

LECTURES

“意志決定と脳”

—自由意志をとらえることはできるのか—

玉川大学学術研究所脳科学研究施設

坂上 雅道

要 旨

私たちは、環境の中の刺激にたいして反応しその結果として報酬を得る。これを繰り返すことによって形成された情報の履歴が今行なおうとする行動を決定する。生物は、それなりの能力で環境の効用を理解し、それに最も適した形の行動決定をしているようである。この効用の理解は自らを含む環境の変化に対応できるダイナミックなものでなければならない。脳・神経系はこのための計算を行なっている、という見方がある。最近、システム神経科学の世界でこのような「意思決定」に関わる研究が急速に盛んになってきた。この背景には、経済学におけるゲーム理論の展開、心理学における効用と選択行動に関する研究の興隆、神経生理学における脳の報酬系情報処理研究の進展などがある。まだ、始まったばかりの新しい研究領域であるが、ここではその動向について最近発表された論文をとりあげて検討してみる。

キーワード：意思決定、自由選択課題、注意、報酬確率、効用

1. はじめに

17世紀の哲学者、René Descartesは、動物の行動には単純なものと複雑なものの2種類があると考えた。単純な行動とは環境中の刺激に対して反応が1対1に決定されるものであり（これは後の反射という概念に近い）、複雑な行動とは刺激に対して反応が予測できないようなものである。Descartesは心身二元論を唱えた哲学者であり、動物は機械的に構成される「体」と機械的には説明できない「心」から出来上がっていると考えた。単純な行動は機械的な構成部分である「体」の仕組みで説明できるが、複雑な行動はその決定に機械的には説明不可能な「心」が関与するため科学的には説明できないものであるとした。19世紀に出現した実験心理学（心理物理学）が「心」の研究を科学の土俵にのせるきっかけを提供して以

来、20世紀の後半になってようやく脳科学もDescartesのタブーの領域に踏み込めるようになってきた。21世紀に入って、未熟ではあるが、脳科学は「自由意志」をも研究の対象としてとらえはじめた。

2. 高次脳機能研究と単一ニューロン活動記録法

「心」の関与がある場合には、覚醒した状態で脳活動を調べる必要がある。1960年代のNIH（National Institute of Health）のEdward Evartsらによる無麻酔ザルを使った単一ニューロン記録法の開発は、高次脳機能研究における新しい時代の幕開けとなった。麻酔をせずに最低限の行動的制約のもと脳から単一ニューロン活動を記録することにより、運動中の脳活動を調べることはもとより、認識や思考も科学的な脳研究の対象とする

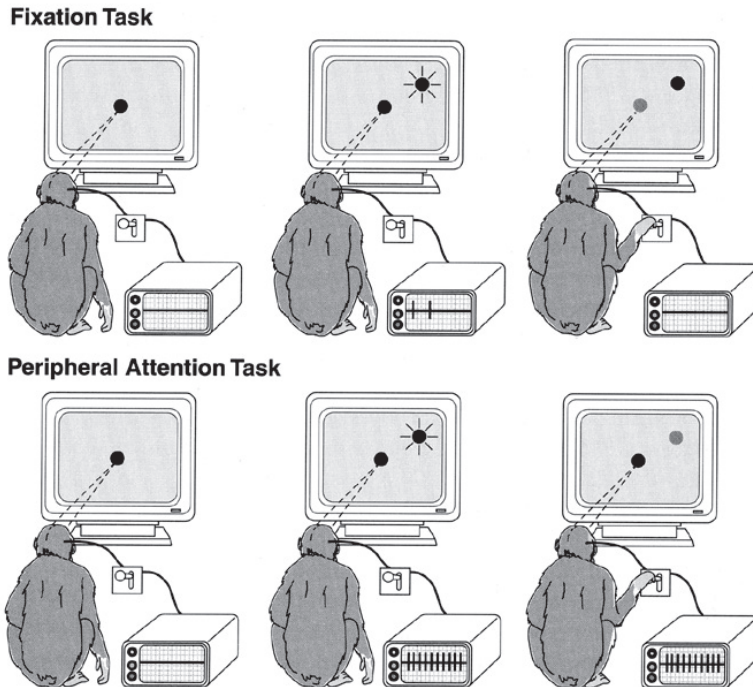


図1. Goldbergらの実験で使われた空間的注意課題。
(文献8, 238ページFigure 10.5を改変)

