-value Function) と呼ぶ。この際、報酬が時折あるいは遅れて与えられる点、及び試行錯誤的に探索する点に強化学習の特徴がある。ある状態である行動をした場合、次の状態を表現する状態遷移確率行列及び期待報酬を環境のモデルと呼ぶ。

強化学習では、探索 (Exploration) と開拓 (Exploitation) のトレードオフが重要である。すなわち多くの報酬を得るには過去に多くの報酬をもたらした行動を選択する (開拓) ことが有効であり、良い行動を発見するには過去に選択しなかった行動も試みる (探索) ことが必要である。累積報酬を最大化する最適方策を求める方法は、次の3種類がある。

#### (1) ダイナミック・プログラミング (Dynamic Programming : DP)

環境の完全なモデルが存在する場合に、ベルマンの最適性方程式を更新式に変換して最適方策を求めるアルゴリズムである。DP は数学的基盤を与え、他の強化学習法は、完全な環境のモデルなしに、より少ない計算量で DP と同じ結果を求めることを目的とする。

$$V^*(s) = \max_a Q^*(s, a) = \max_a E_{\pi} \{R_t | s_t = s, a_t = a\} = \max_a \sum_{s'} P_{ss'}^a [R_{ss'}^a + \gamma V^*(s')] \quad (4 \cdot 6)$$

ただし、 $V^*(s)$  は状態  $s$  における最適価値関数、 $a$  は行動、 $R_t$  は時刻  $t$  における報酬、 $P_{ss'}^a$  は状態が  $s$  から  $s'$  に変化し行動  $a$  を取る遷移確率、 $R_{ss'}^a$  は状態が  $s$  から  $s'$  に変化し行動  $a$  をとる場合の報酬、 $\gamma$  は割引率である。

#### (2) モンテカルロ法 (Monte Carlo method : MC)

環境の完全なモデルを必要とせず、模擬経験によるエピソード終了時の利得を用いて価値関数を更新する。

#### (3) TD 学習 (Temporal Difference Learning)

環境の完全なモデルなしに、またエピソードの終了を待たずに他の推定値に基づいて価値関数を更新する。環境のモデルが不要という点で DP よりも優れており、エピソード終了まで待たなくてよいという点で MC よりも優れている。TD 学習の中でも Q 学習が最もよく用いられている。Q 学習の更新式は以下で表される。

$$Q(s_t, a_t) \leftarrow Q(s_t, a_t) + \alpha \{r_{t+1} + \gamma \max_a Q(s_{t+1}, a) - Q(s_t, a_t)\} \quad (4 \cdot 7)$$

ただし、 $Q(s_t, a_t)$  は時刻  $t$  の状態  $s_t$ 、行動  $a_t$  における行動価値関数、 $\alpha$  は学習率、 $r_{t+1}$  は時刻  $(t+1)$  での報酬である。

### 4-4-4 連想記憶モデル

連想記憶は計算機における機械的な番地記憶とは対照的であり、ヒトの記憶に類似する。中野<sup>10)</sup>、コホネン、アンダーソン (Anderson) らにより同時期に、独立に相関学習を用いた

連想記憶モデルが提案された。相関とは、入力ベクトルと出力ベクトルの相関である。教師あり学習で同じことを実現しようとするると繰り返し計算を必要とするのに対し、連想記憶モデルは入出力ペア群を一巡するのみでよい。教師あり学習では入力から出力への写像がきちんと学習できるのに対し、連想記憶モデルでは入力が互いに直交していない限り、正しい出力に相互干渉項というノイズが加わったものが得られる。ただ、出力が2値の場合、相互干渉項が十分小さければ、正しい出力を求めるのに支障はない。更に、入力が多少の誤りがあっても、誤り訂正も可能である。正しく想起する確率を、記憶させたいパターン数と入力次元数との比率の関数として求めることが可能である。

別のアプローチとして、記憶させたい情報をエネルギー関数の極小値として設定したホップフィールドネットワークにより、連想記憶が可能である。これは、キーとして与えた情報を出発点として、エネルギーの小さくなる方向に状態遷移を繰り返すことにより、出発点の近傍に存在する極小値に辿り着くので、この極小値に記憶した情報を想起できるという仕組みである。

## ■参考文献

- 1) F. Rosenblatt, "The perceptron : A probabilistic model for information, storage and organization in the brain," *Psychological Review*, 65, pp.386-408, 1958.
- 2) M. Minsky and S. Papert, "Perceptrons -- An Introduction to Computational Geometry," The MIT Press, 1969.
- 3) J. J. Hopfield, "Neural networks and physical systems with emergent collective computational abilities," *Proceedings of the National Academy of Sciences USA*, 79, pp.2554-2558, 1982.
- 4) D. H. Ackley, G. E. Hinton, and T. J. Sejnowski, "A learning algorithm for Boltzmann machines," *Cognitive Science*, 9, pp.147-169, 1985.
- 5) D. E. Rumelhart, G. E. Hinton, and R. J. Williams, "Learning representations by back-propagating errors," *Nature*, 323-9, pp.533-536, 1986.
- 6) C. von der Malsburg, "Self-organization of orientation sensitive cells in the striata cortex," *Kybernetik*, 14, pp.85-100, 1973.
- 7) D.J. Willshaw and C. von der Malsburg, "How patterned neural connections can be set up by self-organization," *Proc. R. Soc. Lond. B.*, 194, pp.431-445, 1976.
- 8) T. Kohonen, "Self-organized formation of topologically correct feature maps," *Biological Cybernetics*, 43, pp.59-69, 1982.
- 9) R. S. Sutton and A. G. Barto, "Reinforcement Learning," The MIT Press, 1998.
- 10) K. Nakano, "Association --- a model of associative memory," *IEEE Trans. SMC*, 2-3, pp.380-388, 1972.

## ■脳と神経モデル—4章

### 4-5 言語システム

(執筆者：乾 敏郎) [2008年5月 受領]

#### 4-5-1 主な失語症と責任部位

言語理解や言語産出には、脳の大部分の領域が使われていると考えられる。しかし、言語の重要な側面が、脳の局所的な領域の損傷によって生ずることが古くから知られている。

##### (1) ウェルニッケ失語

ウェルニッケ野の損傷によって言語理解、特に音声認識に障害が出ることが知られている。そのため、ウェルニッケ失語においては、言語の理解や産出に障害が生じ、発話は流暢であるが、ジャargon (Jargon) と呼ばれる新造語がよく見られる。

##### (2) 伝導失語と超皮質性運動失語

左半球の縁上回や弓状束の障害によって、伝導失語が生ずる。伝導失語は、復唱することによって大きな障害が認められるので、縁上回や弓状束は、音声の内的なりハーサルに重要であると考えられている。また、左の補足運動野を中心とした障害によって、超皮質性運動失語が生じる。これは、復唱機能は良好だが、発話を自ら開始することは少なく、エコラリア (反響言語) が見られる。

##### (3) ブローカ失語

ブローカ野を含み、左の6野、8野、9野、10野、46野、及び大脳基底核などに障害が認められる場合には、重篤なブローカ失語が生ずる。ブローカ失語の特徴は、言語産出において、単語間の休止時間がきわめて長く、名詞は正しく使われることが多いが、動詞や、いわゆる文法的な単語である接続詞や前置詞、代名詞などは上手く使うことができない。したがって、いわゆる Open Class の単語は利用可能だが、Closed Class の単語の使用に障害が生ずる。また、単純な文の理解は可能であるが、それを復唱することに障害がある。更に、可逆文の理解にも問題があることが指摘され、ブローカ失語は単に文産出の障害ではなく、文法の理解や使用に障害が生ずることから、特に言語の脳研究においては注目されている。可逆文は、例えば、“The girl was kissed by the boy.”のような文であり、どちらも動作主となる可能性のある文という意味で、可逆文と呼ばれている。このような文を呈示したときに、どちらがどちらにキスをしたのかということが正しく理解できない。特に、受動態の理解障害が顕著にみられる。

##### (4) 名詞と動詞の生成

Damasio らは、大規模な神経心理学的な調査を行い、名詞の生成には、側頭葉が関与していること、また、側頭葉の障害部位によって欠落する名詞のカテゴリーが異なることを発見した<sup>1)</sup>。彼らの結果によると、左の側頭極付近では人名が、左の下側頭葉前部および下側頭葉後部は、それぞれ動物名や道具名に特異的に障害が生ずることが明らかになった。また、Damasio と Tranel は、主として運動前野とブローカ野の44野の障害を受けた患者が、動詞の

生成に障害をもつことを報告している<sup>2)</sup>。

#### (5) 言語領野孤立症候群

ヘッシェル回を含む聴覚経路や、ウェルニッケ野、ブローカ野、これら二つの領野を結ぶ弓状束も正常であったが、それ以外の大脳皮質が広範に障害を受けた失語症が報告されている<sup>3)</sup>。言語領野孤立症候群である。この場合、患者は、自発性の発話はほとんどなかったが、復唱は正常であった。また、ハミングに合わせて歌を歌ったり、障害を受けてから聞いた歌を歌うことも可能であった。また、言語領野孤立症候群では、質問に対して適切に答えられないが、質問に含まれる単語や句に対して、関連のある決まり文句やコロケーションなどによって返答することが報告されている。このように、言語領野孤立症候群では、単語系列の生成と学習が行われるが、意味的な処理はできないことが分かる。すなわち、単語の意味は、これらの古典的言語や以外の領域、特に、側頭葉などに貯蔵されているものと考えられる。

### 4-5-2 脳イメージングによる研究

#### (1) 文の意味処理

Friederici によると、文の意味処理には、左の側頭葉後部ならびに中側頭回が重要な役割を果たしているらしい<sup>4)</sup>。一方、左の中側頭回と同時に意味処理に関連するのが、BA45/47 の領域である。また、Grabowski らは、下前頭溝及び中心前溝の交点付近を LPREM と名付け、これは BA9/44 に対応する<sup>5)</sup>。この部分は、特に、動詞（動作を示す単語）、あるいは、特別な動作と関係する対象の単語を検索する時に働くのではないかと考えられる。また、BA45 は、視覚的に呈示された図に対して、命名（単語の生成）においてよく活動することから、語とそれに対する意味を側頭葉のカテゴリー知識と関連させて統合する場所ではないかと考えられる。

#### (2) 統語処理

また、Friederici (2002) によれば、BA44 の下端と前頭弁蓋は、局所的な句構造構築に必要であり、この部分が統語的なワーキングメモリの一部を構成しているようである<sup>4)</sup>。Inui らは、いくつかのイメージング実験から、BA47 が、日本語の場合、助詞による単語の意味役割付与という働きを実現しているのではないかと考えている<sup>6)</sup>。彼らは助詞判断課題を行った。呈示された一文字が助詞であるかないかを判断する課題である。このような課題を行うと、選択的に BA47 が活動する。また、単文において、助詞を処理するときの脳活動が事象関連 fMRI によって計測され、この場合も BA47/45 で活動がみられた (Ogawa ら)<sup>7)</sup>。また、Ogawa らは、複文を用いて文の階層構造を理解する機能を fMRI を用いて調べた<sup>8)</sup>。実験では、左分岐文、中心埋め込み文、及び等位接続文が用いられた。その結果、中心埋め込み文で左の運動前野の有意な活動上昇がみられることから、この領域が中心埋め込み文における構文的なワーキングメモリを担っているのではないかを考えた。なお、この領域は BA6/8/44 にまたがる領域であり、LPREM を含む領域である。

#### (3) 拡散テンソル解析によるネットワークの同定

Lehéricy らは、ヒトの脳で拡散テンソル解析を用いて、神経線維の経路の同定を行って

る<sup>9)</sup>。それによれば、尾状核の頭部と結合している線維束は、前補足運動野や、BA9、46、45、47、10などに結合しているという。また、被殻の前部は、前頭前野に対してはBA45や47、また、背外側前頭前野 (dorsolateral prefrontal cortex; DLPFC; BA9/46) と結合している。これは、BA45や47が脳基底核ループの一部を構成していることが示唆される。また、Cataniらは、同様の手法を用いて、弓状束を詳細に分析した<sup>10)</sup>。その結果、(1) ウェルニッケ野とブローカ野を直接つなぐ線維束、(2) ウェルニッケ野と下頭頂葉をつなぐ線維束、(3) ブローカ野と下頭頂葉をつなぐ線維束に分かれていることが判明した。これは、伝導失語の諸症状を説明するのに有益な情報である。また、後に述べる乾の予測も実証されたことになる<sup>11)</sup>。

#### (4) 手話の理解

Bavelierらの実験では、英語の文が一単語ごとに呈示され、このとき、健常な実験協力者は、文を読んで後の問いに答えなければならなかった<sup>12)</sup>。このような実験を行うと、英文の処理に対して左半球のブローカ野、上側頭溝、ウェルニッケ野、角回、中心前溝及び中心溝の活動がみられた。Nevilleらは、先天性のろう者がアメリカ手話言語 (ASL) で呈示された文を理解する条件で脳活動を記録した<sup>13)</sup>。この場合も、左半球のブローカ野、中心前溝、及び中心溝、上側頭溝などで活動が見られた。しかし、この場合はウェルニッケ野で有意な活動は見いだせなかった。このように、ASLを理解するときにも健常被験者が英語を読むときとかなり多くの共通する領域が活動することが分かる。したがって、これまで紹介してきた言語領域は、入力の様式によらない言語処理システムであると考えられる。

#### (5) 運動系列予測学習仮説

乾は、運動系列予測学習仮説において、以下のような四つの仮説を考えた<sup>11)</sup>。

- (a) 文法獲得を含む認知機能の発達には、模倣機能が中心的役割を果たす。
- (b) 模倣により対象が自己中心的に表現される。
- (c) 自己中心的表現とは、自己の身体運動を入力とする対象の順モデルである。
- (d) 自己中心的表現は、対象の時間変化の予測を含む。

ここでの模倣は、視覚情報から自己運動情報への変換 (見まねやジェスチャーの理解) や聴覚情報から自己運動情報 (構音指令) への変換 (聞きまね、復唱) が重要となる機能が含まれる。おそらく、予測と模倣の二つの機能が一体化されており、対象の動きを予測しながら模倣を進めていると考えている。さらに乾では、44野の双方向結合により、メンタルシミュレーションが実現され、44野では系列情報の生成を、40野ではそれに対応する視覚または音韻情報が生成され、音韻情報はウェルニッケ野を活動させる<sup>11)</sup>。また、視覚情報は楔前部、高次視覚野を活動させる。一方、44野と45野の相互作用を通じて、視覚・聴覚情報から次の運動指令 (アクションユニット) が予測、生成される。つまり、45野-44野-40野-ウェルニッケ野-聴覚野の経路によって、次の運動情報を予測し、生成するため、スムーズな運動を内部で生成することができると考えている。また、45野と脳基底核、視床からなるループによって、運動系列が学習され、これにより文法知識が獲得されたと考えた。

#### (6) シンボルの接地 Grounding

シンボルは、外界の対象や動き、状況、更には感情などに対応されなければならない。こ

れをシンボルの接地 (Symbol Grounding) という。これまでに述べたように、いわゆる古典的言語野であるブローカ野 (BA44, BA45) や、縁上回 (BA40) などには、いわゆる意味はない。これはシンボル系列を処理するがその意味はこれらの言語や以外のところにある。例えば、色名の生成には色情報を処理する高次視覚野の V4 のすぐ前の領域が重要であるといわれる。単語の意味はこのようにその単語が指し示す属性を処理するところにあると考えられる。したがって、視覚処理の腹側系の周辺に名詞生成に重要な領域があるのである。この場合、腹側系の主な働きは、視覚情報に基づくカテゴリー化であり、このカテゴリー化された情報に名詞 (物の名前) がリンクされているのである。

