

非対称な学習からのリスク態度の反射効果

Reflection Effect in Risk Attitudes from Asymmetric Learning

佐島 玖仁朗[†], 太田 宏之[‡], 宝田 悠[†], 荒毛 将史[‡], 守本 祐司[‡], 石塚 俊晶[‡], 高橋 達二[†]
 Kuniaki Satori, Hiroyuki Ohta, Yu Takarada, Masashi Arake, Yuji Morimoto, Toshiaki
 Ishizuka, Tatsuji Takahashi

[†] 東京電機大学, [‡] 防衛医科大学校

Tokyo Denki University, National Defense Medical College

kuniaki.satori@gmail.com

概要

ヒトは不確実状況下ではどのように意思決定を行い、その意思決定を形作る過程である学習はどのように行われるのだろうか。本研究では、不確実状況下における意思決定モデルであるプロスペクト理論の確率加重関数に焦点を当て、確率加重関数の過大評価、過小評価は、良い結果と悪い結果に対する非対称な学習から起きるリスク態度の反射効果により形成される可能性があることを示す。

キーワード: 意思決定, プロスペクト理論, 確率加重関数, 強化学習, モデルベース分析, マウス, 動物実験

1. 背景と目的

意思決定とは、生物が行動を選択するプロセスのことであり、数学、経済学、心理学、神経科学など多様な分野で研究されてきた。特にリスク状況下でヒトはどのように意思決定を行うべきか、あるいはどのように意思決定を行っているかは古くから多くの関心と呼んできた問題である。

その意思決定に関わる心理学的研究の中で重要な理論として期待効用理論とプロスペクト理論が存在する。期待効用理論とは不確実状況下における意思決定として、ある選択肢の望ましさを選択肢の効用とその選択肢の確率による、期待効用の大きさにより選考が決定されると定義する意思決定理論である [1]。しかし、期待効用理論は理想化された文脈での選択を扱っており、実際のヒトの意思決定が期待効用理論から逸脱することが様々な形で示されている [2, 3]。特にヒトは確率的な事象を直線的に評価しないことが知られており、低確率を過大評価、高確率を過小評価する傾向がある。この傾向はプロスペクト理論で確率加重関数としてモデル化されており真の確率からの歪みを表現している。

Kahneman と Tversky は、期待効用理論からの経

験的な逸脱を説明するために、ヒトの確率に対する主観的評価を表現する以下の確率加重関数を提示した [4, 5]。

$$w(p) = \frac{p^\gamma}{(p^\gamma + (1-p)^\gamma)^{\frac{1}{\gamma}}} \quad (1)$$

この確率加重関数の具体的な形状を図 1 に示した。パラメータ γ は歪みの強さを表し、1 に近いほど直線に近く、0 に近いほど極端に変化する確率加重関数となる。確率加重関数は、低確率で凹、高確率で凸となり、確率の変化は端の 0, 1 付近に近づくほど劇的に現れている。これにより逆 S 字曲線となり低確率を過大評価、高確率を過小評価するヒトの主観的な確率の歪みを表現している。

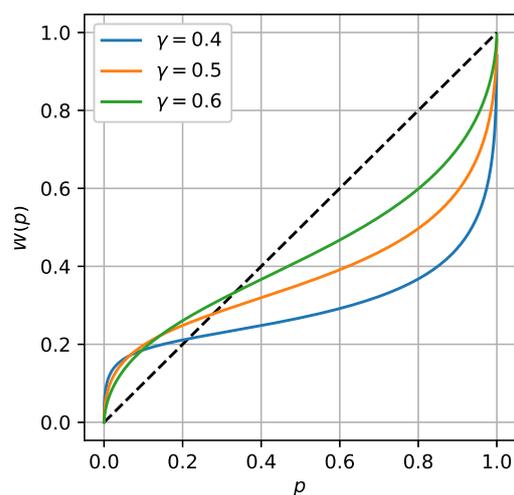


図 1 Kahneman & Tversky (1992) の確率加重関数

これらの理論的枠組みは、意思決定研究に不可欠である。しかし、プロスペクト理論において、その意思決定が形成されるまでの経験的な学習の果たす役割が考慮されていない。ヒトはどのようにして学習を行