ことができるようになった。

この方法を使って、1970年代にはNIHのWurtzとGoldbergのグループによって脳研究に“注意（attention）”という概念が導入されることになる。注意という概念は、ヒトや動物が物理的には同じ刺激環境の中で状況に応じて感覚情報処理の仕方をかえるということを説明するものである。彼らは視覚刺激を手がかりに運動を行わせる課題をサルに学習させ、その課題遂行中に頭頂葉からニューロン活動を記録した（図1）。スクリーン中央に小さな視覚スポット（注視点）が呈示され、サルはこれを見つめることを学習する。課題はこのスポットが暗くなったらレバーを押すことである。ある条件（fixation task）では周辺視野に別の刺激が呈示されるが、サルはそれを無視しなければならない。別の条件（peripheral attention task）では、同様に周辺視野に刺激が呈示されるが、確率50%で周辺刺激も暗くなることがあり、その場合は注視点も暗くなった場合

と同様すぐにレバーを押すことが要求される。どちらの条件でもレバーを押すまでサルは注視点を見つめ続けなくてはならない。刺激も反応も報酬も2つの条件間では全く同じであるが、fixation taskでは周辺刺激を無視して良いが、peripheral attention taskでは周辺刺激に“注意”を払わなければならない。周辺刺激がニューロンの受容野に呈示された場合、1次視覚野のニューロンはどちらのtaskでも同様の応答を示すが、頭頂連合野（7野あるいはLIP, lateral intraparietal area）のニューロンはfixation taskの場合よりperipheral attention taskの場合のほうが周辺刺激に強く応答する。つまり注意を払うと頭頂連合野のニューロン活動は増強される。このような頭頂葉の機能をGoldbergらは内的な行動情報による感覚情報処理の修正と考えた [1, 2]。

Goldbergらの“注意”の研究は、同じ刺激環境にたいする異なる「心」の状態をニューロン活動のレベルで捉えたという意味で画期的であっ

た。これは、まさにDescartesの言うところの神の領域である「複雑な行動」に科学のメスをいれたことになる。

一方、頭頂連合野のニューロン活動を運動情報処理の一部ととらえようとしたMountcastleの流れをくむRichard Andersenのグループは、頭頂連合野ニューロンの活動を“意図”(intention)で説明しようとした。たとえ実際に運動しなくても、たとえば、眼球をそちらに動かす準備として、頭頂連合野は活動するのだという主張である。このように考えて、AndersenらはGoldbergたちの結果は“意図”で説明できるとした[3]。

2つのグループの論争にたいして、PlattとGlimcherはサルに2つの刺激を呈示しそのうちの1つを選択して反応させるという実験を行った[4]。呈示される刺激は、ある時には標的に、またある時には標的に反応するための手がかりになった。この時のLIPニューロンの活動は、注意説でも意図説でも説明できないものであった。彼らは、LIPのこのようなニューロン活動は、行動選択あるいは意思決定(decision-making)機能を反映するものであると考えた。

3. マッチングの法則

Richard Herrnsteinは、ハトに異なる色の2つのキーをつつく訓練を行った[5]。1つのキー(左キー)はVI(variable interval)135秒で、もう1つのキー(右キー)はVI270秒のスケジュールで餌が出てくる。インターバルスケジュールとは、餌をもらってから一定時間(待ち時間)経過した後の最初のキーつつき反応で再び餌がもらえるという、反応にたいして餌を与えるスケジュールである。VI135秒とは、この待ち時間が一定ではなく、平均135秒であることを示す(VI270秒とは平均の待ち時間が270秒)。2つのキーのタイマーは独立に動いており、ハトが2つのキーに均等にキーつつきすると、左キーからは1時間当たり約27回(1時間で可能な分の67%)、右キーからは約13回(33%)報酬を受けることができる。過渡的な状態を経て定常状態になると、ハトは報酬を受けることができる割合にマッチして、左キ