### 4-5-3 言語獲得と処理のネットワークモデル

#### (1) 単文の学習と理解に脳内ネットワークモデル

Dominey, Inui, and Hoen は、言語獲得と理解の脳モデルを初めて提案した (14)。このモデルでは、47 野、45 野、44 野に沿った構造的な処理が明確に分けられないで、ある程度オーバーラップしながら機能分化があると考えている。これは、現実のデータに対応して仮定されたものであり、イメージングデータによる個人差が単に反映されているのか、現実に重なり合いがあるのかは不明である。ここでの主要なネットワークを図 1 に示す。BA47 はリカレントネットワークの構造をしており、これによって現在の closed class の単語のみならず、以前の closed class の単語の履歴を表現することができる。47 野は尾状核頭に投射し、線条体、視床を経由して BA44/45 に信号を伝達する。BA45 には現在の open class の単語が存在するが、文レベルでの統合された意味は Hagoort によって提案されたように、BA44/6 に表現される (15)。つまり、単語レベルの意味 (BA45) は統合され、適切な主題役割を付与されて BA44/6 に文レベルの意味に統合されると考える。

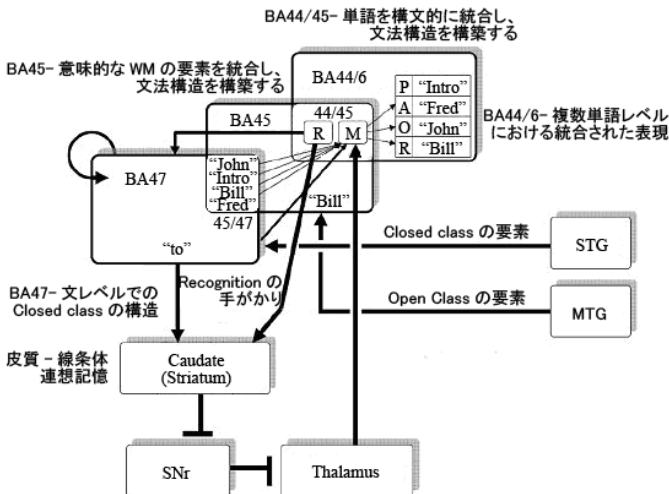


図 4・8 言語獲得と理解の脳モデル



ここで BA44 と BA45 には二つの機能的な領域があると考えられる。すなわち、BA44/45M (変調機能) は BA45 での意味内容を文レベルでの意味の文法的主題役割のフレームに変換し、BA44/6 に送る。したがって、視床皮質入力は、BA44/6 への結合を変更する。第 2 の機能的な領域は、BA44/45R (認識機能) であり、これは BA45 の Open Class の単語を認識し、BA44/6 で符号化された意味の認識を行う。したがって、これら二つの領域は、反対の役割をもっている。すなわち、学習時においては BA44/45R が現在の Open Class の要素に対して、主題役割の付与を検出する (例えば、BA45 における“Bill”が BA44/6 での受け手の役割と一致することを認識する)。これは、BA47 で符号化される文の文脈の中で行われる。

つまり、学習時においては、視覚によって与えられた状況に対して、教師が言葉でする説明すると考える。したがって視覚的に与えられた情報によって単語の Grounding が与えられるのである。この部分は図では省略されている。

一方、BA44/45M はバインディング機能、または対応する変調機能をもっている。すなわち、同じ文脈が BA47 で符号化されたとき、文の同じ種類の理解を行っているときに、BA44/6 における適切な役割に BA45 の要素がバインドされる。BA47 で符号化される文脈が、BA44/45M でのバインディング信号とリンクされるのである。理解時には BA47 の活動は BA44/45M の変調をかけ、現在の Open Class の単語をその対応する主題役割にバインドする。この場合、BA47 で符号化されたときと同じ文脈で、学習時には BA44/45R において認識された主題役割に対応している。

学習の状況を示そう。この場合 BA44/6 におけるイベントシステムは、観察し、符号化する。BA44/45R は BA45 の現在の Open Class の単語 (“Bill”) と BA44/6 のイベントの表現と比較する。そして、それが R (受け手) の役割と一致することを検出する。したがって、BA44/45R の信号は、BA47 における現在の活動を特定化する学習信号を与えることになる。

次に、人工文法の学習について述べる。この場合、BA45/47 のワーキングメモリには常に既に提示された系列要素のうち、7 個が記憶されている。BA44/45R の認識機能によって、現在の反応が BA45/47 の要素と比較される。そして BA47 は、認識機能から入力から入力をとるので、系列の抽象的な繰り返し構造を符号化したものになる。変調 (BA44/45M) ニューロンは BA45/47 の表現を出力する。この変調は BA47 から直接皮質/皮質経路を経たものと、長い尾状核を経た大脳基底核ループを通したものによって、変調される。系列の繰り返し構造を符号化している BA47 の現在の文脈活動を毎回認識機能によって BA44/45M(i)ニューロンと連合する。これは、認識された BA44/47 の内容、すなわち BA45/47i から出力に変調をかける。ABABC をトレーニングした後は、モデルはすべての系列を抽象的な構造 1-2-1-2-3 の形で認識する。または unique, unique, match n-2, match n-2 (detected by BA44/45R)。

Closed Class と Open Class の単語がそれぞれ STG と MTG から入力される。入力文が読み込まれると、BA47 を含むリカレントネットワークによって STG からの Closed Class の単語の順序が符号化される。BA45 は MTG から Open Class の単語を受け取る。BA44/6 は、構造的なワーキングメモリの役割をしており、述語、行為主、対象、受動者などの文法的な意味を表現する。47 野を含むリカレントネットワークの出力によって、BA44/45M は、BA44 からの Open Class の単語を BA44/6 の正しい位置に配置する。

## (2) 複文の学習・理解システム

Dominey and Boucher や Dominey, Hoen, and Inui は、複文を学習し理解するシステムを考案している<sup>16),17)</sup>。これはロボットのカメラで状況を理解させ、同時にその状況を音声で与えることにより、システムが文法を学習するシステムである。前述の Dominey, Inui, and Hoen や Dominey, Hoen, and Inui では、複文は扱えなかった<sup>14),17)</sup>。このシステムは、前述のモデルとほぼ類似した構成である。ここでも、Closed Class の単語と Open Class の単語がその順序とともに別ルートで処理される。図 4・8 の SceneEvent は、外界で起こる状況を記述したものである。それに対して、文が与えられると Closed Class の単語列情報を手がかりに Open Class の単語を SentToScene を通じて適切に対応付けを行う (Gouinding)。

例えば、Agent1 that verb1ed object2 verb2ed object3. という (複文の) 系列に対して、Action1(agent1, object2), Action2 (agent1, object3) という構造に変換できるように学習するのである。

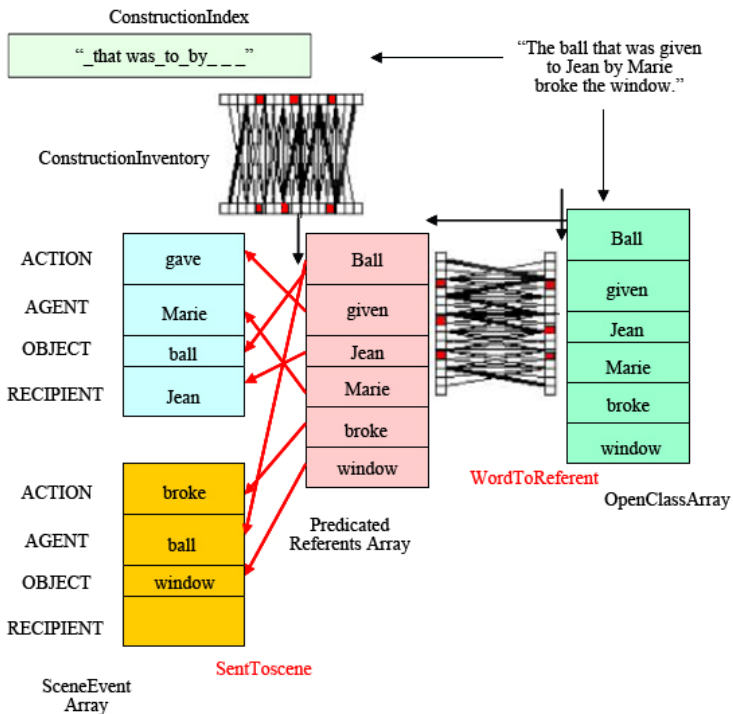


図 4・8 複文処理の学習と認識システム

## ■参考文献

- 1) H. Damasio, et al., “A neural basis for lexical retrieval,” *Nature*, Vol.380, pp.499-505, 1996.
- 2) A. R. Damasio and D. Tranel, “Nouns and verbs are retrieved with differently distributed neural systems,” *Proc. Natl. Acad. of Sci. USA*, Vol.90, pp.4957-4960, 1993.

- 3) N. Geschwind, F. A. Quadfasel, and J. M. Segarra, "Isolation of the speech area," *Neuropsychologia*, Vol.6, pp.327-340, 1968.
- 4) A. D. Friederici, "Towards a neural basis of auditory sentence processing," *Trends. Cogn. Sci.*, Vol.6, No.2, pp.78-84, 2002.
- 5) T. J. Grabowski, H. Damasio, and A. R. Damasio, "Premotor and prefrontal correlates of category-related lexical retrieval," *Neuroimage*, Vol.7, pp.232-243, 1998.
- 6) T. Inui, K. Ogawa, and M. Ohba, "Role of left inferior frontal gyrus in the processing of particles in Japanese," *Neuroreport*, Vol.18, 5, pp.431-434, 2007.
- 7) K. Ogawa, M. Ohba, and T. Inui, "Neural basis of syntactic processing of simple sentences in Japanese," *Neuroreport*, Vol.18, No.14, pp.1437-1441, 2007.
- 8) K. Ogawa, T. Inui, and M. Ohba, "Syntactic processing of complex sentences in left lateral premotor cortex," *Neuroreport*, in press, 2008.
- 9) S. Lehericy, et al., "Diffusion tensor fiber tracking shows distinct corticostriatal circuits in human," *Ann. Neurol.*, Vol. 55, No.4, pp.522-529, 2004.
- 10) M. Catani, D. K. Jones, and D. H. Ffytche, "Perisylvian language networks of the human brain," *Ann. Neurol.*, Vol.57, No.1, pp.8-16, 2005.
- 11) 乾敏郎, "運動系列予測学習仮説," *神経心理学*, Vol. 4, pp.144-149, 1998.
- 12) D. Bavelier, et al., "Sentence reading: a functional MRI study at 4 tesla," *J. Cogn. Neurosci.*, Vol.9, pp.664-686, 1997.
- 13) H. J. Neville, et al., "Neural systems mediating American Sign Language: Effects of sensory experience and age of acquisition," *Brain Lang.*, Vol.57, No.3, pp.285-308, 1997.
- 14) P. F. Dominey, T. Inui, and M. Hoen, "Neural network processing of natural language: II. Towards a unified model of corticostriatal function in sentence comprehension and non-linguistic sequencing," submitted, 2008.
- 15) P. Hagoort, "On Broca, brain, and blinding: a new framework," *Trends. Cogn. Sci.*, Vol.9, pp.416-423, 2005.
- 16) P. F. Dominey and J. D. Boucher, "Developmental stages of perception and language acquisition in a perceptually grounded robot," *Cogn. Sys. Res.*, Vol.6, pp.243-259, 2005.
- 17) P. F. Dominey, M. Hoen., and T. Inui, "A neurolinguistic model of grammatical construction processing," *J. Cogn. Neurosci.*, Vol.18, No.12, pp.1-20, 2006.

## ■脳と神経モデル-4章

---

### 4-6 言語と思考のモデル

※執筆中

## ■脳と神経モデル-4章

---

### 4-7 行動決定システム

※執筆中

## ■脳と神経モデル—4章

### 4-8 行動決定モデル

(執筆者：石井 信) [2009年4月 受領]

#### 4-8-1 報酬得依存の行動決定

ヒトをはじめとする動物は、自らの欲求を満足するような行動を選択することによって、自然環境において生き延びている。ある行動を取ったときに、環境からその行動の良し悪しを表す報酬（あるいは罰）が与えられるとして、より報酬が得られやすい（あるいは罰を回避しやすい）行動を選ぶと生き延びる可能性が高まる。こうした報酬に基づく行動決定は、動物の最もプリミティブな学習形態である。例えば、池のそばで手を叩くと泳いでいる鯉がよく寄って来る。鯉は、音がしたときに人影のそばに行くで報酬（餌）をもらえるという経験を繰り返すことで、この行動を学習していると考えられる。動物は、ある行動で良い結果が得られればその行動を取り続け、一方で悪い結果が得られれば別の行動を試みる。このように、行動とその結果との連合によって、行動の取り方が変わること「効果の法則」と呼ぶ。しかし、実際の環境ではしばしば、ある行動の直後に与えられる報酬のみではその行動が最終的に良かったのかどうかを判断できないことがある。

こうした場合には、行動の系列を考え、その系列をなす行動ごとに得られる報酬の累積和（累積報酬）を大きくするように行動系列を決定することで、環境に適合することができる。しかし、動物が行動決定する際に、その結果として得られる報酬はまだ得ておらず、未来の報酬に基づき現在の行動を決めることはできない。したがって、報酬（または累積報酬）を期待値として予測していることを仮定する必要がある。

単一の行動を考える場合は報酬の期待値、また、行動系列を考える場合は累積報酬の期待値を「価値」と呼ぶ。価値に基づく行動決定の理論では、この価値の学習こそが本質であり、学習された価値を最も大きくするものとして行動を決定する（「食欲な」行動という）ものとして、環境に適応した形での行動決定を説明する。すなわち、“同じような状況では同じような報酬が得られる”ということを繰り返し経験することで、価値を学習し、適切に行動を決定することができる。

#### 4-8-2 強化学習のモデル

強化学習は、経験に基づく累積報酬（価値）の学習により、環境に最も適応した行動系列を決定するためのモデルである。報酬が与えられる環境における行動系列の最適化は、マルコフ決定過程（Markov decision process : MDP）と呼ばれており、強化学習はMDPの解法の一つである。時刻  $t$  で環境の状態  $s(t)$  を観測した後で、エージェント（動物）はある方策（行動決定則） $\pi$  にしたがって行動  $a(t)$  を選択する（すなわち、 $a(t)=\pi(s(t))$ ）ものとし、その後状態  $s(t+1)$  に遷移する。このとき、状態及び行動の良し悪しに応じて報酬  $r(t+1)\equiv r(s(t),a(t))$  が与えられるものとする。このとき、時刻  $t$  からの報酬の減衰和（累積報酬）

$$\begin{aligned} R(t) &\equiv r(t+1) + \gamma r(t+2) + \gamma^2 r(t+3) + \dots \\ &= r(t+1) + \gamma R(t+1) \end{aligned} \quad (8 \cdot 1)$$

を考え、これに基づき方策  $\pi$  を最適化する。ここで、 $\gamma$  ( $0 \leq \gamma \leq 1$ ) は割引率と呼ばれ、環境の不確実性を考慮して、将来に得られる報酬ほど割引引いて評価するためのパラメータであ

る.  $\gamma=1$  のとき, 減衰和は単純和となり, 遠い将来に得られるであろう報酬も考慮したものになり, 一方で,  $\gamma=0$  のとき, 直近に得られる報酬  $r(t+1)$  のみを考慮したものになる. このことから  $\gamma$  がエージェントの衝動性に関連したパラメータであることが分かる<sup>2)</sup>.

