

大脳基底核の計算モデル

中原 裕之

(BSIS Technical Report No.03-01,2003)

0.1 大脳基底核と強化学習

このレポートでは、大脳基底核の機能に関連して、我々が行ってきた二つの研究を紹介する。一つは、報酬信号に基づく自己組織化という観点から基底核の細胞の反応をモデル化した研究である。もう一つは、Actor-Critic モデルの立場から大脳皮質-基底核回路がどのように逐次運動制御・学習を実現しているかを考えた研究である。

これらの研究の紹介に入る前に、強化学習と大脳基底核との関係についてモデルの立場から手短かにまとめておこう。現在、大脳基底核が、機械学習の分野で言われる強化学習を何らかの形で行っていると考えられている。その発端となったのは、大脳基底核にあるドーパミン細胞の反応が、強化学習の一つの手法である Temporal Difference (TD) 学習で用いられる TD 誤差に似ているという指摘である。この指摘の基礎となる Schultz らのドーパミン細胞の実験結果をごく簡単に説明する [18]。彼らの実験では、報酬 (reward) の前に報酬が来ることを知らせる刺激 (条件刺激;CS) を与える。条件付けが不充分だと、報酬が与えられた直後に中脳ドーパミン細胞の一過性 (phasic) の反応がある。その一方、十分に条件付けが行われると、ドーパミン細胞は CS に一過性に反応し、報酬への反応は消える。さらに、CS の前にもう一つ別の CS を与えると、その刺激にドーパミン細胞は反応し、元々の CS への反応は消える。これらのドーパミン細胞の反応が TD 誤差によく似ているというのである。これは、TD 学習が、この実験でどのように行われるか考えるとわかりやすい。学習前には各状態の評価関数は獲得されていないので、TD 誤差は報酬のところに現れる。ところが、評価関数の学習後は、TD 誤差は刺激のところに現れる。つまりドーパミン細胞の反応と同様なことが起こる。この関係が大脳基底核と強化学習、とくに TD 学習と対応付ける議論のきっかけとなった。もちろん以前から、パーキンソン病、ハンチントン病など、大脳基底核に関わる逐次系列の行動障害の事例や、ドーパミン細胞は元々報酬刺激に対する反応が強いなどの状況証拠があったことも見逃せない。

さて、TD 学習と、それと共に用いられる Actor-Critic モデルについて見てみよう。Actor-Critic モデルでは、TD 誤差を使い、Critic モジュールと、Actor モジュールがそれぞれ学習を行う。Critic は各状態の評価関数の学習を目的とし、TD 誤差を 0 にするよう学習を行う。一方、Actor は TD 誤差をプラスの方向に最大化するよう各状態でのアクションを選ぶように学習を進める。この二つの学習で結果的には、一番よい系列を選ぶように学習が進む。Actor-Critic モデルと大脳基底核との対応 [12] は、Critic にあたる線条体 (striatum) 内の細胞は状態の評価を計算し、ドーパミン細胞へ投射する。Actor にあたる線条体内の細胞は現在の状態に応じた運動指令の選択を行い、淡蒼球と視床を経て大脳の運動関連領野へ出力を送る。ドーパミン細胞は線条体からの評価信号と実際の報酬信号との差をとって、TD 誤差を作り出す。ドーパミン細胞群は線条体に拡散的に投射しているので、Actor と Critic の両方の細胞群に影響を及ぼし、特にそれらのシナプスの可塑性に影響を与える。これを利用して Actor と Critic は TD 学習を行うという仮説がたてられた。この仮説は大変有望であるが、一方で、例えば、Actor、Critic と線条体の対応などはまだ仮説の粋を出ず、その意味で、まだ不完全・未確認の点も多い。言い換えれば、研究対象として現在非常に面白い分野である。それでは、まず一つめの研究の紹介を始めよう。

0.2 報酬信号の調節下での尾状核ニューロンの自己組織化

0.2.1 1DR 課題

彦坂研究室では、非対称に報酬を与える記憶誘導型眼球運動課題 (memory-guided saccade task) を用いて、報酬の変化による尾状核 (caudate; caudate は線条体の一部である。) の神経細胞の反応の変化を調べた [13] (図 1)。この課題では、各試行は、注視点が中心に提示されることで始まる。次にキュー刺激が短時間、4 方向のうち 1 方向に与えられる。そして、注視点が消えた後、サルはキュー方向に saccade を行う。これが一試行である。各試行でキューの与えられる方向はランダムになっている。60 試行を 1 ブロックとし、1 ブロック中、全方向に報酬を与える全方向報酬条件 (all-directions-rewarded condition, ADR) と、4 方向のうち 1 方向にだけ報酬を与える一方向報酬条件 (one-direction-rewarded condition, 1DR) の 2 つ異なる報酬条件で、サルにタスクを行わせた。1DR は 4 ブロック行い、各ブロックでは報酬方向が異なっていた。なお、報酬を与えない方向でも、サルは正しく saccade を行わなければならない。行わないと、同じ方向の試行がくり返される。

上記の論文 [13] で報告された結果を簡単にまとめよう。この実験で、各々の 1DR ブロックでの、尾状核神経細胞の各方向のキュー刺激への反応特性

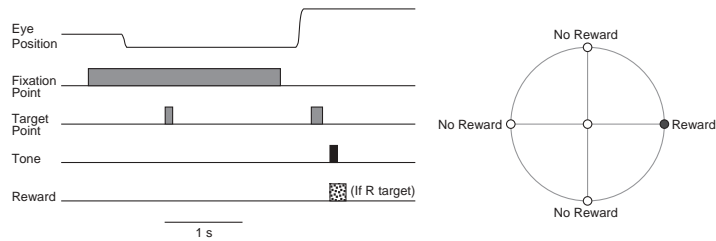


図 1: (左) 記憶誘導型眼球運動課題 (右) 右方向のキューに報酬を与える場合の 1DR

が、そのブロックの報酬方向の影響を受けることが分かった。そして、この報酬方向による反応の変化の仕方が、おおまかに 3 タイプに分かれることが分かった (図 2)。その 3 タイプとは、

a) Flexible type

最も多く見られるタイプで、神経細胞が最も良く反応する方向は、各 1DR ブロックの報酬方向に一致する。

b) Conservative type

各 1DR ブロックを通じて、その神経細胞が良く反応する方向が、一定程度ある方向に限られる。そして、その方向が報酬方向のとき一番強く反応する。

c) Reverse type

他の 2 タイプより少なく、各 1DR ブロックで、その報酬方向に対しての反応が一番弱い。つまり、flexible type と逆の反応を示す。

である。以下、これらのタイプのことを反応タイプと呼ぶ。

0.2.2 強化信号の修飾を受ける自己組織化モデル

さて、少し話を戻して、大脳基底核の次の 3 つの一般的特徴に注目しよう。

- 1). 大脳基底核は大脳皮質のほぼ全体から入力を受け、上丘、前頭葉にその出力を戻す (図 3,6) が、大脳皮質 大脳基底核経路のニューロン数を比較すると、だいたい、大脳皮質 10000 ニューロンに対し、大脳基底核は 100 ニューロン程度である。つまり神経細胞の投射に非常に強い収束がある。
- 2). 大脳基底核内の線条体には、抑制系のニューロン (側抑制や interneuron) が豊富である。

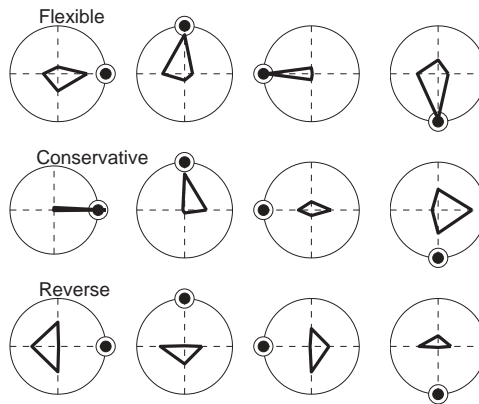


図 2: 実験で見られた 3 つの反応タイプ

3). 線条体は黒質緻密部のドーパミン細胞から豊富な強化信号 (reinforcement signal) を受けている (図 3)。

この 3 つの特徴をふまえた上で、我々が興味をもったのは、上記の 3 つの反応タイプがどのように生まれてくるかということだった。強化信号を利用した Actor-Critic のモデルを考えてみるのも面白いのだが、我々はむしろ自己組織化という観点から考えてみた¹。自己組織化、つまり、神経細胞群が入力情報の効率良い表現を自動的に獲得するという観点は、上記の特徴 (1) から考えても非常に有望である。特に、この自己組織化が、特徴 3) のド - パミンの信号を利用し、また特徴 2) の抑制性の細胞をうまく利用しているのが望ましい。以下に述べるように、我々のモデルはこれらの要件を満たしている。