ーを約67%、右キーを約33%の割合でつついた。Herrnsteinは左右のVIスケジュールの割合をいろいろ変化させてみた結果、ハトのキーつつきにおける選択行動には、次の「マッチングの法則」とよばれる一般法則が成り立つことを見出した。

$$P1/(P1+P2) = R1/(R1+R2)$$

P1は左キーへの、P2は右キーへのキーつつき反応数。R1は一定時間内に左キーに反応することによって得られた報酬の数、R2は右キーへの反応によって得られた報酬数。マッチングの法則はハトに限らず、ラット、ウシ、サル、ヒトと広範囲の動物の行動に当てはまることがその後の研究で明らかにされた[6]。また、上の例では報酬を得る時間的割合が左右キーの違いであったが、報酬の量や質の違いにたいしてもマッチングの法則が成り立つことが知られている[6]。これらのことは、動物がそのときの環境に潜在する労働と報酬の関係にあわせて行動を決定することができることを示している。そのために動物は報酬の量や質を考慮するだけでなく、環境にたいして行った行動の結果の履歴をつくり確率的に反応を生成できなければならない。

* キーにたいする反応の数で報酬がきまるFR(fixed ratio)やVR(variable ratio)ではマッチングの法則は成り立たず、maximizing(報酬率の高いほうへの反応の全面的偏り)がおこる。また、VIスケジュールでも通常交互に左右のキーをつつく傾向が出てくる場合がある。この反応の切り替えをふせぐために、Herrnstein(1961)やその後の研究では、“COD(changeover delay)”を用いている。これは、一方のキーに反応した直後にもうひとつのキーに反応したら、ペナルティーとして一定時間の報酬の遅延がおこるというものである。

4. 意思決定とニューロン活動

確かに、同じ刺激にたいして反応やニューロン応答が異なるという意味では、注意も外からは見えない何らかの心の状態を反映しているといえるが、課題を学習・実行する被験体の側からいうと、

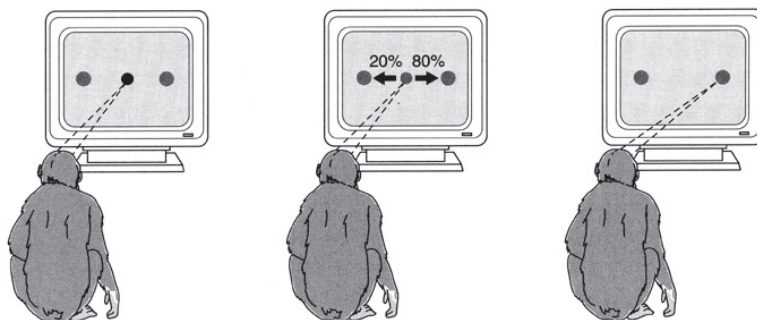


図2. 自由選択課題 (free choice task).
(文献8, 257ページ Figure 10.11 を改変)

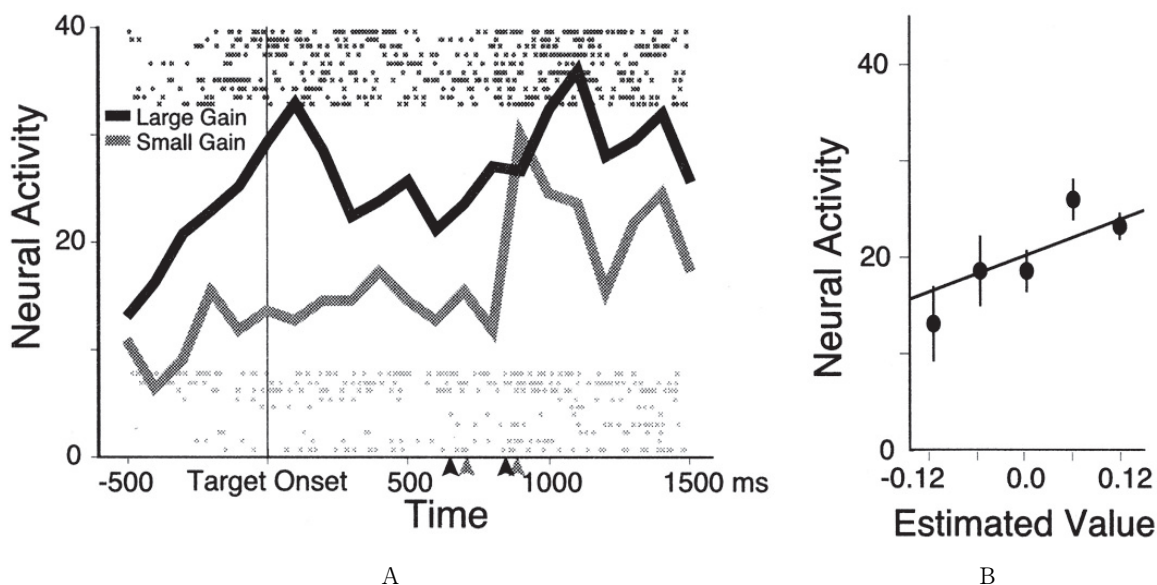


図3. 自由選択課題における頭頂連合野ニューロンの活動.