強化学習では, 累積報酬  $R(t)$  の期待値を価値  $V \equiv E[R(t)]$  とする. 累積報酬の定義式 (8・1) の第二行に示した再帰的關係により,

$$V(s(t)) \sim r(s(t), a(t)) + \gamma V(s(t+1)) \quad (8 \cdot 2)$$

となる. 式 (8・2) では,  $V(t)$  が時刻  $t$  の状態  $s(t)$  に明に依存することを表現した. また,  $\sim$  は右辺と左辺が期待値として等しいことを意味する. この関係式 (8・2) は, ベルマン方程式と呼ばれることもある.

$V$  は状態  $s$  に依存するため価値関数と呼ばれる. ベルマン方程式は, エージェントが価値関数の学習を終了した際に, 時刻  $t$  での価値関数と時刻  $t+1$  での価値関数との間に成り立つ関係性を示したものであるが, 一方で, 価値関数はどのように学習すれば良いのだろうか.

ベルマン方程式 (8・2) の右辺と左辺との差

$$\delta(t) \equiv (r(s(t), a(t)) + \gamma V(s(t+1))) - V(s(t)) \quad (8 \cdot 3)$$

を時間差分 (Temporal Difference : TD) 誤差と呼び, 第一項は報酬  $r(s(t), a(t))$  を得た後での状態  $s(t)$  の価値を, 第二項は報酬を得る前での状態  $s(t)$  の価値の予測を表している. すなわち TD 誤差は, 時刻  $t$  で実際の報酬の獲得により評価される価値の予測誤差である. TD 誤差は各時刻での行動の結果として得られる. もし, 価値が正しく学習されれば, 式 (8・2) が成り立つため, TD 誤差は確率的に 0 である. 実際の経験に基づき, TD 誤差が小さくなるように, つまり式 (8・2) の左辺を右辺に近づけるように, デルタ則,

$$\Delta V(s(t)) = \alpha \delta(t) \quad (8 \cdot 4)$$

によって (正確には確率近似法により), 価値関数を学習するのが TD 学習である. ここで,  $\alpha$  ( $0 < \alpha \leq 1$ ) は学習係数である.

価値関数を学習することができれば, それを用いて各時刻での適切な行動を決定することができる. 簡単のため  $r(s(t), a)$  が  $a$  ごとに評価できるとすると, 式 (8・3) の TD 誤差が最大になる行動  $a(t)$  を選ぶことが考えられ, これは貪欲な方策に対応する. また, 価値関数の学習とは別に, TD 誤差にしたがって, TD 誤差が正となるような行動を再びとりやすく (正の強化), 一方で TD 誤差が負となるような行動を再びとりにくく (負の強化) するような学習を行うことも考えられる. この学習法を Actor-Critic 法<sup>1)</sup>と呼び, 価値関数の学習モジュールを Critic, 行動の強化モジュールを Actor という. ここで Actor-Critic 法では, 価値に基づき行動の強化が行われている点に注意したい.

さて, TD 学習は価値関数に対するデルタ則であり, 行動は価値関数に基づき決定される. この際に, 単純な貪欲方策を用いると, 特に学習初期に悪い価値関数に収束してしまった場合, もはや良い価値関数 (すなわち良い方策) を得ることができなくなる. また, 仮に良い価値関数を得ることができたとしても, その後に環境変化が起こった場合, 現在の価値関数を新しい環境での最適な価値関数へと修正することができなくなる. つまり環境変化に追従した行動決定ができなくなる. そのため, 強化学習の工学応用においてはしばしば, 貪欲方策に対してある程度のランダムさを加え, そのランダムさを状況に依存して加減するヒューリスティクスが用いられる. 例えば, 確率  $(1 - \epsilon)$  で貪欲な行動 (すなわち TD 誤差 (式 (8・3)) が最大の行動) を, 確率  $\epsilon$  でランダムな行動を選択する方策 ( $\epsilon$  貪欲方策) や, 各行動

を, TD 誤差の指数に比例するように選択するボルツマン方策などが用いられる. 前者の  $\varepsilon$  貪欲方策では,  $\varepsilon$  の大きさによりランダムさを加減でき, 後者のボルツマン方策では, TD 誤差に適当な定数をかけたうえで指数を計算するものとして, その定数 (ときに逆温度と呼ばれる) の小ささによりランダムさを加減することができる. この際に, 現在の価値関数に基づき貪欲な行動を選択することでより多くの報酬を得ようとするを「搾取」(exploitation), ランダムな行動を取ることで得られる報酬は少ないものの環境変化に適応しようとするを「探索」(exploration)と呼ぶ. 探索と搾取をいかに制御するかは強化学習における最も重要な課題の一つである.

#### 4-8-3 強化学習の脳内実現

Schultz ら<sup>3)</sup>は, 電気生理学実験によって, 一種の古典的條件付けを行うサルの中脳ドーパミン (dopamine : DA) 細胞の活動が, 式 (8・3) で与えられる TD 誤差と類似することを見いだした. 実験では, サルに光刺激 (条件刺激) を与え, その後で液体報酬 (無条件刺激) を与える試行を繰り返した. 試行の初期では, DA 細胞は報酬が与えられたときにみに反応して発火頻度が高まった. 式 (8・3) では, 当初,  $V(s(t+1))$  も  $V(s(t))$  もゼロなので, 正の報酬があれば, TD 誤差  $\delta$  は正となる. 一方, 条件付け学習が進行すると, DA 細胞は光刺激に対し発火頻度を高めるが, 液体報酬には反応しなくなる. 光刺激提示時には,  $V(s(t))$  はゼロ (すなわち刺激前には報酬は予測されない) が,  $V(s(t+1))$  は正 (刺激により正の報酬を予測する) となるため, 実際の報酬  $r$  がゼロでも TD 誤差は正となり, DA 細胞の反応を説明できる. 一方で, 報酬が与えられたときには,  $V(s(t+1))$  はゼロであるが,  $V(s(t))$  は  $r$  を十分に予測している (すなわち同じ値になっている) ため, TD 誤差はゼロとなり, DA 細胞が反応しないことを説明できる.

実験では更に, 条件付けにより刺激報酬の連合が学習された後で, 光刺激提示後に本来与えるべき報酬を与えないと, 報酬が与えられるべき時刻での DA 細胞の活動が自発発火よりも減少することを示した. この現象は,  $r$  と  $V(s(t+1))$  がともにゼロであるにもかかわらず,  $V(s(t))$  が正であり (光刺激により正の報酬を予測している), TD 誤差が負となるためである, と説明できる. このことから, DA 細胞の活動は TD 誤差を表現しているのではないかと考えられており, 多くの神経科学者が, DA 細胞の投射が多い大脳基底核や前頭前野と強化学習との関連を調べている.

また, そうした知見に基づく計算モデル (仮説) も提唱されている. 例えば, 大脳基底核の線条体において, 黒質緻密部の DA 細胞からの投射を多く受けている Striosome が Critic を, 一方で, 投射ニューロンを多く含む Matrix が Actor の役割を果たすことで, 線条体において Actor-Critic 学習が行われているのではないかと計算モデル<sup>4)</sup>がある. また, 環境が比較的単純であり条件付け学習で十分に説明できるような状況では, 大脳基底核において Actor-Critic 学習が行われるが, 一方で, 環境が複雑であり, 条件付け学習よりも, 状態の変化を予測 (先読み) することで適切な行動を決める方がよいような場合には, 前頭前野と大脳基底核を含む大規模な脳内回路が用いられる<sup>5)</sup>, などの仮説も提案されている. また, 環境に観測できない状態変数が存在する場合, その状態の不確かさを解消しながら行動決定が行う必要がある. これを部分観測マルコフ決定過程 (Partially Observable Markov Decision process : POMDP) という. POMDP を解く際の不確かさの解消には前頭前野の最前部がかか



わると考えられている<sup>6)</sup>。実際に、大脳基底核、前頭前野（眼窩前頭野など）、帯状回など脳の広い領域で、報酬依存の活動が観測される。しかし、そこで行われている計算過程の全貌は未だ明らかになっておらず、現在も研究が続けられている。

## ■参考文献

- 1) A.G. Barto, R.S. Sutton & C. W. Anderson, "Neuronlike elements that can solve difficult learning control problems," IEEE Transactions on Systems, Man, and Cybernetics, 13, 835-846, 1983.
- 2) K. Doya, "Metalearning and neuromodulation," Neural Networks, 15, 495-506, 2002.
- 3) W. Shultz, P. Dayan & P. R. Montague, "A neural substrate of prediction and reward," Science, 275, 1593-1599, 1997.
- 4) A. G. Barto, "Adaptive critics and the basal ganglia," In Models of Information Processing in the Basal Ganglia (eds. J. C. Houk, J. Davis & D. Beiser), Cambridge, MA: MIT Press, 215-232, 1995.
- 5) N. D. Daw, Y. Niv & P. Dayan, "Uncertainty-based competition between prefrontal and dorsolateral striatal systems for behavioral control," Nature Neuroscience, 8, 1704-1711, 2005.
- 6) W. Yoshida & S. Ishii, "Resolution of uncertainty in prefrontal cortex," Neuron, 50, 781-789, 2006.

## ■脳と神経モデル—4章

### 4-9 情動システム

(執筆者：木村 實) [2009年1月 受領]

#### 4-9-1 情動とは何か

「情動 (Emotion)」とは何か。喜怒哀楽で生じる心のはたらきであるが、「感情」と「情動」はどのように違うのであろうか。恐怖、怒り、悲しみ、喜びなどの感情にはその当人にしか分からない主観的なはたらきと、外部から観察可能なはたらきがある。後者は、感情に伴う心拍数の上昇や発汗作用などの自律神経系の活動の変化や、顔の表情や筋の緊張の変化などの身体的変化を測定することによって客観的にとらえることができる。更に、食べ物のおいしいにおいに誘われるように接近行動をとったり、逆にヘビや獣のような恐ろしい対象を目にするとそれから逃れようとする動機が生じる。したがって、「感情」を科学研究の対象として取り上げ、動物実験の成果をもとに心のはたらきを調べるためには、客観的に捉えることのできる「感情」の一つの概念である「情動」を研究することになる。

情動に関与する脳の仕組みを理解する研究は長い伝統があり、多くの知識が得られているが、情報表現とその処理の様式については多くの点で不明であり、今後、新しい研究戦略を導入することを含めて多方面からのアプローチによって発展が期待されている。一方、情動の重要な側面は、記憶、感覚認知、運動・行動の発現、価値判断や意志決定などの脳機能に欠かせない役割を担っていることであり、その役割と情報処理様式、神経回路や分子基盤を明らかにする研究が重要であり、その発展が望まれている。

#### 4-9-2 情動に関する仮説

##### (1) James-Lange 説 (1984, 1985)

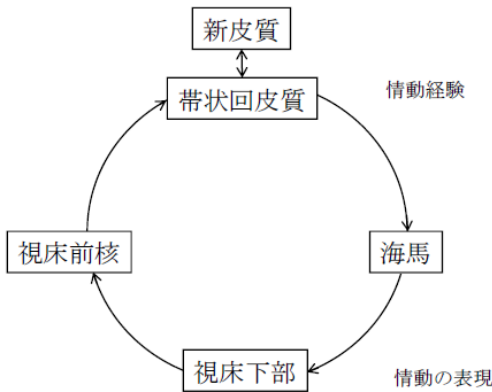
情動が生じるのは生体の生理学的な変化によると考えた。極端に言えば、私たちは泣くから悲しいのであって、悲しいから泣くわけではないと考えたのである。筋の緊張、心臓の拍動や表情そのものが情動であり、それらが取り除かれれば情動も消失すると考えた。しかし、情動が生体の生理学的な変化に緊密に関係しているというものの、心拍数の変化や表情などの変化が観察できないからといって、情動が生み出されないというわけではない。

##### (2) Cannon-Bard 説 (1927)

情動の表出がなくても情動を経験することができると考えた。実験動物が、脊髄の損傷によって下肢や体幹の感覚や運動ができなくなっても、上肢や頭部が正常に機能していれば情動を表出することができることを確認した。また、恐れ的情動は心拍数の増大、内臓の消化機能の低下や発汗の上昇を伴う一方、このような身体反応は、怒りや情動とかわりのない風邪の症状でもみられる。このような事実は James-Lange 説を支持しない。Cannon-Bard 説では、外界や体内の環境からの刺激が感覚系を介して受容され、視床を通して大脳皮質を活性化され、更に大脳皮質の活動によって視床が活性化されることによって情動が生じると考えた。

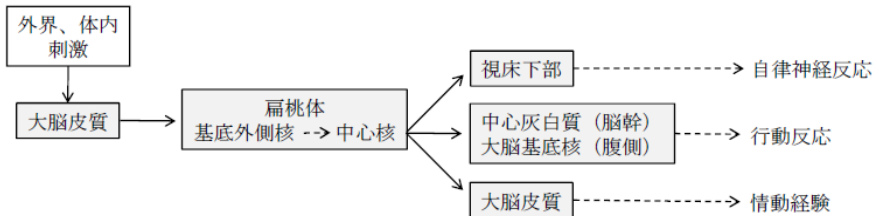
(3) 辺縁系仮説 (パペッツの情動回路)

Papez<sup>1)</sup> は脳の神経接続に関する研究成果と脳の損傷と症状の発現に関する研究成果に基づいて、情動のための神経回路仮説を提唱した (図 4・9)。情動回路では、視床下部乳頭体から中脳へ下る回路が情動表出の中枢である。外界及び体内の感覚情報は視床を介して大脳新皮質に投射されて分析されるが、一部は視床から視床下部へ、更に視床前核を経て大脳帯状回に投射され、この回路が主観的情動体験の座であると考えた。そして、帯状回皮質から海馬体を介して視床下部への投射によって情動体験が情動の表出に導かれると考えた。McLean<sup>2)</sup> は Papez の情動回路を「大脳辺縁系」と名付け、扁桃体、前頭眼窩皮質と側坐核を加えた。海馬で外界と体内 (内臓) からの刺激の連合が行われ、その生物学的評価が視床下部を通して表現されると考えた。



情動経験は帯状回皮質を中心に新皮質の活動によって生み出され、情動表現は視床下部によってコントロールされるとパペッツは考えた<sup>1)</sup>。

図 4・9 パペッツの情動回路 (辺縁系)



大脳皮質で認知された外界や体内の刺激は、扁桃体の基底外側核、中心核で処理された後、一部は視床下部に投射されて自律神経反応を生じ、一部は大脳基底核腹側部と脳幹の中心灰白質に投射されて行動反応を引き起こし、更に一部は帯状回皮質を中心とする大脳に投射されて情動体験を生み出す。

