

第六回研究会の総評：今回は、専修大学人間科学部心理学科の国里愛彦さんに、「ベイズ推論モデルの恐怖条件づけ現象への適用」というタイトルで御講演を頂いた。国里さんとは山下先生を通じて、5年ほど前から交流を持たせていただいている。国里さんから条件づけの研究やそれに対する計算論的アプローチに関する研究をしっかりと拝聴させていただいたのは、実は初めてである。私の中では、国里さんといえば、うつ病に対する CBT の治療効果機序へのベイズ推論モデルからの解明という印象であったので、今回の御発表がどのような内容でいらっしゃるのかとても楽しみであった。サバティカルでアムステルダム大学に1年間研究をされていた後でもいらっしゃるのと相まって。

第六回研究会は、本来4月に予定していたが、新型コロナウイルスの蔓延もあったために6月に延期し Web 開催となった。幸甚にも、20名を超える方にご参加頂いた。Web 形式だと気軽に参加はできるが、やはり相互交流が少なかったためにやや寂しくはあった。今後も Web 形式は続くと思われるので、より一体感のある研究会が出来るように考えていきたい。

さて、御講演の内容であるが、最初に吐露させていただくと、恥ずかしながら、「条件づけは奥が深いとは知っていたがこれほど奥が深いとは知らなかった。計算論という同じ括りではあるが強化学習やベイズの知識が私には足りなかった。結果、私には難解であった」という感想であった。これだけで終わってしまうと総評にはならないので、以下説明させていただく。私は生物物理学的モデルやニューラルネットワークモデルは20歳代の学生のころに基礎を固めたが、強化学習モデルやベイズ推論モデルに関してはしっかりした勉強をしていなかった。例えば、PRML、「パターン認識と機械学習」(C.M. ビショップ, 丸善書店)を読んではいない。このことは最新の計算論的精神医学の論文を読むうえで致命的なことがあると最近わかってきたので、PRML を読まれていない方は、是非早めにお勧めする。ただ、今回の総評を書く上で、私は PRML を買って読む余裕が無かったので、OJT 的に以下のように勉強した。強化学習の理解のために、マルコフ決定過程・ベルマン方程式・ブートストラップ・モデルベース・モデルフリー・動的計画法・モンテカルロ法・TD 学習・Sarsa・Q 学習をキーワードに勉強していった。その後、条件づけと RW モデルについて調べて行ったら、片平さんの総説(片平, 2016.参考文献参照)を見つけ、条件づけの連合学習からのモデルである RW モデルと強化学習の関係が概説されてあったのを読み、国里さんの御発表の背景がわかった。片平(2016)によると、RW モデルは1972年に提出されたが、この数理モデルにより、(古典的)条件づけの多くの現象が説明され、また、その当時まだ知られていない新しい現象の予測もされることになった。さらに、RW モデルのアイディアは強化学習にも取り入れられ、ドーパミン細胞の活動と TD 学習モデルの予測誤差信号とよく対応することが Schultz により発見された。こうして、連合学習モデルと強化学習モデルは神経科学においても重要なものとなった。

条件づけの奥深さについてだが、有名な現象であり、精神作用や精神疾患の機序のあらゆる側面に条件づけは関わっていると考えられるが、私は本当に表面的なことしか知らなかったと感じた。たとえば、ブロッキング・潜在抑制(latent inhibition)・曝露療法後の再発現象の様々な種類について初めて知った。国里さんは、心理学者として、これらの臨床現象を体感しその機序を考えていく中で、モデルをしっかり吟味していつているのだと感じた。

条件づけ現象の説明に革命をもたらした RW モデルであるが、消去期の学習で CS-US の連合強度がゼロになるために再発現象が説明できない。そこで、Redish のモデルが紹介された。このモデルは state-classification process、すなわち、条件づけ環境の刺激・文脈・報酬・刺激からの経過時間などの観察現象を、状態(state)によってカテゴライズし、状態の変換は、報酬の予測誤差が持続的にマイナスなときに