(A) 受容野内に呈示された刺激を選択したときの同一ニューロンの活動

Large Gain ; 高報酬ブロック, Small Gain ; 低報酬ブロック

矢印はそれぞれのブロックでの注視点の消失と眼球運動の開始の平均時間を示す.

(文献8, 266ページ Figure 10.14 を改変)

(B) 反応から得られた報酬値とそれを選択したときの LIP ニューロンの平均発火

(文献8, 267ページ Figure 10.15 を改変)

環境は多少複雑ではあるが、餌を得るためには他の選択肢はない。つまり、通常行われている注意の実験では、同じ刺激にたいして異なる反応が選択肢として用意されているが、同じ課題状況では報酬を得るために必ず同じ反応を行わなくてはならない。したがって、「注意」に関係している

といわれるニューロンは、単一の刺激ではなく複数の刺激の連鎖と反応との関係をコードしているだけではないかという批判が常につきまとう。

この節では、動物の「主体的判断」が脳でどのように作りだされているのかという問題にせまろうとする最近の3つの研究について紹介する。

< Platt & Glimcher, 1999 >

Michael Platt と Paul Glimcher はマッチングの法則にしたがって行動するサルの前頭葉 (LIP) からニューロン活動を記録する実験を行った[7]。課題は、眼球運動によって2つの標的刺激のうち自由に一方を選択する自由選択課題 (free choice task) である (図2)。目の前に置かれたディスプレイの中央に注視点が呈示され、サルがそれをしばらく見つめていると注視点の両側に2つの標的刺激が呈示される。中央の注視点が消えるのを合図に、サルはどちらかの標的刺激に目を動かして選択する。原則的にはどちらを選択することも許されるが、左と右では報酬の量が異なっている。するとサルは前節で述べたマッチングの法則にしたがって、報酬の量の割合で左右の標的刺激を選んだ。そのときの典型的な LIP ニューロンの活動例が図3A に示されている。黒の実線・灰色の実線ともにサルがニューロンの受容野に呈示された標的刺激にたいして眼球運動を行なったときのニューロン活動を示しているが、黒は報酬が多いときのブロックでの応答、灰色は報酬が少ないときのブロックでの応答である。この課題では一定の試行数のブロックごとに左反応での報酬量、右反応での報酬量が決まっているので、サルは数試行課題を行なってみるとブロック内でのそれらの報酬量がわかる (刺激が呈示される前からどちらの刺激の報酬量が多いか知っている)。同じ刺激にたいする同じ運動でも、その知識に比例して LIP ニューロンのスパイク数は変化してくる。さらに Platt と Glimcher は、報酬量をパラメトリックに変化させて LIP ニューロンの応答を調べているが、選択行動から推測されるサルの主観的な報酬価がスパイク数に比例することを見出した (図3B)。一連の実験から、Glimcher は行動決定 (意思決定) の要因として、(1) 報酬の頻度と (2) 報酬の価値 (質と量) の2つが重要であり、LIP ニューロンは注意や意図をコードするというより、確率的知識をもとにした意思決定プロセスの結果を反映していると考えた [8]。

< Sugrue, Corrado & Newsome, 2004 >

Platt と Glimcher の実験を発展させる形で、

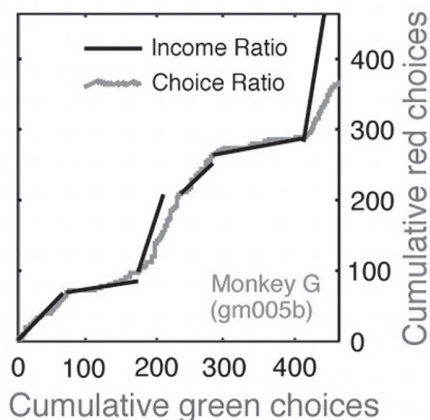


図4. 報酬比率を変化させたときの選択行動の変化。
(文献9, Figure 1Bを改変)