図 4・10 情動経験と表出の神経回路

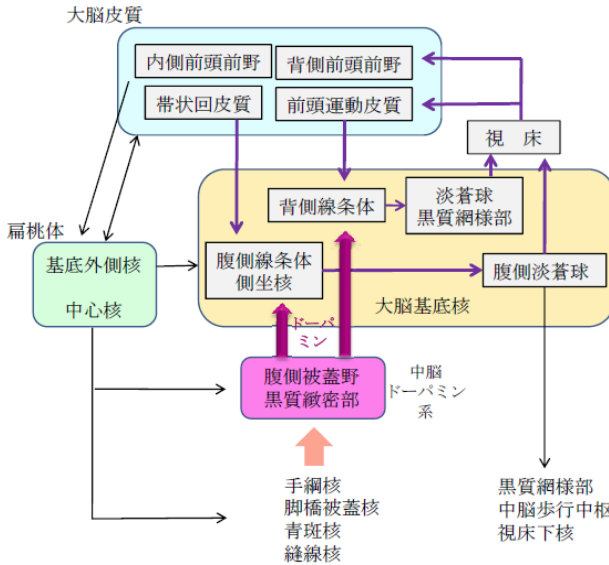
LeDoux<sup>3)</sup> は、外界と体内に生じる際だった情報を受容し、受容した情報の生物学的価値評価と意味の認知、その結果、喜怒哀楽の主観的体験と身体的、非身体的表の三つの連続する脳内情報処理によって、情動の諸現象が生じると考えた。生物学的評価とは、過去の体験や記憶に基づいて、外界と体内に生じる際だった情報が、自分にとって報酬または有益なものか、逆に嫌悪または有害かを判断する過程である。この過程には辺縁系の中でも特に扁桃体が重要な役割を果たす(図4・10)。扁桃体和側頭葉を破壊されたサルでは、刺激の生物学的価値判断が悪いので、危険なもの、通常は嫌悪感で避けるものに対して平気で接近してしまうことが知られている。

#### 4-9-3 ドーパミン細胞による大脳基底核の認知・行動機能の情動的コントロール

大脳基底核疾患は、パーキンソン病における筋の固縮、安静時振戦、ハンチントン舞蹈病やヘミパリスミアみられる四肢を中心とした顕著な不随意運動が大きな特徴である。もう一つ、あまり注目されていないが極めて重要な症状がある。それは、仮面様容顔として記述されているように喜怒哀楽の感情・情動が表情として現れないことである。Denny-Brown と Yanagisawa (1976)<sup>4)</sup> は線条体を破壊されたサルが自分から進んで探索行動をとることがないこと、何に対しても表情を変えることがなく、餌を見ても好奇心を表さないが口に入れば唾液を流すことなどを観察し、動物に欠如しているものは“activating “set” or “pump primer” for a certain act, the preparation for the mechanism preparatory to a motor performance oriented to the movement” であると結論している。また、パーキンソン病の患者さんの症状を見て Sacks (1983)<sup>5)</sup> は、“Some of them would sit for hours not only motionless, but apparently without any impulse to move, although they might move quite well if the stimulus or command or request to move came from another person. Such patients were said to have an absence of the will or “abulia.”” と記述した。また、Mettler (1945)<sup>6)</sup> の次の記述も強調点は共通である。“Animals lacking the striatum always display a certain fatuous, expressionless faces from which the eyes stare vacantly and with morbid intentness.”

大脳基底核は被殻、尾状核と腹側線条体(側坐核)からなる線条体、淡蒼球外節、淡蒼球内節、視床下核と黒質網様部、緻密部によって構成される。黒質は中脳にあるが、大脳基底核の神経回路を構成する主要な神経核であり、大脳基底核に含まれている。大脳基底核をめぐる神経回路のあらましを図4・10に示す。大脳基底核への主要な入力線条体に収束し、基底核の諸核を繋ぐ神経接続を介して特有の情報処理がなされた結果、基底核の出力情報は淡蒼球内節と黒質網様部から視床や上丘、脳幹に向けて送り出される。中脳の黒質緻密部(A9神経)と腹側被蓋野(A10神経)にはドーパミンを含有する神経細胞が集中して存在する。そしてA9神経の軸索は被殻と尾状核に投射する。A10神経は腹側線条体や扁桃体などの辺縁系に投射する中脳—辺縁系ドーパミン系と、前頭葉を中心とする広範囲の大脳皮質に投射する中脳—皮質系とがある。このうち線条体には脳内で最も高濃度のドーパミンが放出される。このようなドーパミンによる標的細胞の強固な支配は線条体では高いが、大脳皮質ではこれに比べて希薄である。ドーパミンの役割の大小を反映すると考えられる。

パーキンソン病に見られる著しい運動の障害は、ドーパミン系が運動のメカニズムに強力な作用をもつと考えさせたが、実際にドーパミン細胞の放電を調べると身体運動そのものに関連する活動は見られない<sup>7),8)</sup>。ドーパミン細胞の放電に影響を与えるのは身体運動ではなく、



大脳基底核は大脳皮質との間に、視床を介して皮質-基底核ループを形成し、緊密な機能連携がある。それらには、背側線条体ループと腹側線条体ループがあり、扁桃体は腹側線条体ループにかかわる。中脳のドーパミン細胞が背側線条体と腹側線条体に投射し、動機づけ、報酬や嫌悪の情報を伝達することによって、情動による認知・行動機能の調節を行っている。

図4・10 ドーパミン系による情動機能を中心とする大脳基底核系の機能的神経回路

動物にとって好ましい食べ物や飲み物などの報酬であることが分かった。Schultz<sup>9)</sup>ら、Sato<sup>8)</sup>ら、Bayer<sup>10)</sup>ら、Nakahara<sup>11)</sup>らはサルの上ドーパミン細胞の活動を精力的に調べ、思いがけず得られた報酬に対しては放電の増大を、期待どおりの報酬では放電を変化させず、期待していた報酬が得られなかったときには放電を減少させることを見出し、報酬予測の誤差情報を担うことを提唱した。報酬予測の誤差を教師信号とする学習理論である強化学習仮説のアルゴリズム<sup>9),12)</sup>そのものであるために、強化学習の大脳基底核モデルとして認められている。強化学習仮説の基本は、結果が出るまで試行錯誤で複数の動作を繰り返し行うことによって、それぞれの動作に価値を割り振ること。ここでいう価値とは、実際に報酬が得られなくてもその動作を行うことによってやがて報酬を得る確率が高くなる場合にはより高い価値(報酬)を与える(式(9・1)、(9・2))。したがって、現在から将来にわたって行う動作について、価値(つまり報酬)を推定して、次の動作を選ぶ場合にはトータルの報酬量が多くなるようにするというポリシーである。

$$P(t-1) = r(t) + \gamma r(t+1) + \gamma^2 r(t+2) + \dots \quad (9 \cdot 1)$$

$$P(t) = r(t+1) + \gamma r(t+2) + \gamma^2 r(t+3) + \dots \quad (9 \cdot 2)$$

$$P(t-1) = r(t) + \gamma P(t) \quad (9 \cdot 3)$$

$$\delta(t) = r(t) + \gamma P(t) - P(t-1) \quad (9 \cdot 4)$$

ここで、 $\delta$ : 時間差分 (TD) 誤差,  $r$ : 報酬,  $\gamma$ : 割引定数

将来の報酬は不確かなのでガンマ ( $\gamma$ ) で割り引いてある。式 (9・1), (9・2) から求められるデルタ ( $\delta$ ) は、報酬量の時間差分である報酬予測と実際の報酬とのズレ、つまり報酬予測誤差であり、これを教師信号とする学習である。あれこれと思案し、また試行錯誤によって突破口を見つけようとするときに、思わず、予想に反していい結果が得られた、逆に予想どおりの結果が得られなかったなどの、“報酬への期待の誤差”の情報が教師信号になっている。したがって、目標やお手本が示されているときに、ベクトル量で得られる誤差を使って目標値にいかにもスムーズに到達するか、精密な動作を実現するかという小脳型の教師付学習ではなく、“ひらめき”と“結果のよし悪し”に頼る試行錯誤によって、見えていない遠くの目標に到達するような学習といえよう。

ドーパミン細胞の担うもう一つの重要な情報は、報酬を予告する事象が現れたときにそれを獲得したいという気持ちの昂ぶり、獲得に向けた行動発現系へのポジティブなはたらきかけ<sup>8),13)</sup>である。動機づけ情報は、報酬に限らず嫌悪情報も含まれるが、中脳のドーパミン細胞には報酬の期待情報を表現する細胞だけでなく、嫌悪への反応を担う細胞も存在することが最近の研究によって明らかになった<sup>14)</sup>。このような動機づけ情報と報酬予測誤差の情報を担うドーパミン細胞は、線条体や大脳皮質にドーパミンを放出して、その部位でのシナプス伝達をコントロールする。ドーパミン系の線条体ニューロンに対する作用は大脳皮質線条体シナプス伝達のゲート作用に留まらず、シナプス伝達の可塑性をコントロールする学習への関与が重要である。線条体の投射細胞に対して皮質線条体投射が一時的に高頻度放電を送り込むこととドーパミン系の活性化が同時に生じると、皮質線条体シナプスの長期増強が生じる<sup>15),16)</sup>。この長期増強にはドーパミン D1 受容体の活性化が必要である<sup>15)</sup>。これは、報酬予測が不十分な学習の初期に、予想に反して好ましい結果が得られたときにはドーパミン細胞が放電を増加させて線条体の細胞にドーパミンを送り込むと、そのときに行った行動を生み出すことに関係する皮質線条体シナプス伝達を強化させるという強化学習の細胞メカニズムである。

このように、動機づけ情報と報酬への期待とその誤差という情動信号を中脳のドーパミン細胞が担い、大脳基底核の線条体、扁桃核、大脳皮質前頭葉に投射することにより、投射部位でなされる運動・行動や感覚認知、価値判断や意志決定などの幅広い基本的脳機能に情動的なコントロールをしている。ドーパミン細胞と大脳基底核の例で示したとおり、情動的なコントロールによって、目に見えていない遠くのゴールに向けて認知的、行動的な処理を可能にしたり、感覚系では、大脳皮質で分析された対象に喜怒哀楽の意味づけをすることに役立っている。

## ■参考文献

- 1) J. W. Papez, *J. Neuropsychiatry Clin, Neurosci.*, 7(1), 103, 1995.
- 2) P. D. McLean, “The limbic brain in relation to the psychoses,” *Physiological correlates of emotion.*, Academic press, 1970.
- 3) J. E. LeDoux, “*Annu Rev Neurosci.*,” 23, 155, 2000.
- 4) D. Denny-Brown and N. Yanagisawa, in *The Basal Ganglia*, edited by M. D. Yahr, Raven Press, New York, pp.115, 1976.
- 5) O. W. Sacks, *Awakenings.*, Dutton, New York, 1983.
- 6) F. A. Mettler, *J Neuropathol Exp Neurol*, 4, 99, 1945.

- 7) M. R. DeLong, M. D. Crutcher and A. P. Georgopoulos, *J Neurosci*, 3(8), 1599, 1983.
- 8) T. Satoh, S. Nakai, T. Sato and M. Kimura, *J Neurosci*, 23(30), 9913, 2003.
- 9) W. Schultz, P. Dayan and P. R. Montague, *Science*, 275(5306), 1593, 1997.
- 10) H. M. Bayer, and P. W. Glimcher, *Neuron*, 47(1), 129, 2005.
- 11) H. Nakahara, et al., *Neuron*, 41(2), 269, 2004.
- 12) R. S. Sutton, and A. G. Barto, "Reinforcement Learning," The MIT press, Cambridge, 1998 ; R. S. Sutton, *Mach Learn* 3, 9, 1988.
- 13) P. Redgrave, T. J. Prescott and K. Gurney, *Trends Neurosci.*, 22(4), 146, 1999 ; T. W. Robbins and B. J. Everitt, *Curr Opin Neurobiol*, 6(2), 228, 1996.
- 14) M. Matsumoto and O. Hikosaka, *Nature*, 459(7248), 837, 2009.
- 15) J. N. Reynolds, B. I. Hyland and J. R. Wickens, *Nature*, 413(6851), 67, 2001.
- 16) W. Shen, M. Flajolet, P. Greengard and D. J. Surmeier, *Science*, 321(5890), 848, 2008.

## ■脳と神経モデル—4章

### 4-10 情動システムモデル

(執筆者：鮫島和行，銅谷賢治) [2009年11月 受領]

#### 4-10-1 情動のシステムと計算論的アプローチ

私たちが主観的に感じる喜怒哀楽の「感情」は主観的感觉であり，科学として論じることが難しい．一方，自律神経系の活動や行動などから自然科学として研究することのできる対象として「情動」という言葉が使われる．危険がせまれば心拍数が上がり，手には汗を握る．このような自律神経系を賦活する現象は，動物などを使って客観的現象として医学分野で研究されてきた．また，不安や強迫観念などの「感情の障害」と薬物の関係は精神医学で扱われてきた．近年の非侵襲計測技術の発達によって，感情を誘発する状況と脳活動との対応関係が明らかにされつつある．しかし，主観的体験である感情と脳の対応部位のみで「情動」の機能的意味を本質的に理解することは困難である．

計算論的神経科学は，脳の機能を，(1) その機能が解いている問題設定，(2) 問題設定を解く方法，(3) 解く方法を実現する物理現象，の三つのレベルを一貫して，主に数理モデルを用いて理解するという脳研究の一つアプローチである．計算論的神経科学では，ヒトを含む生物が行っている情報処理を，その置かれている環境と目的とする望ましさを定式化し，望ましさを達成するアルゴリズムをモデル化し，そのアルゴリズムを実現する神経回路網や分子回路でおきる現象として理解しようとする．

計算論的アプローチでは，快や不快といった情動情報は，その情報が行動の発現や環境への適応に対する目的として機能する，と考える．例えば，生命維持に必要な食物などの探索行動や種の維持に欠かせない性行動の発現などは快情報による行動といえ，逆に捕食者との遭遇等による恐怖情報や体の物理的な損傷にもなう痛みなどの不快情報も生命や種の維持には不可欠である．快・不快情報は，不確実で変動する環境内で生物が生き残るために，自らの行動を環境に合わせる変化させる機能をもっていると考えることができる．すなわち，快情動を引き起こす刺激を再び得るような行動を強化し，不快な刺激には近づかないように行動を変化させる．このように，快・不快情動情報を報酬や罰としてとらえ，意思決定や行動へ反映させる問題は，従来の心理学，機械学習，経済学，行動生態学においてそれぞれ研究されてきた．しかし，近年，脳における報酬や価値の予測と意思決定の学習メカニズムの研究として様々な概念やモデルが融合しつつある．本稿では，これらの報酬に基づく意思決定の研究の中心的な枠組みを提供する計算理論として近年注目されている強化学習モデルと，脳における強化学習の実現方法の研究について紹介する．

#### 4-10-2 報酬の予測と Resocla-Wagner モデル

生存に不可欠な食べ物や飲みものなどの報酬を得ることを目的とする反応は古くから生理学や心理学においてモデル化されている．パブロフの犬で知られる古典的条件づけ (classical conditioning) は，ある刺激の後に報酬が得られると，報酬が得られた後に生得的に誘発される反応 (よだれなど) が，その刺激を提示するだけで誘発される現象である．これは，刺激を提示するだけで一定時間後に報酬が得られるということを予測するというモデルによって