実は、我々のモデルは、甘利らが 70 ~ 80 年代に提案した抑制系の可塑性を利用した視覚野の自己組織化モデル [2, 4, 3] を基礎としている。温故知新である。まずは、このモデルを少し数式を出しながら説明しよう。基本となるのはニューロンモデルで、それは以下ようになる。

$$y = f(u(t))$$

$$u(t) = w(t) \cdot x(t) - w_0(t)x_0(t)$$

y, u はそれぞれニューロンの出力、内部状態、 w, w_0 は興奮性、抑制性の荷重値、 x, x_0 はそれぞれ興奮性、抑制性の入力、 f は出力関数である。興奮性、

¹実は、このニューロンの各タイプを Actor-Critic と対応付けようと試みたこともある。例えば、Critic が各条件で報酬に対する反応、予測をだしているならば、Critic は flexible type のニューロンであろう。各方向に対応する Actor が競合して役割をはたすなら、Actor は空間選択性が高く、一方で報酬の修飾も受けよう基底核から視床へは、直接経路と間接経路の 2 つがあることも考えると、conservative type が直接経路の、reverse type が間接経路のそれぞれの Actor であろうと考えた [19]。これらは一定程度の成功を収めているが、眼球運動との関係など、まだ私自身納得していない。

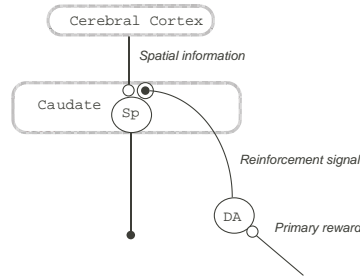


図 3: 尾状核 (caudate) の神経細胞 (Sp) とドーパミン (DA) 細胞のダイアグラム

抑制性の荷重値の学習則は、各々減衰つきのヘップ学習となっている。

$$\begin{aligned}\tau \dot{w}(t) &= -w(t) + cf(u(t))x(t) \\ \tau \dot{w}_0(t) &= -w_0(t) + c'f(u(t))x_0(t)\end{aligned}$$

ここで、 c, c' はそれぞれ興奮性、抑制性の荷重値の学習に対する学習係数である。第一項は、発散を抑える項で、第二項はヘップ学習の項である。ここで大切なのは、興奮性、抑制性の各々の学習則は比較的簡単なヘップ学習則であるが、いわゆる通常のヘップ学習と決定的に違うのは、まさしく興奮性、抑制性の両方の学習則を同時に考えることにある。このモデルでは、例えば方位選択性を示すなど、入力情報の特徴に応じて効率のよい表現が得られることなどが明らかにされていた。

大脳基底核のモデルに戻ろう。線条体 (神経細胞) の出力を y 、大脳皮質からの入力を x 、その入力に対する荷重値を w と書くことにする。ドーパミン細胞の出力、強化信号を α とする。この α が興奮性入力 x を修飾するニューロンモデルを考える (図 3、4)。

$$y(t) = f(u(t))$$

$$u(t) = w(t) \cdot x(t)(1 + \alpha(t)) - w_0(t)x_0(t)$$

で、 u はニューロンの内部状態、 f は出力関数、 w_0, x_0 はそれぞれ抑制性の荷重値、入力を表す。簡単のため、ここでは抑制性入力 x_0 を定数とする。先ほどのモデルとほんの少し、ド - パミン信号が入っているところ、だけ違う。このほんの少しの違いが様々な現象を生み出すところが面白いのである。学習則は先ほどと同様に

$$\begin{aligned}\tau \dot{w}(t) &= -w(t) + cy(t)x(t) \\ \tau \dot{w}_0(t) &= -w_0(t) + c_0y(t)x_0\end{aligned}$$

で与える。ここで、 c, c_0 は興奮性、抑制性の荷重値更新に対する学習係数である。

このニューロンモデルと学習則は、次の2つの実験事実を基にしている：
 (1) ドーパミンの入力により線条体ニューロンの内部電位が増加あるいは減少する [6]，一方で、ドーパミン入力により線条体ニューロンのシナプス荷重が変化する [20, 5, 17]、ことも知られている。上記のニューロンモデルは、(1)を反映している。さらに、(1)つまりニューロンモデルの出力に及ぼすドーパミンの影響は、間接的に学習則に影響を及ぼす。つまり、上記(2)も満たしている。

以下に、尾状核（線条体）神経細胞の3つの反応タイプがこのニューロンモデルの自己組織化メカニズムから生まれてくることを見ていく。つまり、ドーパミンの強化信号による修飾のもとで可塑的に変化する興奮性と抑制性のシナプス荷重の効果のバランスの違いにから、異なる反応タイプが生まれてくるのである。

解析としては、4方向の報酬条件の各々の1DRで、ニューロンが学習を終えたとき、つまり平衡状態（最終的な安定状態）でそれぞれの response type を示すことを、そして、またブロックが変わった時に、一つ前のブロックの平衡状態を新しいブロックの初期状態とした時に、その初期状態からの学習後に対応する反応タイプの平衡状態になることを確かめる、という二つが求められる。ここでは、簡単にその道筋を追ってみよう。

まず解析を簡単にするために、出力関数 f として、ステップ関数

$$f(u) = \begin{cases} 1 & (u > 0) \\ 0 & (\text{otherwise}) \end{cases}$$

を用いることにする。各報酬方向（4方向）の入力を $x_i (i=1,2,3,4)$ と表し、各 x_i のくる確率を p_i とする。 p_i は1DRブロック内では1/4になる。また各入力 x_i に対する強化信号を $\alpha(x_i)$ で表す。

さて、各方向の入力で平均した学習則は、

$$\begin{aligned} \tau \dot{w} &= -w + c \sum_i p_i y_i x_i \\ \tau \dot{w}_0 &= -w_0 + c_0 \sum_i p_i y_i x_0 \end{aligned}$$

で表せる。この時、シナプス荷重の平衡状態 \bar{w}, \bar{w}_0 は、 $\dot{w} = \dot{w}_0 = 0$ を満足するので、

$$\begin{aligned} \bar{w} &= c \sum_i p_i y_i x_i \\ \bar{w}_0 &= c_0 \sum_i p_i y_i x_0 \end{aligned}$$

と表せることになる。ここで、荷重値が収束した後のニューロンの内部状態を \bar{u} とすると、

$$\bar{u} = \bar{w} \cdot \mathbf{x}(1 + \alpha) - \bar{w}_0 x_0$$

となる。入力 x に対し、 $\bar{u} > 0$ ならば、このニューロンは発火する。そこで、 $\bar{u}(x) > 0$ となる x の集合を受容野 R と定義する。この R を用いて、

$$p_R = \sum_{\mathbf{x}_i \in R} p_i$$

$$\mathbf{w}_R = \sum_{\mathbf{x}_i \in R} p_i \mathbf{x}_i / p_R$$

とおく。 p_R は全入力のうち R に含まれる入力がかかる確率、 \mathbf{w}_R は R の重心を表す。いま、ニューロンの出力関数としてステップ関数を用いているので、平衡状態でのニューロンの内部状態 \bar{u} は

$$\begin{aligned} \bar{u}_j &= \bar{u}(\mathbf{x}_j) \\ &= cp_R(1 + \alpha_j) \left(\mathbf{w}_R \cdot \mathbf{x}_j - \frac{c_0}{c(1 + \alpha_j)} x_0^2 \right) \\ &= cp_R(1 + \alpha_j) K(\mathbf{x}_j) \end{aligned}$$

と書ける。ここで、 $K(\mathbf{x}_j)$ は

$$K(\mathbf{x}_j) \equiv \mathbf{w}_R \cdot \mathbf{x}_j - \frac{\lambda}{1 + \alpha_j}$$

で定義され、 λ は、

$$\lambda = \frac{c_0}{c} x_0^2$$

である。この λ は、大雑把に言って、抑制ニューロンの与える影響の大きさを表す。 $K(\mathbf{x}_j) > 0$ でニューロンは反応し、 $K(\mathbf{x}_j) \leq 0$ で反応しない。ただし、 \mathbf{w}_R は各 1DR ブロックの条件に応じて出てくる値なので、この $K(x)$ に関する条件は、必要条件であり十分条件ではない。十分性の確認には、 R 内のすべての x に対し、 $K(x) > 0$ であり、また R 外のすべての x に対し、 $K(x) < 0$ であることをチェックする必要がある。

0.2.3 モデルと3つの反応タイプ

より具体的に解析を進めるために、入力 x を決めよう。大脳皮質からの入力は、各方向に選択的な部分と、いろいろな方向の間に共通の部分の入力があるだろう。そこで、大脳皮質からの入力 x_i を、各方向で共通する部分とキュー方向を特徴付ける部分に分けて、各方向でオーバーラップしている入力大きさを M 、各方向に選択的な入力大きさを N_i とする (図 4)。入力の方向優位性は M, N_i の大小で表現される。このことは、大脳皮質からの入力大きさと各方向の強さを示す実験データを反映している。各方向の入力