Sugrue, CorradoそしてNewsomeはLIPニューロンが選択行動とそれにとまなう報酬の履歴から意思決定に関わる「効用 (utility)」をいかにコードするかについての研究を行なった [9]。使われた課題はPlatt & Glimcher (1999) の課題とほぼ同じ自由選択課題である (ただし、2つの標的刺激は赤と緑で、それぞれの色に眼球運動を行なったときの報酬確率の比が変えられた)。図4は赤選択と緑選択の時の報酬比率を1:1, 1:3, 3:1, 1:1, 1:6, 6:1と変化させていったときの赤選択行動 (縦軸) と緑選択行動 (横軸) の累積反応数を示したものである。サルは短期的な経験をもとにマッチングの法則にしたがった行動パターンを示していることがわかる。このときのLIPニューロンの活動は受容野の中の刺激を選ぶか受容野から離れた位置に呈示された刺激を選ぶかで大きく異なる。受容野内の刺激に眼球運動を行う場合は、刺激呈示から実際の眼球運動が起こるまでの間、多くのLIPニューロンが強く活動する。このときの活動の強さは、選ばれなかった刺激との報酬比率に比例した。すなわち、選ばれた刺激の報酬比率が高いほど活動は強く、低いほど弱い (図5)。一方、受容野から離れた位置に呈示された刺激が選択された場合は、逆にその刺激の報酬比率が高いほどLIPニューロンの活動は

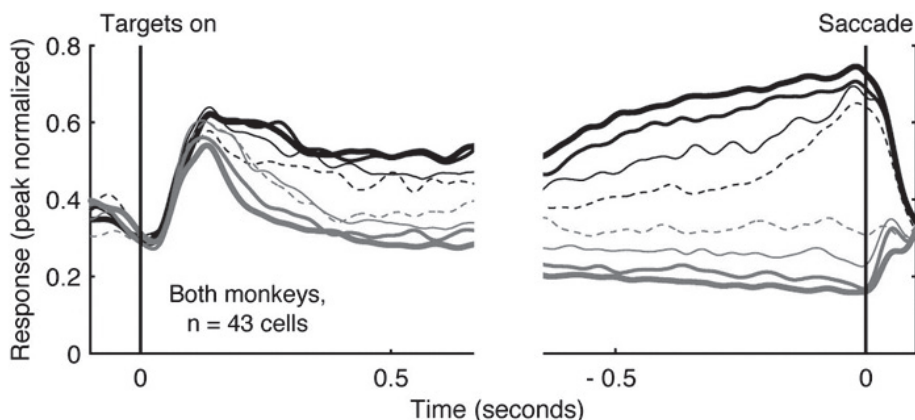


図5. 自由選択課題における受容野内の刺激が選択されたとき（黒）と受容野外の刺激が選択されたとき（灰色）のLIPニューロンの活動（ポピュレーション）. 線の太さは選択された刺激の報酬比率の高さを示す.
(文献9, Figure 5を改変)

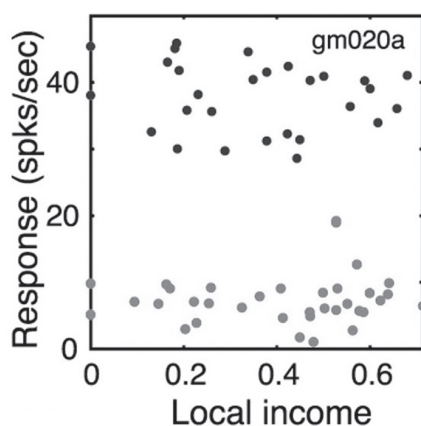


図6. 単独刺激呈示課題における受容野内の刺激（黒）と受容野外の刺激（灰色）にたいするLIPニューロンの応答. 横軸は報酬量（比率）
(文献9, Figure 3Cを改変)

弱い. さらに彼らの実験で興味深いのは, 同じLIPニューロンの活動でも, 標的刺激が1つしか呈示されず選択の余地がない場合には報酬量が変わってもニューロン応答にほとんど変化がないことである (図6).