説明される。Rescola-Wagner モデルは、その報酬予測が報酬予測誤差に応じて更新されてゆくアルゴリズムとして条件付けをモデル化している。すなわち、刺激  $s$  から報酬  $r$  が得られることを学習するために、刺激から報酬の予測  $v(s)$  を行い、その予測と結果のずれである報酬予測誤差  $\delta = r - v(s)$  を減少させるように  $v(s)$  を  $v(s) \leftarrow v(s) + \alpha\delta$  のように更新する。ここで  $\alpha$  は更新ステップの大きさを調節するパラメータである。古典的条件づけでは、ある刺激が報酬を予測するというモデルであるが、ある刺激に応じて行動が行われた条件で報酬が与えられる環境を経験すると、動物は刺激から誘発される行動を強化する。これを説明するモデルは、道具条件づけ (Instrumental Conditioning) として知られている。道具条件づけでは、刺激から行動への連関を直接強化するモデルと、刺激からまず報酬が関連付けられ、報酬を導く行動を連関させる  $s-r-o$  モデルが考えられている。

Rescola-Wagner モデルで説明される条件付けは、今日ではパブロフ型の条件付け、だけではなく、忘却や部分的刺激に対する条件付けや、二つの刺激の一方を条件付けした後に二つを同時に提示して学習しても、あらかじめ条件付けされていなかった刺激に対する報酬予測が学習されないブロッキング現象や、その他抑制条件付け、オーバーシャドウなど様々なバリエーションに対しての説明を与え、その適用範囲は広い。しかし、二次条件付け、すなわち刺激 A と報酬を条件付けした後、刺激 B と刺激 A を条件付けした場合、刺激 B から報酬を予測する現象は Rescola-Wagner モデルから直接には説明されない。二次条件付けを説明するためには、単に連合させるだけでなく、時間的な前後関係を導入する必要がある。

#### 4-10-3 状態遷移と強化学習

強化学習は、このような刺激、報酬、行動の関係を学習する枠組みを時間的に拡張した考え方もいえる。環境と学習者の中で状態、行動、報酬の三つの変数による相互作用を仮定する。このとき、現在の状態と報酬は環境によって規定され、過去の状態やそのときの行動に依存して決まる。学習者は環境から刺激を受けることによって環境の状態を知り、よりたくさんの報酬が得られるような状態へと導くような行動を行う。現在の環境の情報が完全に学習者に知らされているならば、状態や報酬は行動を行う直前の状態と行った行動のみで決定される。このとき環境がマルコフ決定性をもつという。このとき、学習者の最適な行動は、将来にわたって時間的に割引かれた報酬の総量  $V(s(t))$  を最大化する行動となる。

$$V(s(t)) = E[r(t) + \gamma r(t+1) + \gamma^2 r(t+2) + \dots | s(t)]$$

このとき、 $\gamma$  は将来の報酬に対する時間的割引パラメータと呼ぶ。

#### 4-10-4 報酬予測誤差と TD 学習則

では、どのようにある状態から将来にわたる報酬の総量、これを状態価値と呼ぶ、を見積もるのだろうか。その一つの学習アルゴリズムとして時間差誤差 (Temporal Difference Error) を使った評価伝搬による学習方法が知られている。ある時刻における状態  $s(t)$  から将来を予測した状態価値は  $V(s(t))$  は、その場で得られる報酬  $r(t)$  と次の時刻における状態  $s(t+1)$  から将来を予測した状態価値  $V(s(t+1))$  の和によって表現される。すなわち 1 時刻進んだ時間的割引を考慮して二つの時刻の価値には次の等式が成り立つ。

$$\begin{aligned}
 V(s(t)) &= E[r(t) + \gamma r(t+1) + \gamma^2 r(t+2) + L \mid s = s(t)] \\
 &= r(t) + \gamma E[r(t+1) + \gamma r(t+2) + L \mid s = s(t+1)] \\
 &= r(t) + \gamma V(s(t+1))
 \end{aligned}$$

状態の評価である状態価値が学習された後には、この等式の右辺と左辺の差である時間差誤差  $\delta = r(t) + \gamma V(s(t+1)) - V(s(t))$  が 0 となるはずである。そこで、時間差誤差を小さくする方向へ変更するように状態価値関数の近似関数を変更するアルゴリズム  $V(s(t)) \rightarrow V(s(t)) + \alpha \delta$  によって、状態価値関数を得ることができる。ここで  $\alpha$  は学習率であり、状態遷移を考慮しない場合は時間差誤差の第 2 項がなくなり Rescola-Wagner モデルと同値になる。

また、同様にして、ある時刻  $t$  における状態  $s(t)$  から行動  $a(t)$  を行った後に得られる報酬総和の予測値  $Q(s, a) = E[r(t) + \gamma r(t+1) + \gamma^2 r(t+2) + L \mid s = s(t), a = a(t)]$ 、これを状態-行動価値関数または行動価値関数と呼ぶ、も時間差誤差を定義することができる。ただし、行動価値の場合は、状態間の遷移を 1 時刻差とした場合、次の行動を行ってからの時間差誤差に基づいて更新を行う場合 (SARSA : State-Action-Reward-State-Action Algorithm) や、状態遷移の後に現在考える最善の行動の価値に基づいて更新を行う場合、

$$\delta = r(t) + \gamma \max_{a' \in A} Q(s(t+1), a') - Q(s(t), a(t)) \quad (\text{Q-learning})$$

など、種々のアルゴリズムが考えられる。

学習者の行動則 (Policy ; 状態から行動への変換  $a = \pi(s(t))$ ) は、行動価値を学習した場合は、現在の状態におけるよりよい選択を自然に導くことができる。例えば、完全に自己の行動価値が最適であると信じて行動価値最大となる行動を選択する。

$$a = \operatorname{argmax}_{a' \in A} (Q(s(t), a'))$$

または、ある行動価値に応じた確率でその行動をとるという方法

$$a \sim p(a) \propto e^{\beta Q(a)}$$

などを考えることができる。ここで  $\beta$  は行動の乱雑性を決めるパラメータである。ランダムな行動則からはじめて行動価値を推定し (Policy Evaluation)、その後このような行動則を適用して再度行動価値を推定する、推定された行動価値を使って更に行動則を求める、という繰り返し (Policy Iteration) を行うことによって、ある環境における最大報酬が得られる行動則を求めることができる。その他の方法として、状態価値関数と行動則とを独立に構成し、両者のパラメータを同時に状態価値から求められる時間差誤差によって学習する、という方法 (Actor-Critic Method) もある。

#### 4-10-5 脳における強化学習

大脳の深部に存在する神経核群である大脳基底核と中脳のドーパミン細胞は、このような報酬に基づく行動選択と評価の学習にかかわることがこれまで、動物を用いた神経生理学的研究やヒトを対象とする fMRI などの非侵襲脳活動計測を用いた研究で示されてきている。

大脳基底核は、大脳深部にある線条体、淡蒼球、黒質などの様々な核からなる核群の総称で、大脳皮質から入力を受け、視床を介して大脳皮質に帰るループ結合回路をもつ。線条体は大脳皮質の広範な領域から入力を受け、淡蒼球内節/外節や黒漆網様部を通じて視床の一部を抑制/脱抑制する。視床は大脳皮質と結合をもつことや、ある皮質の特定領域が大脳基

基底核を通じて同じ皮質領域へその結合を返す閉ループ回路の存在が指摘されている。

この大脳皮質と大脳基底核を結ぶループ回路が強化学習を実現しているのではないかという仮説が提唱されている。その根拠として、中脳ドーパミン細胞が TD 誤差に非常に似た活動を示すこと、その中脳ドーパミン細胞の主なターゲットである線条体において、報酬の有無/量/確率に応じた神経活動が発見されていること、ドーパミンの濃度に依存したシナプス可塑性が観測されていること、古典的条件付けや道具条件付けなどの学習中の報酬予測誤差と脳の血流変化との相関を fMRI によって求めると線条体と相関が認められるなどがあげられる。これらの動物やヒトを用いた神経生理学的な知見から Barto らは大脳基底核において価値関数と方策の学習が時間差誤差によって学習される Actor-critic モデルを提案している。また、Doya は線条体の神経活動が状態価値と行動価値を表現し、行動価値に依存して大脳基底核の出力核である淡蒼球で行動選択が行われるのではないかという仮説を提案している。このモデルでは、感覚野のみならず連合野において環境の状態が構成されその状態に応じて状態の価値、またはこれから選択する行動の行動価値を線条体の投射細胞が表現する。投射細胞が淡蒼球に情報を送る際に競合がおき、それら行動価値から確率的に行動選択がおき、その選択結果を大脳皮質に戻すことで行動の実行が行われる、というモデルである。その後得られる報酬によって報酬予測誤差がドーパミン細胞によって出力され、線条体にある皮質-線条体シナプスにおいてドーパミン濃度が高いときには強化が、低いときには減弱されることによって行動価値関数を学習する。線条体は、行動そのものを出力するのだろうか、それとも行動価値を出力するのだろうか、この疑問に答える実験を筆者らは行った。サルに自由に二つの行動から一つを選択させる課題を訓練し、その選択の後に与える報酬を行動ごとに確率的とした。行動選択はレバーを右または左に倒すことで行わせる。例えば右を行った後は9割の確率で報酬を与えるが、左を行った場合は5割とする。するとサルは試行錯誤を十試行ほど繰り返した後より多く報酬を得られる選択を行うようになる。すなわち、報酬の経験を通じて行動選択を学習する強化学習課題となっている。それを更に一定期間ごとに変動させたところ、サルは報酬の確率に応じて行動を変化させた。このとき、左レバー倒しのときの報酬確率に応じて反応を変化させるが、右レバー倒しのときの報酬確率には変化が見られない細胞、すなわち左の行動価値を表現できる細胞が多数見つかった。このことから、線条体では報酬に応じた行動価値が学習され、それをを用いて行動選択が行われていることが示唆される。

#### 4-10-6 強化学習のメタパラメータと神経修飾物質仮説

前述のように強化学習を実現するモデルでは、学習される価値関数  $V(s)$ ,  $Q(s, a)$  や方策  $\pi(s)$  などを表現するパラメータが学習される。直接的に学習されるパラメータ以外にも、学習そのものを制御するメタパラメータがある。外界の情報を受け入れる感受性を制御する学習率パラメータ  $\alpha$ 、未知の環境や行動への情報探索の強さを決める探索パラメータ  $\beta$ 、外界の未来に得られる報酬の感受性を制御する割引率パラメータ  $\gamma$  である。Doya はこれらの学習や行動を制御するメタパラメータが脳内において神経修飾物質に対応し、脳においてこれらの神経修飾物質の制御が環境に適応し生存を可能にする学習に重要である、という仮説を提案している。この仮説では、学習率がアセチルコリンに、逆温度がノルアドレナリンに、時間的割引率がセロトニンにそれぞれ対応している。

### (1) アセチルコリンと記憶の更新

アセチルコリンは海馬や大脳皮質における神経生理学的な研究で細胞や神経回路のレベルで記憶の更新と読み出しの制御にかかわっているのではないかということが示唆されている。また、マイネルト核のコリン作動性細胞が死滅すると、記憶に障害が生じることも知られている。線条体において、報酬予測誤差の信号がコリン作動性のインターニューロンにおいて観測されていることから、皮質線条体シナプスのドーパミン依存性可塑性に何らかの影響を与える可能性も考えられる。

### (2) ノルアドレナリンと探索/利用のバランス

既にこれまで学習された価値をもとに最良の選択をするのか、それとも確率的な行動を行い新たな情報を得て学習を促進するのか、という問題は「探索と利用のジレンマ問題 (Exploration and Exploitation Dilemma)」として知られている。この探索の度合いは、Softmax 行動選択戦略を用いたときには、行動価値から行動選択確率への変換関数の傾き  $\beta$  によって制御される。ノルアドレナリンは、覚醒-弛緩レベルに関係していることが知られている。また、サルの実験などでノルアドレナリン作動性細胞のある LC 核の神経細胞が刺激予測の正確な反応にかかわることが分かっている。脳内のノルアドレナリンレベルは行動選択の正確性にかかわる可能性がある。

### (3) セロトニンと将来の時間的割引

価値関数は将来に得られる報酬の総和であるが、その総和は将来へ行けば行く程割引かれる。これは生物としては死ぬまでに食べ物を得なければならないという制約や、環境が常に変動していてあまり将来のことを正確に予測していてもその意味がないためその報酬量を割引いて考えるというモデルである。Tanaka らはヒトを対象とした強化学習課題において様々な割引率で求めた価値関数と脳活動との相関を求めたところ、線条体と島皮質にその相関が認められ、前方-腹側へ行くにしたがって低い  $\gamma$  (短い時間スケール)、後方-背側がより大きな  $\gamma$  (長い時間スケール) に対応した状態価値に対応することを見いだした。また、セロトニンの前駆物質であるトリプトファンを経口でコントロールし、被験者に報酬までの時間間隔が異なる選択課題を行わせ脳活動を計測したところ、トリプトファンを多く接種したグループの背側の線条体がより価値に相関し、逆にトリプトファンを少なく接種したグループは腹側の線条体がより価値に相関した。セロトニンのトランスポータの拮抗薬は鬱病に効果があることなどが知られており、将来の報酬に対する行動発現、意欲の制御にかかわっている。大脳基底核を含む脳のセロトニンレベルは将来の報酬に対する割引を制御する可能性がある。

## 4-10-7 環境や生物の性質とメタパラメータ

これらのメタパラメータは独立に制御されるとは限らない。むしろ環境内で生存確率をあげるために要請されている学習の最適性から考えると互いに関係して制御されるのが望ましいだろう。環境の変動性 (Volatility) がよい例である。ある一定の環境 (すなわち状態遷移と報酬確率が定常) において学習を行いこれが全く変動しない場合、学習率は経験数に反比例して減衰させるのが最適となり、同様に探索も学習初期にはランダムに後期には価値の最

大をとる貪欲戦略が最適となる。ところが自然環境であれば、同じ状況で同じ行動を行っても同じ報酬が得られ続けるのではなく変動する可能性がある。この場合、新しい環境を検出して学習率を制御すると同時に探索も増やす必要がある。Beherens らは、探索性 (Volatility) の検出と学習率の制御に前帯状回皮質が重要な役割を果たすことをヒトの脳活動から示している。変動性とも関係するが、環境に対する将来予測性 (Predictability) も学習率と将来割引の両方に関与するだろう。将来の予測がより確実な場合には学習率を高くとることで適応を早くし、同時に将来の時間スケールを長くとることで、短期的には損でも長期的には得をする戦略をとることができるが、将来の状態がランダムである場合は学習により時間がかかり、長期的に得をする戦略に意味がなくなる。実際、ランダムに遷移する環境における行動選択から割引率を推定すると、決定論的に推移する場合より低く推定される。また、決定論的に遷移する場合には、腹側と背側の両方の線条体が報酬予測に相関するが、ランダムに遷移する場合には腹側の線条体のみが相関する (Tanaka 2006)。

## ■脳と神経モデル—4章

### 4-11 運動系システム

(執筆者：松波謙一) [2008年5月 受領]

運動系システムについてモデルを絡めて記述する。この観点から Hodgkin-Huxley モデルは神経生理学の基盤となる<sup>1)</sup>。最初に中枢神経の可塑性で重要な、樹状突起とスパイン (棘; spine) についてモデルで述べる。モデルで述べるのは、生理学的データがまだ少ないと思うからである。