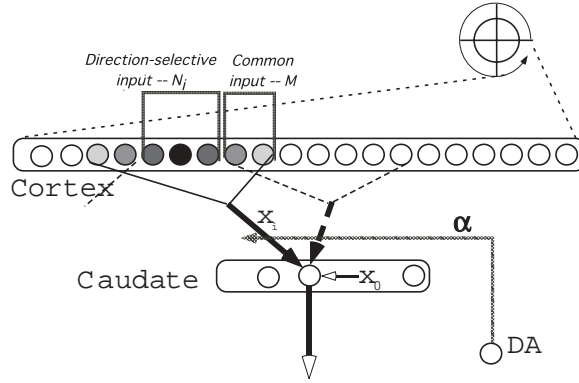


図 4: 自己組織化モデルのダイアグラム

を最も単純な形で書くと,

$$\begin{aligned}
 \mathbf{x}_1 &= (\overbrace{1, \dots, 1}^M, \overbrace{1, \dots, 1}^{N_1}, \overbrace{0, \dots, 0}^{N_2}, \overbrace{0, \dots, 0}^{N_3}, \overbrace{0, \dots, 0}^{N_4}) \\
 \mathbf{x}_2 &= (1, \dots, 1, 0, \dots, 0, 1, \dots, 1, 0, \dots, 0, 0, \dots, 0) \\
 \mathbf{x}_3 &= (1, \dots, 1, 0, \dots, 0, 0, \dots, 0, 1, \dots, 1, 0, \dots, 0) \\
 \mathbf{x}_4 &= (1, \dots, 1, 0, \dots, 0, 0, \dots, 0, 0, \dots, 0, 1, \dots, 1)
 \end{aligned}$$

とかけることになる。このときの入力間の内積は、

$$\mathbf{x}_i \cdot \mathbf{x}_j = \begin{cases} M + N_i & (i = j) \\ M & (i \neq j) \end{cases}$$

で表せる。また、1DR ブロック中で報酬を与える方向を \mathbf{x}_i とすると、この方向に対する強化信号 α_i は $\alpha_i = \alpha$ で、その他の3方向に対しては $\alpha_j = 0 (j \neq i)$ としよう。

以上の準備の下で解析を行うと、尾状核神経細胞の3つの反応タイプの条件式を求めることができる。詳しくは論文を見てもらうことにして答だけを書くと、各タイプは、

a) Flexible type

$$M + \frac{N}{2} \leq \lambda < M(1 + \alpha)$$

ここでは、 $N = N_i$ として簡単にして示した。

b) Conservative type

1の方向に空間選択性が高いとし、入力間の関係を

$$\mathbf{x}_i \cdot \mathbf{x}_j = \begin{cases} M + N_1 & (i = j = 1) \\ M + N & (i = j \neq 1) \\ M & (i \neq j) \end{cases}$$

$$N_1 > N$$

とすると、条件式は

$$M + \frac{N}{2} \leq \lambda < \min\{M(1 + \alpha), M + \frac{N_1}{2}\}$$

となる。

c) Reverse type

$$(M + \frac{N}{4})(1 + \alpha) \leq \lambda < M$$

但し、 $N = N_i$ とした。なお、このタイプでは $\alpha < 0$ である必要がある。 $\alpha < 0$ とは、ドーパミン細胞からの影響が荷重値の可塑性に負の効果があることになる。実験でもそのような現象が起こりうる事が確かめられている。

という条件にまとめられる。実際、この条件でシミュレーションをしてみると、それぞれのタイプが現れることが確かめられる (図 5A)。

さて、上記の条件は、ニューロンの出力関数 $f(\cdot)$ にステップ関数を使っていた。ステップ関数は、ニューロンが発火しているか、していないかの二値表現の関数だが、これの代わりに、シグモイド関数を用いると、(正規化した)平均発火頻度を表現できる。シグモイド関数ならば、ニューロンの内部状態 u の大小をニューロンの反応 y の大きさに反映することができる (図 5B)。これにより、生理実験データにより類似した結果をモデルを得ることができる。実際、図 2 と図 5B を比べると互いに似ていることが分かるだろう。なお、これらのニューロンの刺激方向に関する反応特性は、 M や N_i の各方向の入力の内積を表現している項の大小と対応づけることが可能である。また今は詳しくは論じないが、ADR の条件で flexible type の反応を見ると、1DR で報酬が与えられたときと同じ方向への反応を比べると弱くなっている場合があるが、このモデルでも同様な結果が得られる。

0.2.4 まとめ

大脳基底核の線条体神経細胞のモデルとして、ドーパミンの強化信号に修飾されながら自己組織化するニューロンモデルを考えた。このモデルは、線条体内の豊富な抑制性神経細胞、ドーパミン細胞による強化信号、大脳皮質から基底核へ強く収束する投射などの基底核の重要な特徴にうまく対応している。実験で得られる尾状核神経細胞の異なる反応タイプ [13] が同じモデルからパラメータの違いにより生じうることを明らかにした。このモデルは、線条体の神経細胞が、ド - パミンの強化信号を利用しながら、大脳皮質からの入力の効率良い表現を獲得しうることを示している。

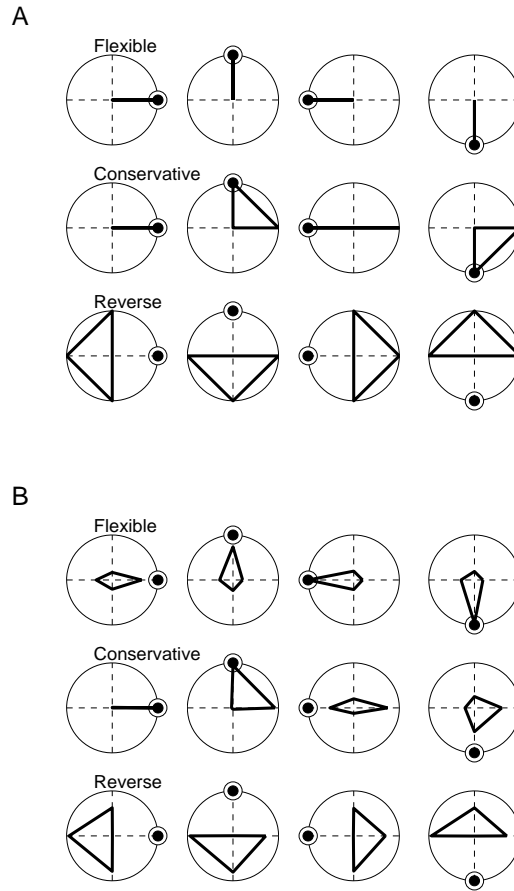


図 5: 自己組織化モデルのシミュレーション。A ステップ関数が出力関数。B シグモイド関数

0.2.5 質疑応答

Q: 大脳基底核は大脳皮質とループを形成しているので、このモデルによる大脳皮質への影響は?

中原: このモデルの動機のひとつは、大脳皮質へのループを選ぶとき、大脳皮質の入力の表現を効率よく作ることにあり、その理由のひとつが、大脳皮質と線条体のニューロン数が違うところにある。その点では、ループがどのような影響を及ぼすかを調べるのは面白い。時間遅れの報酬に対する強化信号の自己組織化を数理的にどうするかも面白い問題だと思う。また、この場合皮質との対応関係を考えないといけないと思う。

Q: モデルの抑制性を入力はどこからくるのか?

0.3. 複数の皮質－大脳基底核－視床回路の共同による逐次運動系列の学習, 制御¹¹

中原：側抑制か interneuron のどちらの可能性もある。最近の実験データから、個人的には interneuron の可能性の方が高いと思う。最近の実験結果は、大脳皮質から interneuron への直接の投射が重要であることを示唆している。

Q：自己組織化は通常、個々のユニットが同じ特性をもたないように、相互抑制なり、実効的に lateral inhibition を受けるものを入れるが、今のモデルでは？

中原：今のモデルではない。単純に言えるのは、尾状核ニューロンの出力を全部まとめステップ関数を通したものを inhibition にするものが考えられる。

Q：ここでの α は、TD 誤差か報酬そのものなのか？

中原：TD 誤差の形式にはなっていない。実験で見られるドーパミンの反応を α とおいている。単純には報酬そのものを反映していると思ってもらってもよい。学習の最初の数回でドーパミンはあるレベルまで達するので、それを表現している。このモデルでは TD 誤差は入っていない。

Q：このモデルには、学習の目的を表す objective function はあるのか？

中原：明示的な objective function はない。安易に objective function を導入するよりも、豊富な抑制ニューロンとその可塑性、Hebb 学習の自己組織化、更に強化信号の modulation という、シンプルなモデルからスタートして何が起きるかを見たかった。

0.3 複数の皮質－大脳基底核－視床回路の共同による逐次運動系列の学習, 制御

それでは、二つ目の研究の紹介に移ろう。この研究では、とりあえず Actor-critic の仮説を信用した上で、大脳皮質－大脳基底核回路が逐次系列運動の学習・制御にどのように貢献しているかということ調べた。特に、我々は多重表現仮説を提唱して、その仮説に基いてモデルを構築し、彦坂研究室の逐次系列運動実験（2x5 課題）での実験結果とシミュレーション結果の対応を調べた [14, 16]。この研究で着目したのは、一つには、ドーパミンの強化信号を通じた運動系列の学習で、もう一つは、大脳皮質－大脳基底核ループは幾つかの分かれたループになっているという解剖学的な特徴だった。