LIPニューロンの活動は, 報酬がどのような確率で与えられるかについて, 短期的な経験をもとに類推する機能と関係しているようである. この

機能は選択肢がある場合には働くが, 選択の余地がない場合には働かない. さらに, そのときの報酬の質や量もその行動決定要因としてLIPニューロンの活動に影響を与える.

< Barraclough, Conroy & Lee, 2004 >

我々のまわりの環境が我々にとってどのような意味を持っているのか. これを知るためには環境内の刺激にたいする反応とその結果の報酬に関する経験を蓄積し, 刺激—反応—報酬の関係を把握する必要がある. このとき, 環境内の刺激は常に静的なものとは限らない. それが我々にとって利害が競合する敵の行動による場合もある. 敵の行動を知り, それを行動決定要因の一つとしなければならない 때가往々にしてある. この場合, 環境内の効用は動的であり, 確率的にしかとらえることができない. Rochester大学のDaeyeol Leeたちのグループは, コンピュータを競合する対戦相手とし, サルにゲームを行わせた [10].

課題は再び眼球運動を使った自由選択課題である. サルは呈示される2つの標的刺激のうちの1つに眼球運動を行なうが, 報酬が得られるかどうかは対戦相手であるコンピュータの選択に依存する. コンピュータも試行ごとにどちらの刺激を選ぶか決める (もちろん, その結果はサルにはわから

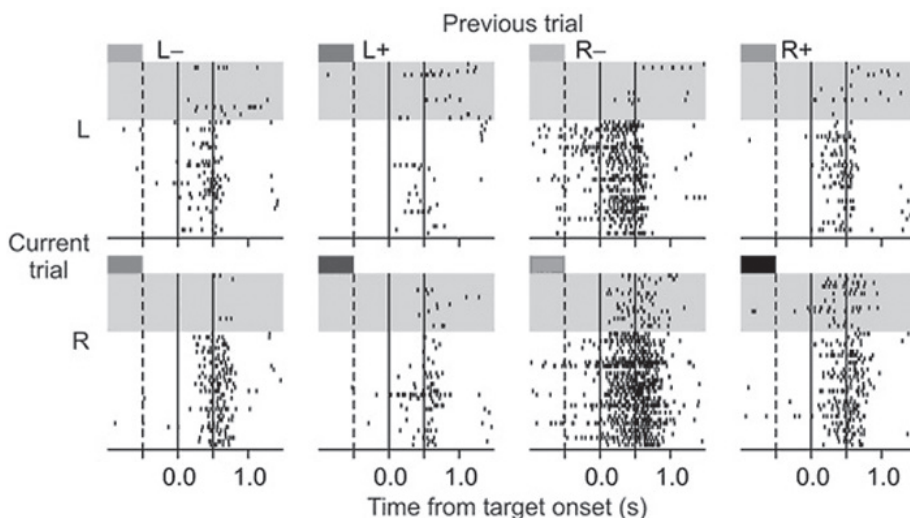


図7. 前の試行での反応とその結果である報酬の有無をコードする前頭連合野ニューロン。
 上段；現試行で左刺激が選択された場合のラスタグラム
 下段；現試行で右刺激が選択された場合のラスタグラム
 列は前試行でそれぞれ、左試行が選択され報酬が得られなかった場合 (L-), 左試行が
 選択され報酬が得られた場合 (L+), 右試行が選択され報酬が得られなかった場合 (R-),
 右試行が選択され報酬が得られた場合 (R+) のラスタグラム。
 (文献10, Figure 6を改変)

ない)。サルが選択した標的刺激がコンピュータが
 選択した刺激と同じであれば、サルの勝ちで報酬
 がもらえる。サルとコンピュータの選択が異なれ
 ば、コンピュータの勝ちでサルは報酬がもらえな
 い。実験者らはコンピュータの選択基準を3段階
 もうけた。アルゴリズム0では、左右の標的刺激
 をランダムに選んだ。アルゴリズム1ではサルの
 過去の反応履歴を考慮し、その傾向と正反対の選
 択を行なった（サルが80%の確率で左を選ぶ場
 合は、コンピュータは80%の確率で右を選んだ）。
 アルゴリズム2では、サルがどちらに反応したか
 だけではなく、その結果報酬が得られたかどう
 かまで考慮してサルの報酬獲得率が最小になる
 ように標的刺激を選択した。アルゴリズム0では、
 サルがどのように行動しようと報酬を得る確率
 は0.5であり、このときサルは一方の刺激にたい
 して反応の偏りをみせた。アルゴリズム1と2
 では、サルの最適行動は前に行なった反応とは
 無関係にランダムに同じ確率で左右の標的刺激
 を選択することであり、実際サルはそれに近い
 行動を示した。