#### 4-11-1 Core Conductor Theory 及び Hodgkin-Huxley (H-H) モデル

Core Conductor Theory については詳細な総説がある<sup>2)</sup>。その基礎となるケーブル方程式は2本の偏微分方程式で表される電信方程式から導かれる。インダクタンスと遺漏電流がない場合が理想ケーブル (トムソン・ケーブル) である。神経線維ではインダクタンス (=0) を無視し、下記の所謂ケーブル方程式が得られる (式 (4・1))。これを距離定数で規格化した式などの詳しい解説がある<sup>3)</sup>。また、H-H 方程式を拡張し、44のパラメータを駆使した心筋の KYOTO モデルもある<sup>4)</sup>。

$$\lambda^2 \frac{\partial^2 V_m}{\partial x^2} = \tau \frac{\partial V_m}{\partial t} + V_m ; \quad \lambda = \sqrt{\frac{r_m}{r_i}} \quad (\text{Space Constant}), \quad \tau = C_m r_m \quad (\text{Time Constant}) \quad (11 \cdot 1)$$

ここで、 $V_m$  = 膜電位、 $r_i$  = Core Resistance [KΩ/cm]、 $r_m$  = 膜抵抗 [KΩ・cm]、 $g_m$  = 膜コンダクタンス [mS/cm]、 $C_m$  = 膜容量 [μF/cm] である。

日本語による H-H モデルについては、優れた解説書があるので詳述は避ける<sup>5)</sup>。ここでは、数学的に新しい表現で述べられている岡本によって記述された方法について述べる (岡本教授からは有益なコメントを頂いた。改めて謝意を表したい)<sup>6)</sup>。

$$J_m = C_m \frac{dV_m}{dt} + I_{ion}, \quad I_{ion} = i_{Na} + i_K + i_L \quad (11 \cdot 2)$$

$$i_{Na} = g_{Na} m^3 h (V_m - V_{Na}), \quad i_K = g_K n^4 (V_m - V_K), \quad i_L = g_L (V_m - V_L) \quad (11 \cdot 3)$$

$$\frac{dx}{dt} = \alpha_x (V_m) (1-x) - \beta_x (V_m) x \quad (x = m, h, n) \quad (11 \cdot 4)$$

ただし、外部刺激電流 ( $J_{ext}$ ) がある場合は、式 (11・2) において  $J_m = J_{ext}$  とする。また、 $\alpha_x, \beta_x$  は Rate Constant に相当するもので、例えば  $m$  については

$$\alpha_m (V_m) = \frac{0.1(V_m + 45.68)}{1 - e^{-(V_m + 45.68)/10}}, \quad \beta_m (V_m) = 4e^{-(V_m + 70.68)/18}$$

であり、 $h, n$  についても類似の式で表わされる。

また、各定数の標準値としては下記の数値が用いられている。

$$C_m = 1[\mu\text{F}/\text{cm}^2], \quad V_{Na} = 53.4[\text{mV}], \quad V_K = -80.5[\text{mV}], \quad V_L = -59.5[\text{mV}] \\ g_{Na} = 120[\text{mS}/\text{cm}^2], \quad g_K = 40[\text{mS}/\text{cm}^2], \quad g_L = 0.24[\text{mS}/\text{cm}^2]$$

$V_{Na}, V_K, V_L, g_{Na}, g_K, g_L$  は定数であるので、 $I_{ion}$  の値は  $V_m, m, h, n$  を指定すれば、確定する。した

がって、 $I_{ion}$ を $I_{ion}(V_m, m, h, n)$ と表記できる。また、 $V_m, m, h, n$ は時間の関数であるので、 $I_{ion}$ も $V_m, m, h, n$ を介して時間の関数となる。一方、外部刺激電流( $J_{ext}$ )は既知の時間の関数であるので、

$$f_V(t, V_m, m, h, n) = \{J_{ext}(t) - I_{ion}(V_m, m, h, n)\} / C_m \quad (11 \cdot 5)$$

は $V_m, m, h, n$ および時間( $t$ )に関する既知の関数になり、式(11・2)は

$$\frac{dx}{dt} = f_x(t, V_m, m, h, n) \quad (11 \cdot 6)$$

と書ける。同様に式(11・4)は

$$\frac{dx}{dt} = f_x(t, V_m, m, h, n) \quad (x = m, h, n) \quad (11 \cdot 7)$$

したがって、

$$f_x(V_m, m, h, n) = \alpha_x(V_m)(1-x) - \beta_x(V_m)x \quad (x = m, h, n) \quad (11 \cdot 8)$$

は $V_m, m, h, n$ による既知の関数となる。 $V_m, m, h, n$ を並べ4次元ベクトル $\mathbf{x}$ とし、この $\mathbf{x}$ および時間 $t$ に依存して決まる4次元ベクトル $\mathbf{f}(t, \mathbf{x})$ を

$$\mathbf{x} = \begin{pmatrix} V_m \\ m \\ h \\ n \end{pmatrix}, \quad \mathbf{f}(t, \mathbf{x}) = \begin{pmatrix} f_V(t, V_m, m, h, n) \\ f_m(m, h, n) \\ f_h(m, h, n) \\ f_n(m, h, n) \end{pmatrix} \quad (11 \cdot 9)$$

と定義すれば、式(11・2)(11・4)は

$$\frac{d\mathbf{x}}{dt} = \begin{pmatrix} dV_m/dt \\ dm/dt \\ dh/dt \\ dn/dt \end{pmatrix} = \begin{pmatrix} f_V(t, V_m, m, h, n) \\ f_m(m, h, n) \\ f_h(m, h, n) \\ f_n(m, h, n) \end{pmatrix} = \mathbf{f}(t, \mathbf{x})$$

すなわち

$$\frac{d\mathbf{x}}{dt} = \mathbf{f}(t, \mathbf{x}) \quad (11 \cdot 10)$$

と表現できる(外部刺激電流がなければ $d\mathbf{x}/dt = \mathbf{f}(\mathbf{x})$ となる)。

#### 4-11-2 樹状突起についての Passive Membrane モデル

脳の錐体細胞などで見られるように、神経細胞(ニューロン)は多数の樹状突起をもち、それに1万から2万のシナプスがついている。この構造はニューロンの連合作用や学習に重要と考えられてきた(例えば、Hebb, 1949)。Rallはこの樹状突起やスパインに注目しモデルを提出した<sup>7),8)</sup>。彼はまず、樹状突起をPassive Membraneとみなし、 $n$ 本の樹状突起を1本の等価のケーブルに変換した一様分布モデル(Distributed Model)を提出した(図4・11A)<sup>7)</sup>。その際、ケーブル方程式に相当する下記の式が導出される。式中、 $Z$ はElectrotonic Distance、 $\lambda$ は(ここでは)Characteristic Length(特性距離)であり、 $dZ/dx = 1/\lambda$ の関係がある。

$$d^2V/dZ^2 = V + dV/dt \quad (4 \cdot 11)$$

また、導出の途中において、インピーダンス・マッチングに相当する下記の式が成り立つことを要求する(二分の三乗則)。詳細は文献に当たられたい<sup>7),8)</sup>。

$$r^{3/2}n(dS/dx)^{1/2} = \text{定数} \quad (11 \cdot 12)$$

しかし、抵抗や容量などが一様に分布するという一様分布モデルは計算量が多くなり、(当時の)コンピュータでは計算時間がかかった。そこで Rall は、更に条件を付けた Compartment Model (分室モデル) を提出した。これは集中定数回路に相当する。分室の数は 10 個とした (図 4・11A)。すると  $i$  番目の分室について (ただし、分室  $j$  から  $i$  についての速度定数を  $\mu_{ij}$  とする)。

$$dv_i/dt = \sum_j \mu_{ij} v_j + f_i ; \text{ただし, } f_i = (\varepsilon_i + \beta J_i + \chi_i) / \tau$$

各 1 個の分室を分布モデルの  $\Delta Z$  の増加に対応させてあるので(特に、分室が一系列の場合は、 $\Delta Z = 0.2$ )、 $\mu_{ji} = \tau^{-1}(\Delta Z)^{-2}$  となる。また、 $C_i$  は  $\Delta Z$  の長さの容量、 $g_{ij}$  はコンダクタンスとなる。そして、刺激パルス  $\varepsilon, J$  (EPSP, IPSP に相当) を与えて計算すると (図 4・11A 左) が得られる。

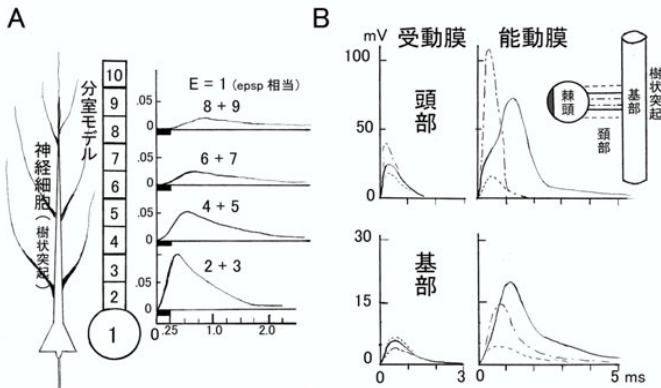


図 4・11

A: 神経細胞樹状突起の分室モデル (1~10 分室)。細胞体(1)近傍の分室(2,3)に興奮性入力(横太線)した場合の EPSP は急峻で大きく、樹状突起末梢部(8,9)の入力では鈍で低振幅である。

B: Spine (棘) 膜が受動膜・能動膜部 (黒色) の有無により、棘頭部と基部での EPSP・活動電位の大きさ・時間経過に及ぼす差異を示す (文献 7), 10), 11) を参照し改変)。

#### 4-11-3 Spine (棘) の Active Membrane モデル

その後、樹状突起については Active Membrane (能動膜) を取り入れたモデルが主流となる<sup>9)~11)</sup>。さて、スパイン表面積は  $1.5\mu\text{m}^2$  とし、能動膜はその 12.5% で Spine Head (棘頭) のシナプス下膜にあるとして結果を定性的に述べる。まず Spine Stem (棘茎) の抵抗 ( $R_s$ ) が増加すると、一過性の電圧変化 (EPSP 相当) に対応する活動電位発生の閾値は下がる (その関係はほぼ直角双曲線で近似的できる)。またあらかじめ樹状突起に電圧を負荷しておけば、EPSP の時間経過は (ない場合に比べて) 変わる。その際、付加電圧を Spine Base (棘底) からの距離に関係なく一様に (10mV で) 分布した場合に比べて、棘からの距離に比例して



(5mV/λの勾配で)電圧が増すと、EPSPの立ち上がりは早くなる。逆に-5mV/λの勾配で減少させると、EPSPは抑制される。その影響は棘頭で強く棘底では弱い傾向としては等しい。また、棘茎の長さや太さが樹状突起に生じるPSPに与える影響については興味をもたれてきたが、能動膜を仮定した場合には、非線形が生じる；すなわち、EPSPの増幅を最大にする最適値が存在する。

次いで、棘頭部におけるH-Hチャンネルの密度(5\*H-Hと10\*H-H)の影響について考えよう( $R_{ss}=1000M\Omega$ ,  $g_{Na}=120mS/cm^2$ ,  $g_K=36mS/cm^2$ ,  $g_L=0.3mS/cm^2$ ,  $R_m=1400\Omega/cm^2$ とする)。棘頭における活動電位(AP)は、10\*H-Hの方が5\*H-Hの場合より立ち上がりが急峻であり振幅も大きい。しかし、棘底では両者の差は極めて少ない。また、これからも予想されるように、コンダクタンスの違い(例えば、0.2nSと0.4nS)も極めて大きな差異を棘頭でのAPやEPSPに起こす。しかし、棘底のEPSPでも差が生じる。また受動膜と比較すると、能動膜(0.4nS)の場合と比べ、受動膜ではAP、EPSP共に極めて小さくなる。この違いは能動膜で見られる非線形性をもたらす増幅作用による。

EPSPとIPSPの時間関係は重要である。能動性棘( $R_{ss}=630M\Omega$ )で、IPSPがEPSPと同時に起きたときには、予想されるように、抑制効果が一番大きい。そして、与えるIPSPの時間がこれより前後にずれていくと、抑制効果は減弱する。グラフ的にはV字曲線を描く。しかし、一つ非常に重要な現象が生じる。IPSPをある時点より早めに生起させると、EPSPに対して促進的に働く。これはH-H式での、過分極に対する、 $n$ ,  $h$ (不活化係数)の振る舞いに帰因する。いわゆる、“過分極によりInactivationが外される”という現象である。別の一因は $m$ にかかわる外向きK電流の減少である。これらによりPost-inhibitory Reboundも生じる。そしてこの抑制の時間的効果は $R_{ss}$ の影響を受け、 $R_{ss}=1000M\Omega$ にすると、V字の幅は狭くなり、同時に、IPSPの促進現象もなくなる。

その他に、興奮性チャンネル(H-Hチャンネル)の“経済性”について考察されている。つまり、一定数のチャンネル(このモデルでは10\*H-Hチャンネルで、換算すると約1050個のNaチャンネル数)がある場合、10\*H-H個を一箇所に集中した方が効果的か、それとも5\*H-H個ずつ分散したほうが効果的かという問題である。結論から言うと、能動的な棘頭での活動電位には大きな違いを生じない(Naの平衡電位から予想できる)。一方、棘底でのEPSPには大きな違いを起こす。一個の棘頭に10\*H-H個を集中させると、受動膜時の0.9mVに比べ、5.1mVと5.7倍(5.1/0.9)の振幅増になる。一方、2個の棘に5\*H-Hずつ分散した場合は、EPSPの振幅は8.4mVと大きい、受動膜でも1.8mVであるので増幅比は4.7(8.4/1.8)となる。これは受動膜の場合では隣の棘から電気緊張的に伝わってくる膜電位が上乘せされる結果、受動膜の場合のEPSPが1.8mVとなるからである。膜電位がNaチャンネルのパラメータ( $n$ ,  $m$ ,  $h$ )に影響を与えることが分かっているから、上記の現象は、とりもなおさず、Hebbの意味での(短期的な)シナプス効率を変えたことになる。彼らは詳しい“Discussion”を展開しているが、それは原典に当たられたい。その他、Caイオンの関与を考えたモデルも提出されている。

#### 4-11-4 システムとしての脳(大脳)

今まで述べてきた樹状突起、棘、イオンチャンネルはミクロのレベルであった。これは神経系の機能の基盤であるが、ここで話を一転し、脳の“マクロ”のレベルについて話していこう。大脳では細胞構築学的違いに基づくBrodmann(1914;完成版)やSarkisov(1955)の

脳地図 (図 4・12 左) があり、場所による機能の違いがある。運動野についてみると、“身体部位の再現”があり、身体各部が運動野の表面に体部位別に割り当てられている。更に微視的に見ると直径  $500\mu\text{m}$  の円柱構造 (機能円柱) を単位としている (例えば Asanuma H; 1967)。機能円柱は 1 万前後のニューロンで充填される。ニューロンの解剖学的微細構造 (細胞体・樹状突起・棘・軸索) は篠田らによる HRP トレーサー法により見事に示されている。