この解剖学特徴を押さえておこう (図 6)。解剖学的な研究から、大脳基底核は大脳皮質とループ構造を持つことが分かってきた。大脳皮質から大脳基底核には皮質全体から入力が入ってきて上丘など脳幹の運動中枢に加え、視床を介して主に frontal cortex に出力を戻している。特徴的なのは、実はその

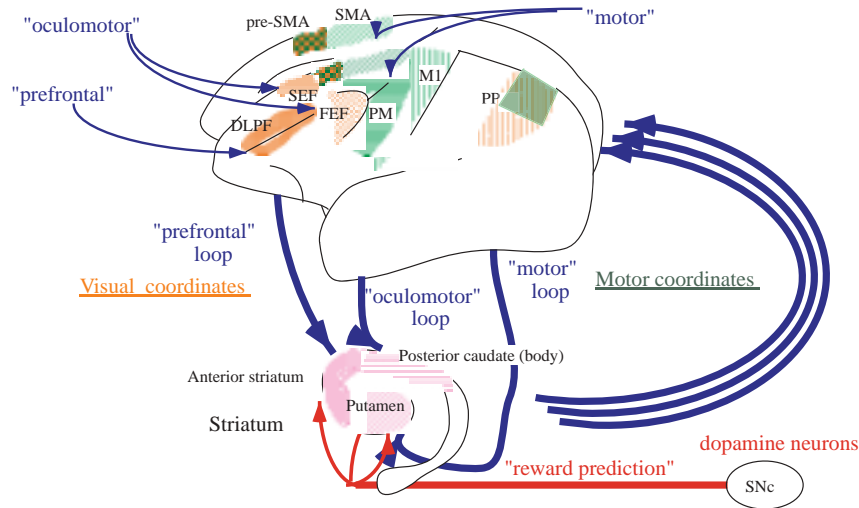


図 6: 大脳皮質-大脳基底核の並列ループ。略称は、dorsolateral prefrontal cortex (DLPF:前頭前野背側部), supplementary eye field (SEF:補足眼野), frontal eye field (FEF:前頭眼野), supplementary motor area (SMA:補足運動野), pre-supplementary motor area (pre-SMA:前補足運動野), premotor area (PM:運動前野), posterior parietal (PP:頭頂部後部)、primary motor area (M1:第一次運動野), striatum (線条体; anterior, 前部, posterior, 後部), caudate (尾状核), putamen (被殻), substantia nigra pars compacta (SNc:黒質緻密部)。

大脳皮質-大脳基底核ループは、いくつかの分かれたループを束ねたようになっていることである [1, 11] (図 6; 以下の略称は図のキャプションを参照のこと)。例えば、大脳皮質-大脳基底核ループのうちで、motor loop は、運動関連領域、SMA, M1, PM などが主に線条体後部に投射してそれがまたループを還している。また、oculomotor loop は、SEF, FEF などが尾状核の body のあたりに主として投射しループを還している。あるいは prefrontal loop は、DLPF を中心として PP も含め線条体前部に投射をして、ループを還している。これら複数のループが同時にドーパミンの強化信号の修飾を受けている。このような構造は、何のためにあるのだろうか。

0.3.1 2x5 課題

さて彦坂研で行われた逐次系列学習課題、2x5 課題、をまず説明しておこう (図 7A)[9]。サルがパネルの前に座って、4x4 のボタンのうちのランダムに 2 つのボタンが点灯する。これが第 1 セットである。サルは、この点灯した二つのボタンを、実験者が予め決めた順番で押さなければならない。ただし、サルはその順番を知らないの、試行錯誤しながら (trial and error)、

0.3. 複数の皮質－大脳基底核－視床回路の共同による逐次運動系列の学習, 制御13

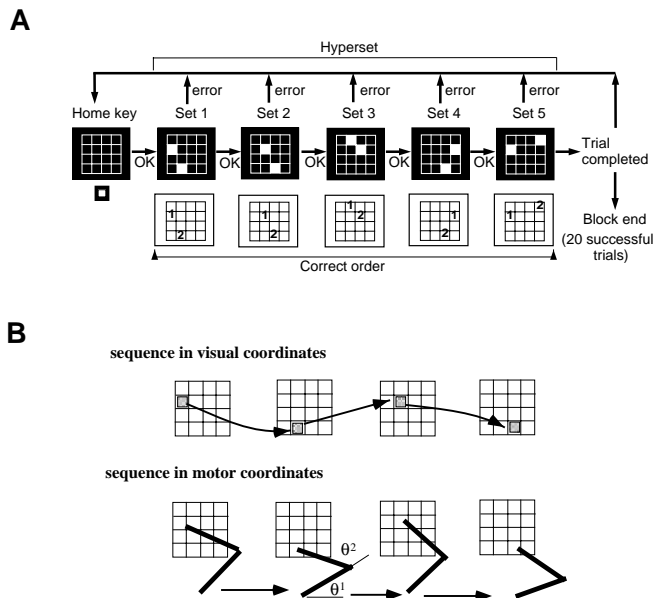


図 7: A 2x5 課題 B 視覚座標と運動（関節）座標

どちらのボタンを先に押し、次にもう一つのボタンを押すのかを学習しなければならない。失敗すると最初からやり直しとなる。もし、正しい順序で押すことができると、第2セットに入り、2つのボタンが点灯する。サルはまたここでも試行錯誤しながら正しい順序でボタンを押すことを学習しなければならない。もし、失敗するとまた最初に戻る。5セットが1単位のハイパーセットとなっている。従って、もし第1セットから第5セットまで続けて成功すると、そのハイパーセットを1試行成功したことになる。各々のハイパーセットでは、5セットは固定されている。

1ブロックの実験では、20回の成功試行をするまで、同じパイパーセットが続けて提示することになっている。サルは一日に例えば20個くらいのハイパーセット（つまり20ブロック）を経験している。このうちの半分はその場限りで経験するハイパーセットで、残りの半分は毎日同じハイパーセットを経験している。作りうるハイパーセットの組み合わせの数はものすごい数になる（約 7.6×10^{11} 個）ので、このように毎日新しいハイパーセットを経験させつつ、同時に同じハイパーセットを経験させることができるのである。これを半年とか一年とか、もっとやる。以下、サルが毎日経験して習熟したものを learned hyperset と呼び、その場限りのものを new hyperset と呼ぶ。

この課題は我々日常の系列学習を抽象的に巧みに表現している。我々の日常生活で、朝起きて、歯を磨く、服を着替えるといった行動は、大人の我々にとっては何でもない自動的な行動になっているが、実はこれらの行動は子どもの頃学習してできるようになったのである。一方、例えばある日誰かが

テニスを始める，スキーを始める，ワープロを始めるなど，そのときその場で我々大人も新しい系列課題を学習する。このように、我々の日常生活には、鍛えられて自動的になった系列があるのと同時に、その場その場でこなしていく新しい系列の両方がある。それを非常に抽象的に表現してサルに行わせている。例えば、learned hyperset のパフォーマンスは、日常生活で行っている歯を磨くとか服を着替えることに対応する。実際、一年半くらい課題をやらせないでおいて、久々にやらせても猿はちゃんと覚えていて、本当に植え付けられた系列となっている。

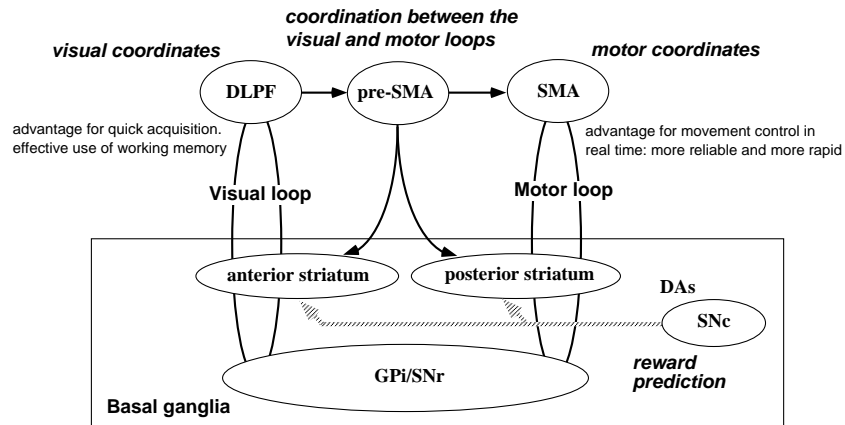


図 8: 大脳基底核の並行ループとその機能

これから、さまざまな実験とシミュレーションの対応を見て行くが、たとえば、線条体前部に GABA agonist であるムシキモル (muscimol) を注入して、その部分の神経活動を抑えたときには、learned hyperset に対する学習は全然変わらないのに、new hyperset に対する学習は悪くなる。それに対し線条体後部にムシキモルを入れたときには、learned hyperset に対するパフォーマンスは悪くなるのに、new hyperset に対するパフォーマンスは悪くならない。このように、様々な機能分化が実験的に観察が可能である。これらの実験に対し、我々は、多重表現仮説を提案し、その上で、そのモデルのシミュレーション結果と実験結果を比べることによって仮説の妥当性を検討した。