アルゴリズム1と2の間では、1では前の試行
 で報酬が得られるとその選択を繰り返す、得ら
 れないと反応を替えるという傾向 (win-stay-lose-shift)
 が2のときより強かった。これらの行動データは、
 サルがコンピュータのとり戦略について理解した
 上で自らの行動を決定していることをうかがわ
 せる。

次に彼らはこの課題（アルゴリズム2）遂行中
 のサル前頭連合野外側部からニューロン活動を
 記録した。約30～40%の前頭連合野ニューロ
 ンが前の試行でどちらの標的刺激を選択したか
 の情報を現在の試行中に保持した活動を示し、
 ほぼ同じ割合のニューロンが現在の試行中に
 前の試行での報酬の有無をコードしていた。さ
 らに、このうちの一部のニューロンは前の試
 行の選択とその結果の報酬の有無を組み合わ
 せた情報をコードしていた（図7）。これらの
 結果は、前頭連合野ニューロンは課題状況にあ
 わせて必要な情報をリクルートし、これらの情
 報をもとに行動の決定を行なっていることを示
 唆している。

5. おわりに

我々生物の生存は、取り巻く環境の中でいかに効率よく食物を摂取し、外敵から逃れるかに依存している。より良く、より効率的に行動するためには、環境の中に潜む効用 (utility) をより良く理解する必要がある。そのために我々は脳・神経系を持っている。神経系の進化・発達はその方向に進んでいるのであろうし、人間社会の中での個人の能力の評価もそれに大きく依存している。一見、自由意志にみえる我々の個々の判断も実は効用をいかに理解しているかと密接な関係にあることは経済学が証明してきた (神の見えざる手 (Adam Smith))。

最近になってシステム神経科学研究者の間でこの問題が急速に意識され始めた。我々の脳は報酬の価値の計算をする。また、環境—行動—報酬の関連についての情報を過去にさかのぼって収集しそれを確率的に行動に結びつけることができる。ここで紹介した「意思決定 (decision-making)」とよばれる一連の研究は、このような機能をより単純化した形で解明していこうとする試みである。この背景には、経済学におけるゲーム理論の展開、心理学における効用と選択行動に関する研究の興隆、神経生理学における脳の報酬系情報処理研究の進展などがある。これら学際的知識の交流は今後さらに盛んになってくるであろうし、高次脳機能理解には多面的なアプローチが不可欠に

なりつつある。

参考文献

1. Robinson DL, Goldberg ME & Stanton GB : Parietal association cortex in the primate : Sensory mechanisms and behavioral modulations. *J Neurophysiol* **41** : 910–932, 1978
2. Colby CL, Duhamel JR & Goldberg ME : Visual, pre-saccadic, and cognitive activation of single neurons in monkey lateral intraparietal area. *J Neurophysiol* **76** : 2841–2852, 1996
3. Gnadt JW & Andersen RA : Memory related motor planning activity in posterior parietal cortex of macaque. *Exp Brain Res* **70** : 216–220, 1988
4. Platt ML & Glimcher PW : Responses of intraparietal neurons to saccadic targets and visual distractors. *J Neurophysiol* **78** : 1574–1589, 1997
5. Herrnstein RJ : Relative and absolute strength of response as a function of frequency of reinforcement. *J Exp Anal Behav* **4** : 267–272, 1961
6. Mazur JE : *Learning and Behavior*. Prentice Hall, New Jersey, 1994
7. Platt ML & Glimcher PW : Neural correlates of decision variables in parietal cortex. *Nature* **400** : 233–238, 1999
8. Glimcher PW : *Decision, uncertainty, and the brain*. The MIT press, Cambridge, 2003
9. Sugrue LP, Corrado GS & Newsome WT : Matching behavior and the representation of value in the parietal cortex. *Science* **304** : 1782–1787, 2004
10. Barraclough DJ, Conroy ML & Lee D : Prefrontal cortex and decision making in a mixed-strategy game. *Nat Neurosci* **7** : 404–410, 2004