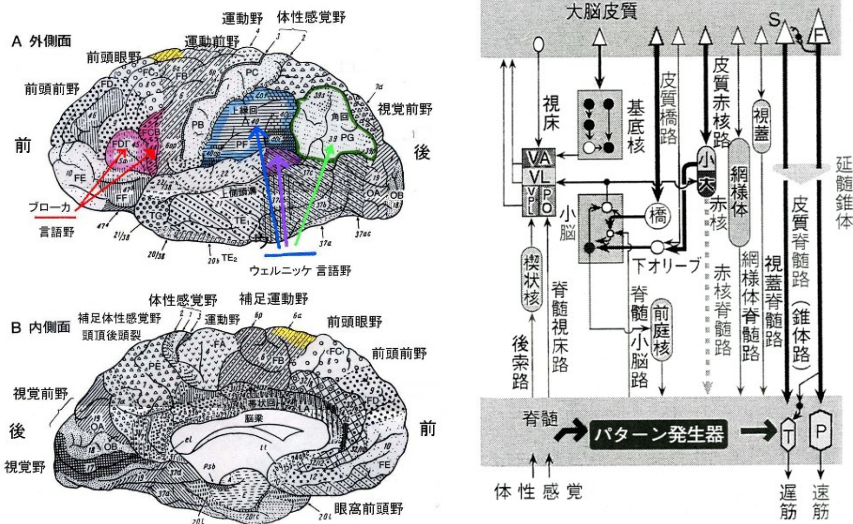


図 4・12 運動野および出力系

左図は大脳皮質細胞構築図の外側面(A)と内側面(B); (Sarkissov, 1955; 改変)。右図は運動野からの出力系。これは背外側系 (外側皮質脊髄路と赤核脊髄路) および腹内側系 (内側皮質脊髄路と網様体脊髄路) に大別できる。前者は遠位筋群 (手と指) を対側支配し、後者は近位筋群 (腕・腿と躯幹) を両側性に支配する。赤核は赤核脊髄路 (ヒトでは痕跡的) を形成する運動細胞型の大細胞性赤核 (サルでも赤核全体の 1/3) と下オリーブ核に投射する小細胞性赤核がある。後者は橋核とともにヒトでは大きく発達する。

運動野からの出力系を見よう (図 4・12 右)。従来、臨床例や研究・観察のやりやすさから錐体路 (細胞) が重視されてきた。運動時にこれら出力系ニューロンがどのように活動するかは、Evarts (1968), Yokota & MacLean (1968) が導入発展させたサルの慢性実験で示される。すなわち、運動野・運動前野・補足運動野・前頭前野・帯状回ニューロン活動が運動・行動との文脈で記録され関連付けられた。概観的に述べるが、詳しくは松波・内藤を参照されたい<sup>12)</sup>。

運動野については Evarts (1968) の錐体路細胞 (PTN) での功績が大きい。すなわち、PTN は(1) 運動に先行して活動し、負荷の大きさに比例して活動度が増す、(2) 準備 (set) の時期から活動を開始する、(3) 皮質経路で反射に関係する、(4) 精密な (小さな) 運動では  $\gamma$  系からのフィードバックかかると活動度が上がる。そして皮質運動細胞は (キネティックな) 運

動のパラメータ (位置・速度・加速度) に関係し、これらを集団的コード化 (Population Coding) をしている。運動前野・補足運動野・帯状回運動野においても、運動そのものに関係するニューロンが多数あるが、それよりは運動のダイナミクスとか set (準備) 期間に活動するニューロンが増えてくる。更に吻側にある前頭前野に至っては、古くは短期記憶との関係で語られたように、運動・行動のストラタジーの記憶 (作業記憶) との兼ね合いで理解されるべきである。また、頭頂葉の破壊は運動失行をもたらすが、前頭前野の破壊でも生じる。このことは、前頭葉と頭頂葉を連結する前頭後頭束など長軸索線維の重要性を浮上させる (これら連合・交連線維の障害が超伝導性障害)。つまり、運動失行とは学習して形成された、“運動記憶の喪失”として理解可能である。統合失調症が作業記憶と関連付けられ、前頭前野に責を帰せられるのもこうした文脈による。それ以来、前頭前野が「精神の座」とも呼ばれ、統合失調症やADHAなどと関係してくるのも、こうした“記憶喪失”の結果、状況に合った正しい運動・行動の遂行が困難になるからである。すなわち、運動・行動の正しい評価ができなくなったということなのだが、前頭前野は評価機能をもつということでもある。帯状回も“運動の評価”や“注意”の機能をもつとされるが、前頭前野が帯状回から系統進化的に発生し、神経結合も双方向性投射が強いことを考えれば、両者の相似性は当然である。ここで運動・行動の遂行を俯瞰的に示す Paillard のスキーマに触れておくと、枚数制限から図を出せない。原典等を参照されたい<sup>14)15)</sup>。辺縁系が取り込まれているのは重要である。

#### 4-11-5 小脳と運動

大脳運動野が壊されると運動そのものができなくなる。一方、小脳障害では、運動はともかくできる。ただ大変拙劣になる。いわゆる、小脳症状と呼ばれているもので、意図振戦、眼振、測定異常 (過大測定・過小測定)、運動障害・運動分解、構音障害、筋緊張低下などである。Kornhuber (1971) は臨床所見から、「小脳は時計だ」と唱えた。つまり、小脳障害はこの時計が狂うために、巨視的には筋活動の順序のタイミングが狂い、微視的には筋の運動単位の発射順序の乱れを生じる。小脳が時計として働くのは、プルキンエ細胞が極めて一定の間隔で配列されているため、信号が一つのプルキンエ細胞から次のプルキンエ細胞へと平行線維を伝わって行くとき、これが「遅延回路」を形成し“時間”を作っているからである。

小脳にも大脳で見られたような体部位再現が見られる。小脳核にも見られるし、その投射先の視床でも見られる。

小脳との兼ね合いで忘れてならないのは、ヒトで重要になる、小細胞性赤核との関係であろう。小脳 (歯状核) - 小細胞性赤核 - 下オリーブ核 - 小脳と閉じる、いわゆる、小川 (小川鼎三) の三角形である<sup>16)</sup>。小細胞性赤核の生理学的研究は皆無であり、その機能は解剖学的データによって推測するしかない。ただ、最近の fMRI のデータを参考にすると、高次機能と係わっている可能性も大きい。これら下オリーブ核のモデルや adaptive model 等については次節のモデルの所などで述べられよう<sup>17)</sup>。

小脳の構造を“微視的”にみると、結晶構造といわれるほど整然とした構造をなす。入力系は下オリーブ核からの登上線維と橋核・被蓋網様体からの苔状線維である (図 4・13 右)。後者からの軸索は表層で T 字状分岐し平衡線維となり、プルキンエ細胞と連絡し単純スパイクを発生する。その際“遅延回路”を形成することは述べた。単純スパイクの数は運動と相関する。登上線維はプルキンエ細胞の樹状突起に蔦のように絡みつき無数のシナプスを形成

する。これらは同時に活動し大きな EPSP とこれに重畳する複雑スパイクを発生する。複雑スパイクは短期的には増幅説や同期パルス説で説明されるように速度増加に関係しよう。長期的には学習と関連づけて考えられているが<sup>18,19)</sup>、「4-12 運動系システムのモデル」を参照されたい。

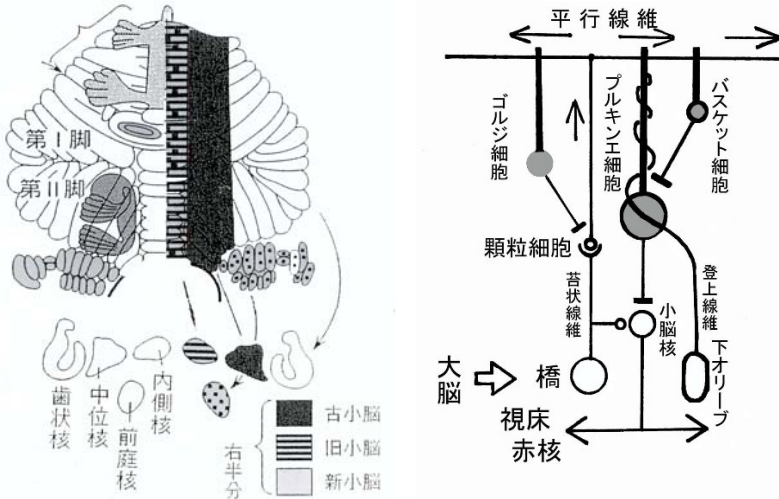
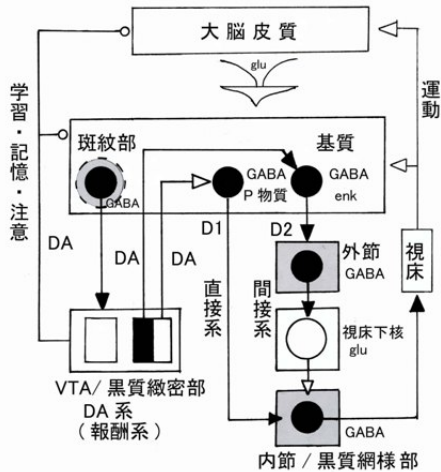


図 4-13 小脳の図。左図：左側に小脳の体部位再現及び小脳核。右側は系統発生図と古・旧・新各小脳皮質から小脳核への投射様式を示す (Snider & Eldred (1952) 改変)。右図：小脳の解剖学的回路。灰色の細胞は抑制性 (GABA)。

#### 4-11-6 大脳基底核と運動

大脳基底核は霊長類では新線状体 (被核と尾状核) と旧線状 (淡蒼球) に別れる。古くから錐体外路症候群と関連付けられてきた。ハンチントン舞踏病 (被核), アテトーゼ, パリスムス (視床下核), パーキンソン病 (黒質緻密部) である。新線状体はオタマジャクシ様であり、頭・体部からは前頭前野に投射が強く、ワーキングメモリに深くかかわる。したがって、広義の運動・行動の計画・評価、その結果としての整合性ある運動・行動の発現・遂行にもかかわる。

基底核のニューロンの基本構成を見ると、glutamate 作動性で興奮性の視床下核と、抑制性 GABA 作動性ニューロンが主管する被核・尾状核系に大別できる (図 4-14)。そして大脳から、被殻→外節→視床下核→淡蒼球内節/黒質網様部という間接路 (被殻と外節で二重の抑制系になっている。機能的には脱抑制による興奮を起こす) と、被殻から淡蒼球内節/黒質網様部に直行する直接回路がある。この両回路の機能の強弱により、パーキンソン病の寡動、ヘミパリスムスの運動過多が説明できる。ドーパミン欠乏症状と、D1 アゴニスト・D2 アゴニスト投薬による症状改善も説明できる<sup>20)</sup>。また、新線状体は黒質網様部に投射する基質 (matrix) と、黒質緻密部に投射する斑紋部 (Patch) で構成されるが<sup>19)</sup>、Barto (1995) は強化学習と関連づけ、Actor-critic モデルを提出している<sup>17),22)</sup>。



※色は抑制性. glu ; glutamate. DA ; ドーパミン.

図 4・14 大脳基底核の回路図

直接路と間接路がある。黒質緻密部の DA 欠損はパーキンソン病を惹起する。治療薬は前駆物質 L-Dopa である。DA 系には D1 受容体と D2 受容体がある。DA 系は報酬系でもあり学習・記憶に重要である。統合失調症にも関係する。

## ■参考文献

- 1) A. L. Hodgkin and A. F. Huxley, "A quantitative description of membrane currents and its application to conduction and excitation in nerve," J. Physiol., 117: 500-544, 1952.
- 2) W. Ral, "Core conductor theory and cable properties of neurons," In: "Handbook of Physiology. The Nervous System, Cellular Biology of Neurons," Bethesda, MD ; Am. Physiol. Soc., sect.1, vol.1, p.39-98, 1977.
- 3) 渡辺 昭, "膜の生物物理," 吉岡書店, pp.77-129, 1966.
- 4) 野間昭典, 松岡 達, 皿井信明, "包括的心筋細胞モデル (KYOTO model)," pp.64-87 ; 岡本良夫 (編著), "心臓のフィジオーム: 電気生理現象のシミュレーション, 分子から臓器まで," 森北出版, p.314, 2003.
- 5) 北里 宏, "興奮性膜の一般生理," 南江堂, 1972.
- 6) 岡本良夫 (編著), "心臓のフィジオーム: 電気生理現象のシミュレーション ; 分子から臓器まで," p.314, 森北出版, 2003.
- 7) W. Rall, "Distinguishing theoretical synaptic potentials computed for different soma-dendritic distributions of synaptic inputs," J. Neurophysiol., 30: 1138-1168, 1967.
- 8) 内菌耕二, 松波謙一, "神経細胞・神経回路網," 大島正光他 (編集), "ME 入門講座 第5巻 生体とシミュレーション," 金原出版, pp.10-66, 1970.
- 9) C. Koch and I. Segev (eds), "Methods in Neuronal Modeling," MIT Press, p.671, Cambridge, Massachusetts, USA, 1998.
- 10) I. Segev and W. Rall, "Computational study of an excitable dendritic spine," J. Neurophysiol, 60: 499-523, 1988.
- 11) J. P. Miller, W. Rall and J. Rinzel, "Synaptic amplification by active membrane in dendrite spine," Brain Res.,

- 325: 325-330, 1985.
- 12) 松波謙一, 内藤栄一, “最新運動と脳 (2 版),” サイエンス社, 2009.
  - 13) 久保田競 (編), 松波謙一, 船橋新太郎, 桜井芳雄 (著), “記憶と脳,” サイエンス社, 2002.
  - 14) J. Paillard, “Apraxia and the neurophysiology of motor control.” *Philos. Trans. R. Soc. London B Biol. Sci.*, 25: 111-134, 1982.
  - 15) 松波謙一, “運動と脳,” p.62, 紀伊国屋書店, 1986.
  - 16) 小川鼎三, “脳の解剖学,” p.247, 南山堂, 1953.
  - 17) 銅谷賢治, “計算神経科学への招待,” 数理科学 (臨時別冊) SGC ライブラリ 60, 115 頁, 2007.
  - 18) G. E. Alexander and M. D. Crutcher, “Functional architecture of basal ganglia circuit: neural substrates of parallel processing,” *TiNS* 13(7): 266-271, 1990.
  - 19) M. Ito, “The Cerebellum and Neuronal Control,” Raven Press, New York, 1984.
  - 20) 川人光男, “脳の計算理論,” p.457, 産業図書, 1996.
  - 21) C. R. Gerfan, “The neostriatal mosaic: multiple level of compartmental organization,” *TiNS* 15:133-139, 1992.
  - 22) A. G. Barto, R. S. Sutton and C. W. Anderson, “Neuronlike adaptive elements that can solve difficult learning control problems,” *IEEE Transactin on Systems, Man and Cybernetics*, 13: 834-846, 1983.