0.3.2 多重表現仮説

まず、多重表現仮説を手短かに説明しよう [14, 7, 16]。先ほど解剖学的な特徴について述べたが、なんのために、あのような分離した複数のループがあるのだろうか？機能分化を示している実験は、それぞれのループの違いに対応するのだろうか？それぞれのループでの領野の様々な実験での神経細胞

の反応の違いとどう関係するのだろうか？これらの問いかけを検討するなかで、各ループでの表現系、言い換えれば、座標系が重要なポイントなのではと考えたのである (図 6,8)。

同じ運動系列はさまざまな座標系を使って表現できる (図 7B)。たとえば、同じ系列を視覚座標系で表現することもできるし、腕の関節角度を使うような運動座標系で表現することもできる。それぞれの表現系にはそれぞれの特徴がある。視覚座標系の方が学習が早く運動座標系の方が学習は遅い一方で、運動座標系で学習してあれば運動はロバストにかつ迅速に遂行できるという特徴の違いがある。先ほどの actor-critic のスキームで考えると、それぞれのループというのは異なる座標系を持ちつつも、ドーパミン細胞からの報酬信号の影響を同様に受けている。つまり、異なる座標系をもった actor に対応する。ならば、視覚座標の方が学習が早いから、新しい系列を学習するときには、視覚座標を主に使った方が良いというのは想像がつく。

その一方で actor として異なる座標系、異なるループで学習を行っていたら、当然お互いの意見に相違が出る。また pre-SMA のデータを見てみると、様々なタスクの切り換えのときに発火が多いなどの結果がいろいろある。更に、pre-SMA というのは、visual loop(大脳皮質－大脳基底核の1つのループである prefrontal loop をここではそう呼ぶことにする) と motor loop の間にあってお互いを繋いでいる。具体的に言うと、皮質では、DLPF と SMA をつないでいるし、基底核の線条体でもこの各々のループの投射部である線条体前部と線条体後部の中間に投射している。この pre-SMA が個々のループのコーディネーションをしているだろうというのが、われわれの仮説のもう一つ重要な点である。

0.3.3 多重表現仮説に基づいた大脳基底核並行ループモデル

図9を見てもらいたい。この図では、大脳皮質だけで示しているが、基本的には我々が考えたのはこのような構造のネットワークモデルになる。視覚性入力が入るとDLPFに入ってきて、運動性入力が入るとSMAに入ってくる。DLPFからある形でコーディネータを介して情報がmotor networkに伝わる。DLPFでは視覚座標系で出力を決めているので、実際に運動を出力するには運動座標系にPMvを介して変換してM1を介して最終的な運動出力となるという構造になっている。

それぞれについてもう少し具体的に見ていこう。視覚座標系では二つの機能がある。visual immediate mapping というのは現在の視覚入力に対して出力を計算する。context prediction というのは、一つ前までの情報をもとにして、これから入ってくる入力に対する出力を予測する。visual context prediction であれば視覚座標系、motor context prediction であれば運動座標系でそれぞれ予測を行う。DLPFにもう一つあるといわれているのは作業記

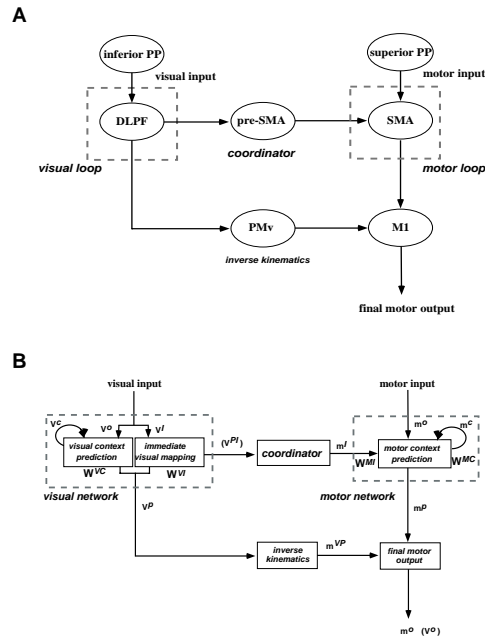


図 9: ネットワークモデル

憶である。その特徴は、短期的に inputs を保持しておいて、必要がなくなればすぐに消去できることにある。作業記憶という、短期的に記憶が保持できるとうことが良く注目される。そのアドバンテージは確かに大事だが、その場ですぐキャンセルできるというのも実は重要である。例えば各ブロックで各ハイパーセットの学習を行っているときに、一つのハイパーセットで覚えた入出力表現が次のブロックでは干渉する可能性がある。その意味で、作業記憶でその場で情報を保持して、次のブロックでその情報をキャンセルできるという機能が immediate mapping に入れてある。作業記憶をもう少し複雑にすることはできるが、このミニマムモデルでも十分実験との対応が見つかる。具体的には、immediate visual mapping は、ブロックごとに初期化されて学習する。また、その情報をコーディネータを介して motor context prediction に情報を送る。そして、その情報と motor context prediction の出力との和が取られて、motor network の出力が決められる。この motor network の出力と、PMv を介して送られてきた visual network の出力が MI で併されて、final motor output が決められ、それによって action が起きる。

以上の議論を数式として見ていこう。visual network では、immediate visual mapping で、現在の視覚入力 (v^I) に対する重み (W^{VI}) がかけられ、visual context prediction で、視覚座標に基づくコンテキスト情報 (v^C) に対する重み (W^{VC}) がかけられ、その和に対する softmax によって出力が表現される。

つまり、

$$\mathbf{v}^P(t) = S(\mathbf{W}^{VI}\mathbf{v}^I(t) + \mathbf{W}^{VC}\mathbf{v}^C(t))$$

で与えられる。ただし、 S は softmax function で $S(\mathbf{u})_i = e^{\zeta \mathbf{u}_i} / \sum_k e^{\zeta \mathbf{u}_k}$ を表す。visual network での状態空間は 16 次元で、現在の視覚情報 (\mathbf{v}^I) では、どこに点灯しているボタンがあるかという 0,1 のバイナリー表現になっている。visual context prediction の入力 (\mathbf{v}^C) は、視覚座標系での出力ベクトル \mathbf{v}^O と現在の visual context prediction の状態ベクトル \mathbf{v}^C の差の減衰和で表される。

$$\mathbf{v}^C(t+1) = \mathbf{v}^C(t) + \frac{1}{\tau_V}(\mathbf{v}^O(t) - \mathbf{v}^C(t))$$

ただし、 τ_V は減衰の時定数である。つまり、視覚座標系での過去の行動を時間的に荷重平均したものになっている。

motor network では、コーディネータから渡された入力 \mathbf{m}^I を、重み (\mathbf{W}^{MI} ; 実際にはこのモデルでは単位行列である) にかけて、それと運動座標に基づくコンテキスト情報 (\mathbf{m}^C) を motor context prediction の重み (\mathbf{W}^{MC}) にかけて、それらを加えたものに softmax をかけることで出力が得られる。

$$\mathbf{m}^P(t) = S(\mathbf{W}^{MI}\mathbf{m}^I(t) + \mathbf{W}^{MC}\mathbf{m}^C(t))$$

motor network の状態表現は、2次元の関節座標上に 64 個の正規化ガウス関数 (normalized Gaussian) を配置し、その activation によって表現される 64 次元のポピュレーションベクトルによって実現されている。ここではその詳細は省く。motor context prediction の入力は運動座標系での出力ベクトル \mathbf{m}^O と現在の motor context prediction の状態ベクトル \mathbf{m}^C の差の減衰和で表される。運動座標系での過去の行動を時間的に荷重平均したものになる。

$$\mathbf{m}^C(t+1) = \mathbf{m}^C(t) + \frac{1}{\tau_M}(\mathbf{m}^O(t) - \mathbf{m}^C(t))$$

τ_M は減衰の時定数である。

最後の motor output の選択は、motor network の出力と visual network の出力を運動座標系になおした後で、積による平均をとっている。

$$p_j(t) = \frac{\mathbf{m}_j^{VP}(t)\mathbf{m}_j^P(t)}{\sum_k \mathbf{m}_k^{VP}(t)\mathbf{m}_k^P(t)} \quad (1)$$

これは情報幾何でいう e-projection になっている。実際に決められる系列は、この運動座標系で表現された式の確率分布に基づいて、どのボタンを押すかランダムに選ぶ。また、各ボタンを押すときに毎回ネットワークが回る。