なされる。状態とはおそらく、おそらく精神の内的モデルであろう。ある状態における行動価値は、強化学習の Q 学習によって更新されていく。このモデルにより、再発現象のうちの ABA デザインでの復元効果が綺麗に説明される。すなわち、ABA デザインでは消去期に報酬が来ないので、そのときの予測誤差はマイナスが持続する。そのために消去期は、文脈が違うということも相まって、獲得期と違う状態と判断される。こうして2つの状態がモデル内で作成される。テスト期では文脈が A に戻るなので、状態は獲得期の状態と判断され、条件反射を出し、結果、再発現象となると。一方、Redish のモデルは、予測誤差がマイナスの場合に新しい状態を作るという前提をおいていたがゆえに、潜在制止の文脈依存性や、ABC デザインと AAB デザインの復元効果の説明ができない。そこで、次に、Redish のモデルの画期的なアイデアである state-classification process を引き継ぎ、彼らのモデルでの限界を改良するモデルとして Gershman の潜在原因モデル(latent cause model)が紹介された。「動物は観察現象から潜在原因を探る」というのがこのモデルの根底となる前提であるが、この潜在原因は Redish のモデルでは状態(state)ということになる。潜在原因モデルでは、潜在原因はベイズ推論を用いて探られる。そして、潜在原因の事前分布は、Redish モデルのときとは違い、Chinese restaurant process というルールが仮説される。このルールは、「ある状態の潜在原因は、基本的にはこれまでに経験したと判断された回数が多い潜在原因が確率的に選ばれやすいが、ある一定の確率で新しい潜在原因も選ばれる」というものである(“Rich get richer.”)。である。動物は、今の状況での観察現象から潜在原因を推測するのだが、潜在原因からどのように観察現象が生じるかは多項分布で確率的に決まっているので、潜在原因の推測はベイズ推論で更新されていく。この方法により、Redish モデルでは再現できなかった潜在制止の文脈依存性や ABC デザインの復元効果が再現された。Gershman(2010)の Figure2 に、ABC デザインの復元効果の再現とそのときの潜在原因の関係を示す綺麗な図が示されている(ABC デザインとの比較のために ABA や ABB デザインの結果も示されている)。まず獲得期 A と消去期 B では、違う潜在原因 a と b が生じることが示された。そして、ABA デザインではテスト期に潜在原因が a となる確率が多く、そのために条件反射が出やすい(復元現象)こと、ABB デザインではテスト期に潜在原因が b となる確率が多く、そのために条件反射は出にくい(復元現象が生じない)ことが示された。ABC デザインでは、いくつかのモデル動態が考えられるが、この論文でのシミュレーションの結果に関しては次のように解釈された。テスト期 C が A や B と変わらない状況と判断され、新たな潜在原因は作成されず、潜在原因はあるときは a、あるときは b と判断され、a と判断されたときに条件反射が生じる。したがって、ABC デザインでは、ABA と ABB の約半分の頻度で条件反射が生じるような復元現象となる。ただ、もしそれまでに A の経験の方が B よりも多かったのなら、C の状況は Chinese restaurant process のルールにより、A と同じ潜在原因 a と判断され、結果、多くの条件反射が起きることもモデルの動態としては起こりうるし、テスト期 C が A や B とは全く違う状況と判断されたら、新たな潜在原因 c が作成され、そのときには潜在原因 c に関連した確率(多項分布)により条件反射が出ることもある。ベイズ推論モデルも導入して、Gershman の潜在原因モデルにより、条件づけ現象の再現や説明が進んだと思ったが、まだまだ道のはあり、このモデルでは、自然回復などの時間経過による効果は検証できないとの国里さんの説明があり、Gershman が自身のモデルを発展させた Latent cause-modulated model(LCM-RW モデル)が紹介された。このモデルは、これまでのベイズ推論を使って潜在原因の事後分布を更新していく潜在原因モデル(構造学習)と、デルタルールを使って、選択された潜在原因のもとでの観察データの尤度を最大化することによって求められる CS-US の連合強度を更新していく連合学習モデルからなる。連合学習モデルのほうは、RW モデル

に似たモデルとなっている。この2つの学習は交互に実行されて、expectation-maximization (EM) algorithm の変異版のように扱われる(E ステップ=構造学習、M ステップ=連合学習)。ちなみに、Gershman(2017)では、この EM アルゴリズムを、「大脳皮質の計算を理解するフレームワークとして Friston は示唆している」と紹介している。LCM-RW モデルでは、自然回復などの時間経過による効果の再現のために、これらの2つの学習モデルという大きな仕掛けに加え、構造学習の際にどの潜在原因が出現するかの事前分布として、power law temporal kernel を導入している。これにより、消去期とテスト期の間隔が長くなると、獲得期と消去期の2つの潜在原因が記憶圧縮され、テスト期に潜在原因として選ばれる確率が等確率になる。その結果、自然回復が生じてしまう。私としては、Redish や Gershman が条件づけ現象に強化学習やベイズというモデルを採用したというのも重要であるが、現象にマッチしたサブ仮説を本当に上手く所々に取り入れているのと感じる。すなわち、Redish なら「予測誤差が持続的にマイナスな状況では、新しい状態が採用される」であり、Gershman なら「power law temporal kernel による記憶圧縮」である。条件づけの実験を両者がその程度してきたかはわからないが、現象そのものをいかに味わうかがやはり重要と考える。

この総評の内容があっているかの自信は、大変申し訳ないがあまり無いが、この1か月ほど国里さんの御発表を追従していく間に、条件づけとそれへの計算論的アプローチの魅力に取りつかれた感がある。もし若いころに出会っていたら、国里さんに弟子入りしてライフワークにしたかもしれない。私のように感じられた参加者がいらっしゃったら、是非国里さんにアプローチをとられてください。(沖村幸 2020年8月31日)

【参考文献】

片平健太郎.「動物の学習・行動の数理モデルにおける定性的評価と定量的評価」. 動物心理学研究, 第66巻第2号, 2016.