## ■脳と神経モデル—4章

### 4-12 運動系のモデル

(執筆者：藤田昌彦) [2008年5月 受領]

運動は運動神経が指令信号を筋骨格系や外眼筋運動系に送り込んで実現される。反射運動や随意運動は目的と制御対象の特性に応じて、適切な運動指令を構成している。運動系のモデルはそのような運動指令の形成と実行を含むべきであるが、ここでは筋骨格系の具体的な制御は含まず、運動指令の形成に焦点をあてる。また、運動には物理的な正確さと、それが運動主体にもたらす価値・報酬という、密接に結びついた二つの側面があるが、ここでは前者の物理的正確さ、目標軌道・目標点を基準にして適切な速さ、精度で運動を実現する機構を扱う。経年的な身体変化や、傷害などで運動が正確さを失うとそれを補償する機構が働くし、新規な運動課題を練習すると徐々にうまくなる。これらは適応、運動学習と呼ばれる。この運動学習を調べることで、運動を正確に実現する機構を明らかにすることが期待できる。関連する成書や解説論文も多い<sup>1)~5)</sup>。

#### 4-12-1 運動学習の計算論

システムの入力と出力を長く観察すると、その特性、振る舞いが見える。制御理論はこれをシステム同定といい、数理モデルを統計的手法で推定する。これは学習によっても達成できる(図4-15 A)。共通の入力を受けるシステムとモデルの出力を比較して、その情報を基にして、モデルを修正して対象システムの振る舞いを模倣するように学習を進める。

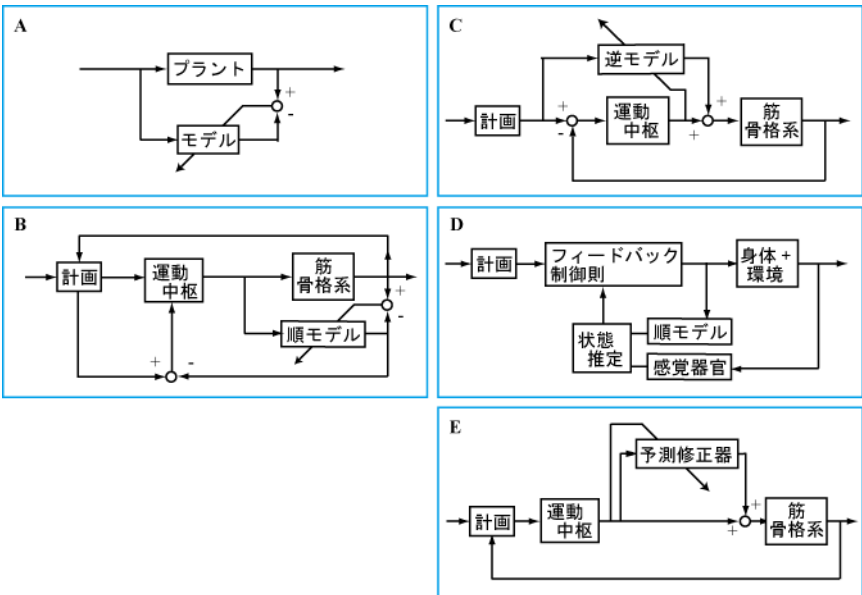


図4-15 運動学習理論のブロック図

## (1) 順モデル

一般に、運動によって周りの世界や身体末梢からもたらされる感覚情報は変化する。そこで、運動中枢は動きの指令を出すとき、そのコピー (Efference Copy) を脳の他の領域に送って、運動主体自身の動きによる変化成分を差し引いて、外界の動きや身体への作用を検出させているという説がある。動かない外界という知覚を形成するはずだから、空間の恒常性知覚の源泉と考えられる。運動制御を考えてみる。感覚フィードバックは時間遅れが大きいので、可能な運動制御はゆっくりしたものに限られる。そこで Efference Copy の考えを発展させて、筋骨格系を対象システムとしてそのモデルを中枢内部に獲得することができれば、運動指令や関節・筋からの固有受容感覚を基にして、感覚フィードバックに先行して運動状態を予測することができる。目標と照らし合わせて、より早く正確な制御が行える (図 4・15B)。これを順モデル (forward model) という<sup>6)~8)</sup>。順モデルの出力と運動状態が一致しないとき、その誤差を基準として順モデルの修正を行う。

## (2) 逆モデル

逆モデル (Inverse Model) はシステムの入出力特性を逆転する。ある動作をシステムに実行させたいとき、その希望する振る舞いをまず逆モデルに通すと、対応する運動指令に変換される。これを対象システムである筋骨格系に与えると、元の希望する振る舞いが実現できる。単純明解な原理である。ただし安定な逆モデルが存在するとは限らないし、一般にフィードバックのない前向き制御系は外乱の影響を補正できない。そこで逆モデルの学習機構として、これにフィードバック系を随伴させて、運動を確実に実行させる (図 4・15C)。かつ、そのフィードバック制御器の出力信号は、その時点における逆モデルの出力を補正していると考えて、それを参照して学習を進めれば対象システム (筋骨格系) の逆モデル形成が期待できる。Kawato ら<sup>7),9)</sup> は順モデルと逆モデルを含む随意運動制御の理論を提案するとともに、一般ニュートン法を用いて逆モデル学習の収束性を証明した。フィードバック誤差学習理論である。フィードバック系は目標入力のサーボ系として動作するので、運動実行前に目標軌道を計算する階層的な構造になっている。これもこの理論の特徴であり、軌道計画について多くの展開がなされてきた<sup>2),4),10)</sup>。

一般にヒトが道具を使って運動すると、対象システムである筋骨格系の特性はその負荷によって道具ごとに変化する。様々な道具使用の習熟過程は道具ごとの逆モデルの学習とみなせる。場面に応じて逆モデルの切り替えを行う必要が生じるが、これは順モデルと逆モデルをペアにして学習しておいて、順モデルのうち最良の予測を示すものとペアの逆モデルを選択すればよい<sup>11)</sup>。理論は順逆モデルをペアにしたモジュール群が責任信号を介して一斉に学習を進めてそれぞれ専門化するという筋道を示している。

更に、制御対象の順逆モデルの形成という能力が我々人間にあるのなら、そのような能力を基にしてより一般的な対象のモデル形成という学習も考えられる。他者の内部モデルの形成と考えれば、社会的なコミュニケーションの場面に拡張して適用できるだろうというアイデアもあり、このような自己や他者の内部モデル形成の能力の欠損で精神疾患を説明する試みもなされている<sup>12),13)</sup>。順逆モデルは内部モデル理論として展開されている<sup>6),14)</sup>。



#### 4-12-2 最適フィードバック制御と順モデル学習

筋をドライブする運動神経の信号にはその大きさに比例した標準偏差をもつ雑音加わると推定されている。そのために、目標に向かう運動を速くすると正確さが落ちる。これに着目して、Harris & Wolpert<sup>15)</sup>は終点誤差の分散最小モデルを提出した。眼球の運動は位置の時間に関する3次までの微分方程式で近似できる。制御入力の大さきの2乗に比例した分散をもつ白色雑音が運動指令に加わるとすると、一定角度の回転をするには終端点の分散と回転速度の間にトレードオフが生じる。両者の和をコストとして最適問題を解いて得られたSaccade (サッカド、サックード、衝動性眼球運動、跳躍性眼球運動)の速度特性は、振幅ごとの実験データとよく一致するものであった。同様な解析は単関節、2関節の腕の平面運動でも示されており、距離と精度から2点間の到達運動の所要時間を記述するFitt's lawも最適化問題として説明できることが示された。到達運動の軌道を予測する躍度最小<sup>16)</sup>やトルク変化最小<sup>10)</sup>などの最適基準と比較すると、運動主体にとって基準の意味が理解しやすい。

運動にはエネルギー消費、疲労などの負の側面と、目標達成という正の側面がある。これを運動制御のコストと運動の正確さがもたらす報酬の和を基準として、感覚フィードバックによる制御系の最適化問題という枠組みに乗せた理論<sup>17),18)</sup>がHarris & Wolpert<sup>15)</sup>と同じ頃に提案されている。「運動協調の理論」というTodorov & Jordanの論文タイトル<sup>17)</sup>が示すように、この理論は十分な熟練後の運動制御を表現しており学習機構を示した理論ではない。この理論の特徴は、フィードバック誤差学習理論が前提とする運動実行前の軌道計画を必要とせず、力学条件を満たす最適解を実行することである。ただしオンラインで最適解を計算する負荷が大きいこと、個々の計算結果を他の運動制御に転用できないという不効率がある。

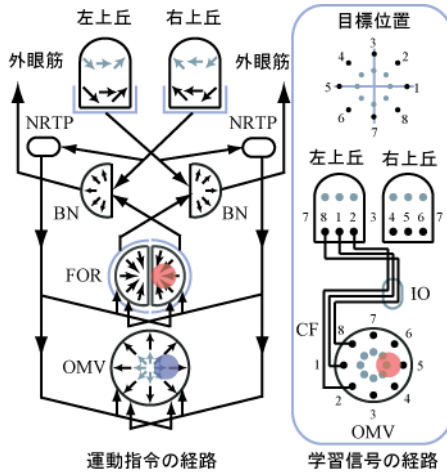
この枠組みで運動学習を実現するために、運動指令と固有受容器の信号を入力として、状態予測を行う順モデルを導入する試みがある。順モデルと視覚などの感覚器官からの情報によって状態推定を行って最適制御を進める。運動誤差は順モデルの推定と感覚フィードバックの不一致として現れると考えて、これを基に順モデルの修正として運動学習を行う<sup>19),20)</sup>。

#### 4-12-3 場所符号連合学習

Kawatoら<sup>7),9)</sup>が提案したフィードバック誤差学習理論—逆モデルの優れてオリジナルな点は、学習信号が運動性起源であることを示唆したことである。単純な反射運動の適応を検討している間は、感覚器からの誤差信号が学習を導いていのようにみえる。しかし一般の随意的な運動の学習を検討してみよう。運動の実行には手、肘、肩、胴、脚など多数の系が関与している。網膜像として得た運動誤差を個別の部分系へどのように割り当てることができるだろうか？ 感覚情報を基に課題遂行のために関連する筋骨格系全体に運動指令を送るのは運動中枢である。この中枢は初めの運動(主運動と略す)の企画・実行だけでなく、運動が誤差を残したときにも修正すべく働く。この修正運動が試行の繰り返し過程で少しずつ主運動に取り込まれるなら、うまく学習が進むに違いない。このような考えから場所符号化連合学習理論(FFAL)が提案されている。図4・15Eに示すように主運動指令から直ちに修正運動指令を連想する仕組みがあり、その出力と主運動指令が下位運動中枢や筋骨格系に合流して主運動が実行される。この連想は小脳で形成される瞬き応答の古典的条件付け<sup>22)</sup>と同じである(連合理論)。修正運動指令と同時に、学習信号が連合学習装置(小脳)の修正運動に対応するユニットに到達して、約200ms前に送り込まれていた主運動の信号がそのユニットの

出力をドライブするように変化を起こす。ここでの学習信号の役割は、ユニットの選択に主要な機能があり、詳細な時間的情報は不要である。瞬き応答の古典的条件づけにおいて無条件刺激を遅らせて与えると条件応答も遅れて出現するという遅延条件づけと同じ仕組みが予測修正器による運動のタイミング制御に働くだろう。この予測修正器は複数の運動中枢が直列、並列に回路を構成していても、それぞれに随伴させて働くことができる。高次、低次中枢における大域的、局所的運動評価・修正によって学習が進むと期待できる。

この理論が具体的に展開されている運動系は今のところ、眼球運動 saccade 系であり、**図 4・16**に示す。



IO : 下オリーブ核, NRTP : 橋蓋網様核

図 4・16 脳幹 saccade 系の場所符号連合学習モデル

上丘は脳幹における saccade 中枢である。ここを微小電極で刺激すると場所に応じた振幅・方向の Saccade が生じる（以下、Saccade を図の矢印で表現）。この指令は直接経路と小脳（虫部皮質 Vlc-VII<OMV と略>と眼球運動関連室頂核<FOR と略>）を介する間接経路とで、脳幹の Saccade-related Burst Neuron (BN; 高頻度な発火で Saccade の水平、垂直成分を形成する) に到達する。そこで、例えば右向き Saccade が目標を越えた結果、左向きの修正 Saccade が生じたとしよう。図 4・16 左の紫色の領域の Purkinje 細胞の活動が次回の右向きの Saccade 中に長期減弱すれば、間接経路からの右向き成分が減るので、その分、Saccade の振幅は減少する。それには学習信号である登上線維 (CF) 信号がこの領域に到達すればいい。左向きの修正 Saccade は右上丘の活動で生じる。図右が示すように、登上線維信号は同側の OMV に投射する。長期減弱の時間条件にしたがって、100~200ms 先行する平行線維活動は第 1 Saccade によるもの以外にない。すなわち、右向きの saccade 指令が前回より弱い右向き要素を小脳に惹起するので、振幅減少が生じると推定できる。

## ■参考文献

- 1) 川人光男, “脳の計算理論,” 産業図書, 1996.
- 2) 銅谷賢治, 五味裕章, 阪口 豊, 川人光男編, “脳の計算機構,” 朝倉書店, 2005.
- 3) 伊藤宏司, “身体知システム論-ヒューマンロボティクスによる運動の学習と制御-,” 共立出版, 2005.
- 4) 阪口 豊, “上肢到達運動制御の計算モデルとその課題,” VISION, 16, 2, pp.93-110, 2004.
- 5) 藤田昌彦, “眼球運動のモデル,” 岩波認知科学 3 運動, 第4章, 岩波書店, 1994.
- 6) M. Ito, “Neurophysiological aspects of the cerebellar motor control system,” *Int. J. Neurol.*, 7, 2, pp.162-176, 1970.
- 7) M. Kawato, K. Furukawa, and R. Suzuki, “A hierarchical neural-network model for control and learning of voluntary movement,” *Biol. Cybern.*, 57, 3, pp.169-185, 1987.
- 8) M. Desmurget, and S. Grafton, “Forward modeling allows feedback control for fast reaching movements,” *Trends. Cogn. Sci.*, 4, 11, pp.423-431, 2000.
- 9) M. Kawato et al., “Coordinates transformation and learning control for visually-guided voluntary movement with iteration: a Newton-like method in a function space,” *Biol. Cybern.*, 59, 3, pp.161-177, 1988.
- 10) Y. Uno, M. Kawato and R. Suzuki, “Formation and control of optimal trajectory in human multijoint arm movement. Minimum torque-change model,” *Biol. Cybern.*, 61, 2, pp.89-101, 1989.
- 11) D. M. Wolpert and M. Kawato, “Multiple paired forward and inverse models for motor control,” *Neural Netw.*, 11, 7-8, pp.1317-1329, Oct. 1998.
- 12) S. J. Blakemore and C. Frith, “Self-awareness and action,” *Curr. Opin. Neurobiol.*, 13, 2, pp.219-224, April 2003.
- 13) S. J. Blakemore et al., “The detection of contingency and animacy from simple animations in the human brain,” *Cereb. Cortex*, 13, 8, pp.837-844, 2003.
- 14) D. M. Wolpert, K. Doya and M. Kawato, “A unifying computational framework for motor control and social interaction,” *Philos. Trans. R. Soc. Lond. B Biol. Sci.*, 358, 1431, pp.593-602, Mar. 2003.
- 15) C. M. Harris and D. M. Wolpert, “Signal-dependent noise determines motor planning,” *Nature*, 394, 6695, pp.780-784, Aug. 1998.
- 16) T. Flash and N. Hogan, “The coordination of arm movements: an experimentally confirmed mathematical model,” *J. Neurosci.*, 5, 7, pp.1688-1703, July 1985.
- 17) E. Todorov and M. I. Jordan, “Optimal feedback control as a theory of motor coordination,” *Nat. Neurosci.*, 5, 11, pp.1226-1235, Nov. 2002.
- 18) E. Todorov, “Stochastic optimal control and estimation methods adapted to the noise characteristics of the sensorimotor system,” *Neural Comput.*, 17, 5, pp.1084-1108, May 2005.
- 19) R. Shadmehr, and J. W. Krakauer, “A computational neuroanatomy for motor control,” *Exp. Brain Res.*, 185, 3, pp.359-381, Mar. 2008.
- 20) H. Chen-Harris et al., “Adaptive control of saccades via internal feedback,” *J. Neurosci.*, 28, 11, pp.2804-2813, Mar. 2008.
- 21) M. Fujita, “Feed-forward associative learning for volitional movement control,” *Neurosci. Res.*, 52, 2, pp.153-165, June 2005.
- 22) R. F. Thompson, “In search of memory traces,” *Annu. Rev. Psychol.*, 56, pp.1-23, 2005.