学習はこの 3 つのベクトル ($\mathbf{v}^I, \mathbf{v}^C, \mathbf{m}^C$) に対する重み ($\mathbf{W}^{VI}, \mathbf{W}^{VC}, \mathbf{W}^{MC}$) に関して actor-critic のスキーマで TD 学習を行う。それぞれの学習則は次のようになる。

$$W^{VI}(t+1) = W^{VI}(t) + \eta^{VI} \hat{r}(t) g(\mathbf{v}^O(t), \mathbf{v}^P(t)) \mathbf{v}^O(t) \{\mathbf{v}^I(t)\}^T$$

$$W^{VC}(t+1) = W^{VC}(t) + \eta^{VC} \hat{r}(t) g(\mathbf{v}^O(t), \mathbf{v}^P(t)) \mathbf{v}^O(t) \{\mathbf{v}^C(t)\}^T$$

$$W^{MC}(t+1) = W^{MC}(t) + \eta^{MC} \hat{r}(t) g(\mathbf{m}^O(t), \mathbf{m}^P(t)) \mathbf{m}^O(t) \{\mathbf{m}^C(t)\}^T$$

ここで、 $\eta^{VI}, \eta^{VC}, \eta^{MC}$ は学習係数、 $\hat{r}(t)$ が TD 誤差である。 $g(\mathbf{p}, \mathbf{q})$ は softmax のゲイン行列であり、この第 i 対角成分は $(p_i - q_i)q_i(1 - q_i)$ で表され、他の成分は 0 である。この学習則は、 $\|\mathbf{p} - \mathbf{q}\|^2$ の自乗誤差に対する、TD 誤差に修飾された stochastic gradient descent を実現している。

我々の仮説では、visual network の学習が motor network よりも早いことがネットワーク全体の特性に大きく寄与すると考えている。この visual network の学習が早い理由の一つとして、運動座標系の不良設定性 (ill-posed) があるが、上のモデルではそれは入れていない。そのため、モデルでは、visual network の学習が motor network よりも早いというのはパラメータを調整することで実現した。その方法は、new hyperset を一つもってきては、それについて 20 回のハイパーセットのトライアルを成功するまでにしたエラーの数を指標にして、学習速度をチェックするというを繰り返して、パラメータ調整を行なった。

0.3.4 実験結果とシミュレーション結果の比較検討

では、シミュレーションと実験の結果の対応を見ていこう (図 10)。まず 1 つのハイパーセットがどのように学習されるか見ていこう。実験とモデルの結果を比較するための指標は、一つのブロック (つまり一つのハイパーセット) が終了するまでの試行回数、または誤り回数である。当然のことながら、誤り回数と成功試行回数の和が、試行回数となる。さて、実験で見られたこととしては、"short term" の学習がある。これは 1 ブロック内でハイパーセットを学習していくときに、最初に間違いを起こしながら、後半のほうが error が少ない。そして、あるハイパーセットが learned hyperset になっていくときに何日にも渡って徐々に学習成果が現れてくるという "long term" の学習がある。この 2 つの学習は、シミュレーションでも見られる。

次に、我々の仮説の一つをテストした。上述のパラメータの調整は、一つ一つの new hyperset に対するパフォーマンスを利用して行われた。しかし、我々の日常はそうではない。つまり、次から次へと new hyperset も来るし、learned hyperset も来る。このような状況のもとで、ネットワークが振舞うかは、必ずしも定かではない。それで、そのような状況で、モデルがどう振舞うか、またパラメータチューニングした visual network のほうが motor network よりも早いかを調べた。visual network のみを用いたモデル (visual only) と motor network のみを用いたモデル (motor only) というモデルを使用して確認した (図 11)。すると、十分 visual only のほうが早かった。更に、この早いのと遅いのをまず単純につなげてみた。これを no-cordinator と呼ぶ。そのパフォーマンスはほぼ visual only と motor only の中間を示している。早いの

0.3. 複数の皮質－大脳基底核－視床回路の共同による逐次運動系列の学習, 制御19

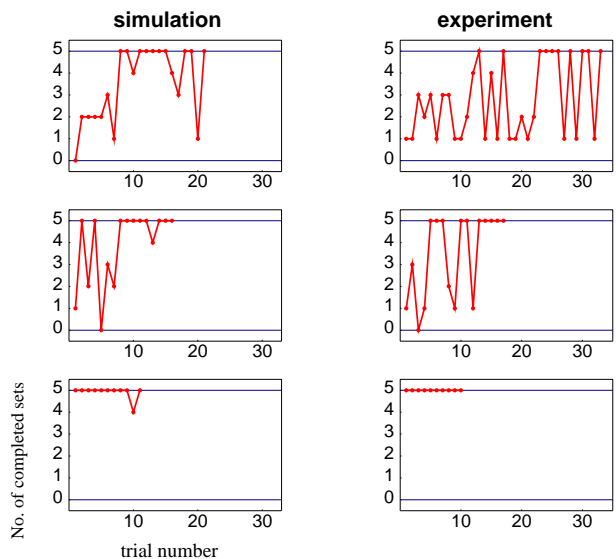


図 10: 実験 (右) とシミュレーション (左) での 1 つのハイパーセットのパフォーマンス例。上から下に経験した日数が増えている。

と遅いのを足しあわせたので、中間になるのはわかるような気がするが、一般には必ずしもこうなるとは限らないので、それを確認したわけである。この no-coordinator のままだと visual only のパフォーマンスを悪くしただけだが、コーディネーターを入れると、つまり我々のモデルだと、visual only よりもパフォーマンスはよくなっている。これは、お互いのコーディネーションを入れることの御利益がこのモデルでは現れていることを示唆している。

もう一つモデルのテストをしたのは、working memory をリセットできることが重要であるという仮説である。我々のモデルには、それが単純化された形で入っていた。どのようにテストしたかという、我々のモデルでは、immediate visual mapping を各ブロック (各ハイパーセット) ごとに初期化している。ここでは結果だけ述べておくと、この初期化なしでは、learned hyperset に対してもパフォーマンスは悪くなっているが、new hyperset に対するパフォーマンスが特に悪くなっていることが分かった。これによって、working memory をリセットできることが特に新しい系列の学習に重要であることが分かった。

以上のテストで、モデルの基本的なふるまいをまずチェックした。その上で、以下に述べるいくつかの実験結果とモデルのふるまいの比較検討を行った。

まずテストしたのは系列依存性である。系列学習で重要なのは、その系列をどのように学習しているかである。例えば、ハイパーセットの学習で、各セットの中で、このパターンの視覚刺激が来たときにはこの順番でボタンを押すとい

Training simulation

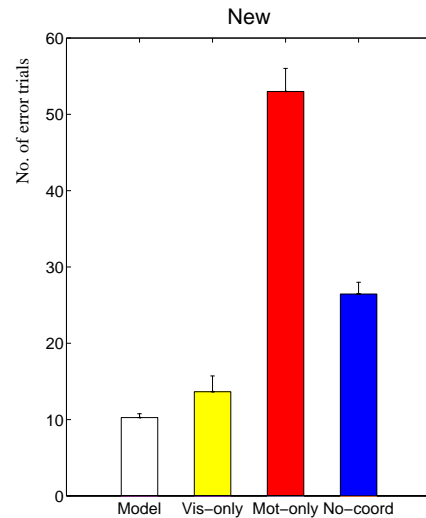


図 11: 各ネットワークの比較

う学習、つまり系列に依存しない学習も可能である。そこで、learned hyperset の系列をひっくり返したハイパーセット (reverse hyperset、図 12B) に対するパフォーマンスを調べた。この結果 (図 12A) を見ると、reverse hyperset に対するパフォーマンスは、ほとんど new hyperset に対するパフォーマンスと変わらない。つまり、learned hyperset の学習は非常にコンテキスト依存、系列依存、で学習されていることがはっきりわかる。

次に、opposite hand simulation。たとえば、アメリカにいて車に乗って、ウインカーを使おうとして思わずワイパーを回してしまう。つまり、習熟したスキル (系列学習) は、エフェクターに依存していることが多い。右利きの人が左手でボールを投げるときこちない。この場合、利き腕の問題があるが、まあそういうことを考えるとわかる。実際に実験で個々のハイパーセットを学習させるときには、このハイパーセットは右手でやらせる、このハイパーセットは左手でやらせるという形で実験を行っている。なお、サルは、人間ほど利き腕が強くない。そこで、一方の手で十分学習したハイパーセットを他方の手でやらせたらどうなるかということが調べられた (図 13)。我々のモデルでは、motor network がエフェクター依存の機能と考えられるので、motor network の記憶を初期化してやることでシミュレーションを行った。この結果が示しているのは、learned hyperset の結果はエフェクター依存なので、手を変えるとそのパフォーマンスは悪くなる。ただし、全く新しい系列に対するパフォーマンスよりは十分良い。即ち、一定程度エフェクター非依存な記憶を持っているということが示された。

次に blockade simulation の話をしよう。これは、GABA agonist である

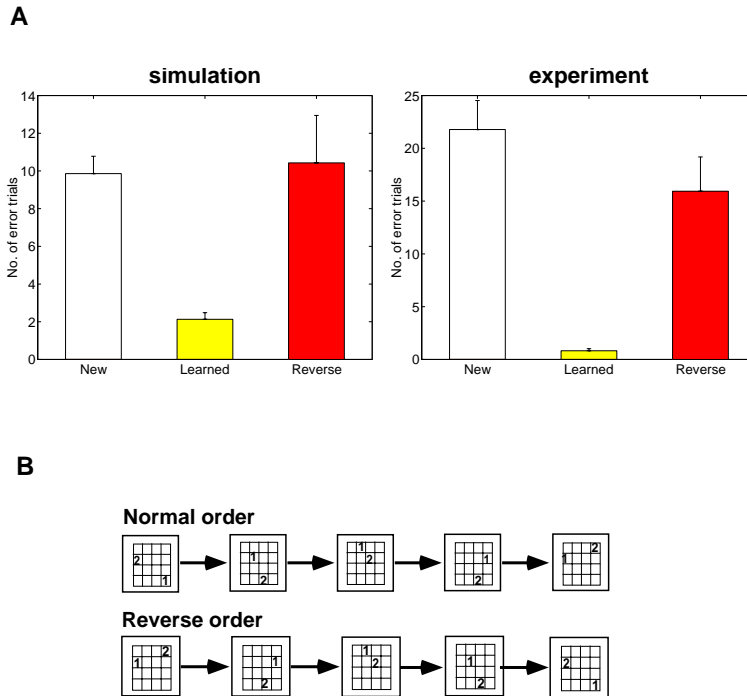


図 12: reverse hyperset simulation

ムシキモル (muscimol) を脳の一部の部位に注入することで、その部分の神経活動を抑えたときの実験結果に対応したシミュレーションである。visual network を blockade している場合、motor network を blockade している場合、coordinator を blockade している場合の3つの場合をテストした(図 14)。それぞれの場合のシミュレーション結果と実験結果を比べるとまあまあ合っているということを示していく。

まず visual blockade から見ていこう(図 15)。実験では線条体前部を blockade した場合に対応する。シミュレーションの結果は、learned, new 両方に対してパフォーマンスが悪くなっている、ただし、new に対するパフォーマンスはより強く悪くなっている。これは次の motor blockade と比較するとはっきりする(図 16)。motor network の blockade も new と learned を両方悪くする。ただし、visual blockade と比較すると learned に対するパフォーマンスがより悪くなっている。実験を見ると、線条体後部を blockade すると learned のみに対して有意に悪くなる。一方で、SMA では new に対してのみ有意に悪くなる。この意味では、我々のモデルは実験と完全な一対一対応はまだしていない。ただし、我々のモデルではネットワークで blockade していて、各ネットワーク内の領域の区別はしていない。これはまだ改善の余地があると言える。コーディネータの blockade を見よう(図 17)。これは実験で

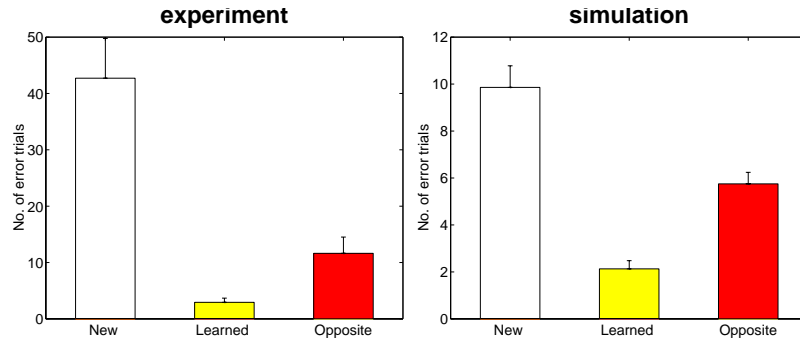


図 13: opposite hand simulation

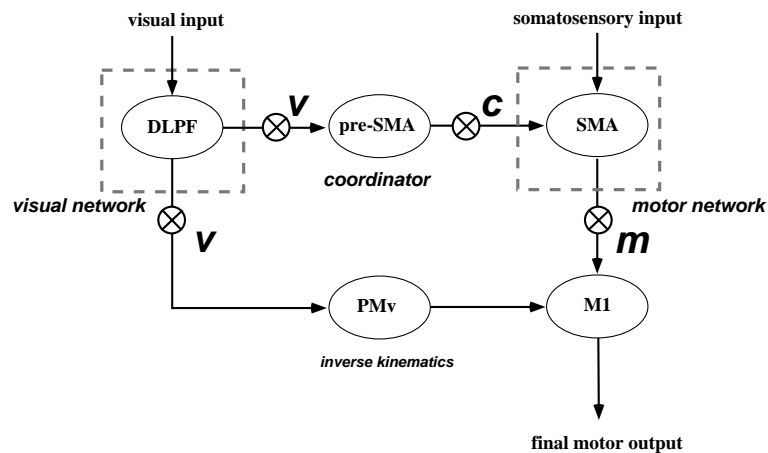


図 14: blockade シミュレーション。(V; visual blockade, Mj; motor blockade, C; coordinator blockade)

いうと pre-SMA の blockade にあたる。実験・シミュレーションともに、new に対するパフォーマンスのみが有意に悪くなっている。このように、それぞれの blockade でモデルのシミュレーション結果と実験結果の間で非常に良い対応が見つかった。

最後に、ドーパミンの投射を消失させた場合のシミュレーションについて述べておこう。Visual network、motor network 各々に入るドーパミンの投射を抑えた場合を調べた。このどちらの場合でも、learned hyperset に対しては、パフォーマンスが落ちるということはない。一方、new hyperset では、motor network への投射が押さえられても、パフォーマンスは悪くならなかったが、visual network への投射が押さえられた場合には、パフォーマンスは有意に悪くなった。これについては実験結果はまだないが、調べてみるのは面白いのではないかと思う。

0.3. 複数の皮質－大脳基底核－視床回路の共同による逐次運動系列の学習, 制御23

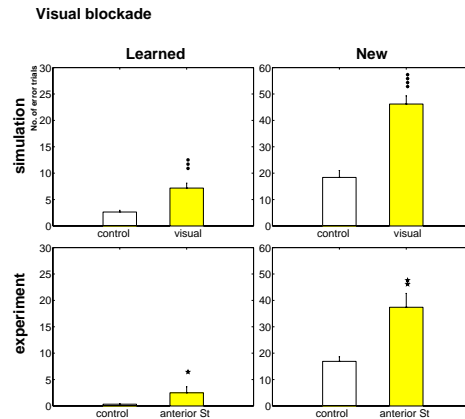


図 15: visual blockade

0.3.5 まとめ

以上まとめると、この研究は、大脳皮質－大脳基底核回路はパラレルに働き、そのループのそれぞれのサブループは、異なる座標系を利用することによって、新規の系列の学習という課題と習熟した系列の遂行という課題を共にうまく両立させて実現しているという、多重表現仮説を提案し、検証しました。その結果として、我々のモデルが 2×5 課題ののさまざまな実験結果と対応する振舞いをすることを示した。

0.3.6 質疑応答

Q：visual network の学習が motor network より早いと考える理由は何か。

中原：まず現在実装したモデルではパラメータチューニングで実現しているので、その理由自体はモデルの中にはない。一般的にそのように考える理由としては、視覚座標系と運動座標系は一般に不良設定な関係にあるので、視覚座標系のほうが学習が早いだろうと考えている。というのは、不良設定というのは、座標系に関する不良設定もあるが、学習に関する不良設定もある。学習における不良設定というのはつまり学習が不安定になってしまうことである。おそらく座標系における不良設定、一対一の対応が見つからないということが学習における不安定性を引き起こすので、学習における不安定性を押さえるには運動座標系で十分学習係数を低くする必要があると推測している。いつかきちんとそこらへんも研究したいと思っている。

Q：なぜ図 8 では、pre-SMA から SMA の一方向に矢印が伸びているのか。

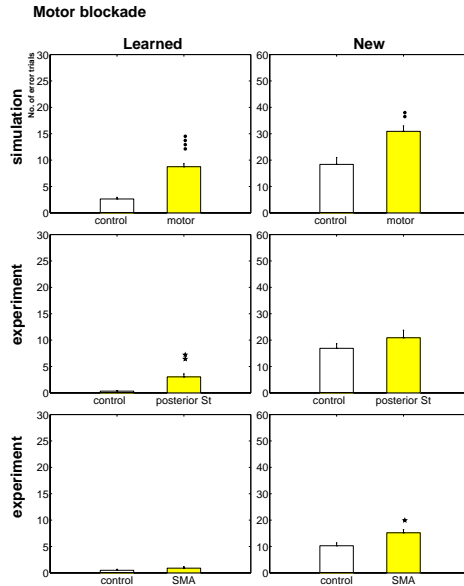


図 16: motor blockade

中原：確かに、pre-SMA がコーディネートしているという立場から考えると両向きのほうがわかりやすいと思う。実際、解剖学から言えば、両向きである。ここで、一方向にしているには理由があって、一つは pre-SMA のニューロンの反応を見ていると、現在の視覚入力が入ってきて行動を起こすまでの間に反応しているニューロンが結構多い。それに基づいて現在のモデルではこの一方向になるように実装している。この機能のモデルを、最近の実験と照らし合わせながら、より洗練させていくのは面白い課題だろう。

Q：SMA の結果はなぜ合わないと言ったのか。また、今後の進展は？これからどのようにニューロンモデルなり学習則を変えていこうとしているか。

中原：今日話したことは、基本的には誤り回数に関することだったが、実際の実験では performance time とか、実制御の時間に対する実験結果も得られている。SMA の結果を見ると誤り回数の結果に関しては、実は new に対して悪くなっている。ただ、その一方で、performance time で見ると、学習された系列に対して SMA の blockade が影響しているであろうという結果がある。我々のモデルでは、posterior putamen と SMA を区別せずに learned に対して誤り回数が特に増えるという結果になっている。そういう意味で、完全には一致していない。このように、実時間の話と誤り回数の話をどう組み込むか。やってみるべきこと

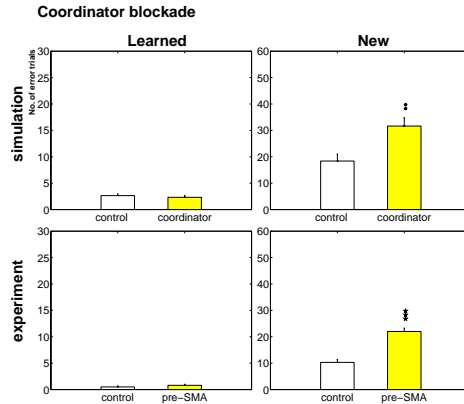


図 17: coordinator blockade

はいろいろある。

Q: この話の short term の学習と long term の学習についての基本的な考え方は、運動学習には限らない非常に広い話に還元されるような気がするし、線条体以外の関係してくるべき部分はたくさんあると思われる。このモデルを外挿していくときに、このモデルの中で最低限含めなければならない拘束条件は何か。

中原: 運動の実制御の話だと小脳を考慮に入れるべきである。大脳－小脳のいくつか別れたループとの対応がつかののかということも重要である。このモデルを拡張していくときの制約については、どのレベルで考えるかによる。基本的には異なる座標系で共同して強化学習するという意味では外挿していくことは十分可能だが、cognitive science の分野あるいは、cognitive neuroscience のレベルまたは、cellphysiology に近いレベルで外挿していくのかで考えるべき優先課題が違ってくると思う

最後に

この前半の研究は、甘利俊一(理研)・彦坂興秀(NIH)との共同研究で、後半は、銅谷賢治(ATR)・彦坂興秀(NIH)との共同研究です。これらの方々の御指導、また後半の研究に関しては博士論文の指導教官の永野三郎(埼玉工大)の御指導に感謝します。本章の詳細な内容は各々の原論文 [15] と [16] を参照ください。この二つの研究を行う際(実際には現在もそれぞれ更なる研究が進行中ですが)には、彦坂研究室の多くの方にお世話になりました。あまりにも多数になるので名前を列挙するのは差し控えますが、彦坂研究室の方々のご協力がなければこれらの研究は不可能でした。本文中では文献の引

用は最小限にとどめています。本章で紹介した個々の実験結果には、それぞれ原論文があり、その内容はここで紹介したよりもずっと豊かです。紹介できていない実験も、新しいモデル研究もあります。本章の内容に興味を持たれた方は、ぜひ原論文に当たられることを勧めます。前半の研究については、現在もぞくぞくと新しい結果が出ていて、総説もまだない状態です。PubMedなどのデータベースで探してみてください。後半の研究については、[10, 7]が詳しい総説です。また、それから数年間の研究をフォローした総説が[8]にあります。これらの総説には、本章に直接関連する研究や、または本文中では紹介できなかった他の研究者の興味深い研究が紹介してあります。そこから原論文に当たられるのが良いでしょう。

最後に、この講義録の原案 (BSIS Technical Report No.01-02、No.03-01にも登録) は、森本淳 (当時、奈良先端大)・石田文彦 (当時、東北大) によって起こされたものに加筆しました。お二人に感謝します。

関連図書

- [1] G. E. Alexander, M. D. Crutcher, and M. R. DeLong. Basal ganglia-thalamocortical circuits: Parallel substrates for motor, oculomotor, “prefrontal” and “limbic” functions. In H. B. M. Uylings, C. G. Van Eden, J. P. C. De Bruin, M. A. Corner, and M. G. P. Feenstra, editors, *Progress in Brain Research*, volume 85, chapter 6, pages 119–146. Elsevier Science Publishers B.V., 1990.
- [2] S. Amari. Neural theory of association and concept-formation. *Biological Cybernetics*, 26:175–185, 1977.
- [3] S. Amari. Field theory of self-organizing neural nets. *IEEE Transaction on Systems, Man and Cybernetics*, SMC-13(9 & 10):741–748, 1983.
- [4] S. Amari and A. Takeuchi. Mathematical theory on formation of category detecting nerve cells. *Biological Cybernetics*, 29:127–136, 1978.
- [5] P. Calabresi, A. Pisani, N. B. Mercuri, and G. Bernardi. The cortico-striatal projection: From synaptic plasticity to dysfunctions of the basal ganglia. *Trends in Neurosciences*, 19:19–24, 1996.
- [6] S. Hernandez-Lopez, J. Bargas, D. Surmeier, A. Reyes, and E. Galaraga. D1 receptor activation enhances evoked discharge in neostriatal medium spiny neurons by modulating an l-type ca^{2+} conductance. *Journal of Neuroscience*, 17(9):3334–42, 1997.
- [7] O. Hikosaka, H. Nakahara, M. Rand, K. Sakai, X. Lu, K. Nakamura, S. Miyachi, and K. Doya. Parallel neural networks for learning sequential procedures. *Trends in Neuroscience*, 22(10):464–471, 1999.
- [8] O. Hikosaka, K. Nakamura, K. Sakai, and H. Nakahara. Central mechanisms of skill learning. *Current Opinion in Neurobiology*, 12:217–222, 2002.
- [9] O. Hikosaka, M. K. Rand, S. Miyachi, and K. Miyashita. Learning of sequential movements in the monkey: Process of learning and retention of memory. *Journal of Neurophysiology*, 74(4):1652–1661, 1995.

- [10] O. Hikosaka, K. Sakai, H. Nakahara, X. Lu, S. Miyachi, K. Nakamura, and M. Rand. *The New Cognitive Neurosciences*, pages 553–572. The MIT Press, 2000.
- [11] J. E. Hoover and P. L. Strick. Multiple output channels in the basal ganglia. *Science*, 259:819–821, 1993.
- [12] J. C. Houk, J. L. Adams, and A. G. Barto. A model of how the basal ganglia generate and use neural signals that predict reinforcement. In J. C. Houk, J. L. Davis, and D. G. Beiser, editors, *Models of Information Processing in the Basal Ganglia*, pages 249–270. MIT Press, Cambridge, Massachusetts, 1995.
- [13] R. Kawagoe, Y. Takikawa, and O. Hikosaka. Expectation of reward modulates cognitive signals in the basal ganglia. *Nature Neuroscience*, 1(5):411–416, September 1998.
- [14] H. Nakahara. *Sequential Decision Making in Biological Systems: The Role of Nonlinear Dynamical Phenomena in Working Memory and Reinforcement Learning in Long-Term Memory*. PhD thesis, Univ. of Tokyo, 1997.
- [15] H. Nakahara, S. Amari, and O. Hikosaka. Self-organization in the basal ganglia with modulation of reinforcement signals. *Neural Computation*, 14(4):819–844, 2002.
- [16] H. Nakahara, K. Doya, and O. Hikosaka. Parallel cortico-basal ganglia mechanisms for acquisition and execution of visuo-motor sequences - a computational approach. *Journal of Cognitive Neuroscience*, 13:1–22, 2001.
- [17] W. Schultz. Predictive reward signal of dopamine neurons. *Journal of Neurophysiology*, 80:1–27, 1998.
- [18] W. Schultz, P. Apicella, and T. Ljungberg. Responses of monkey dopamine neurons to reward and conditioned stimuli during successive steps of learning a delayed response task. *Journal of Neuroscience*, 13(3):900–913, 1993.
- [19] T. Trappenberg, H. Nakahara, and O. Hikosaka. Modeling reward dependent activity pattern of caudate neurons. In *International Conference on Artificial Neural Network (ICANN98)*, pages 973–978, 1998.
- [20] J. Wickens. *A Theory of the Striatum*. Oxford: Pergamon Press, 1993.