い、意思決定をするのだろうか。

Cazé によるとヒトの学習は、良い結果、悪い結果に対して非対称な正と負の学習率を用いて学習しており、特に Cazé の行った 2 本腕バンディットタスクの理論的解析によると、低確率環境で正の学習率に対して負の学習率を低め、高確率環境で正の学習率に対して負の学習率を高めるといふ、環境の報酬確率の分布に応じて正と負の学習率のバランスを調整するという仮説を立てている [7]。この非対称な学習を環境の確率に応じて調節するという仮説に従うと、低確率環境で過大評価、高確率環境で過小評価、という環境の確率に応じたリスク態度の反射が起きる。

我々は、低確率環境で過大評価、高確率環境で過小評価という性質はプロスペクト理論の確率加重関数と同じ傾向を示しており、主観的確率は非対称な学習の調整によって形成されるのではないかと仮説を立てた。

しかし、この Cazé の仮説はヒトや動物の実験では確認されていない。Gershman はヒトを対象とした 2 本腕バンディットタスクの実験を行い、その行動データから強化学習を用いたモデルベース分析を行い、正と負の学習率のパラメータ推定を行った [8]。しかし、負の学習率の方が正の学習率よりも常に高いという結果となり、このヒトの研究はこの仮説を支持するものではなかった。

本研究ではプロスペクト理論を学習という側面から解析することによって、プロスペクト理論の確率加重関数は学習によって形成される可能性があることを示す。そのために、動物であるマウスを用いたモデルベース分析実験から Cazé の仮説を検証する。その後、非対称な学習の調整によって生じる過大評価、過小評価がプロスペクト理論の確率加重関数と同じ傾向であることを示し、学習と確率加重関数の関係を考察する。

2. 強化学習

ヒトや動物は、良い経験と悪い経験の両方から学ぶ。成功した行動は強化され、失敗した行動は抑制されることでより適切な判断ができるようになる [9]。

代表的な強化学習モデルの一つである Q 学習では、期待された報酬と得られた報酬の差分である報酬予測誤差を用いて行動の価値 Q を更新する。期待された報酬よりも多くの報酬を受け取ったときに正の報酬予測誤差が発生し、報酬に起因する行動が強化される。逆に、期待された報酬よりも少ない報酬を受け取ったときに負の報酬予測誤差が発生し、起因する行動選択の減少をもたらす [10, 11]。

先行研究によると、中脳のドーパミン神経細胞は、

正の報酬予測誤差と負の報酬予測誤差と似た発火パターンを示すことが明らかにされている [12, 13]。また、ドーパミン神経細胞は行動価値を出力すると推定される大脳基底核の線条体に投射を行う [10, 11, 14]。大脳基底核は、行動の促進と切り替えのための直接的な経路と間接的な経路から構成されており [15, 16]、Frank らは複数の研究を通じて、パーキンソン病患者におけるドーパミン受容体伝達経路に関する遺伝子の多型やドーパミン補充薬の投与によって、ヒトの学習率が異なることを報告している [17, 18, 19]。その学習率が異なることから Differential Learning Rate (DLR) という学習モデルが提唱されている。行動 a_t が選択され、報酬 r_t が与えられた各試行 t ごとに、 i 番目の穴の行動価値 Q を以下の式 (2) ように更新される。

$$Q_{t+1}(i) = \begin{cases} Q_t(a_t) + \alpha^+(r_t - Q_t(a_t)) & (a_t = i, r_t = 1) \\ Q_t(a_t) + \alpha^-(r_t - Q_t(a_t)) & (a_t = i, r_t = 0) \end{cases} \quad (2)$$

パラメータ α^+ と α^- は、それぞれ選択された行動に対する正と負の学習率を表し、行動選択によって報酬が得られた際には α^+ で更新され、報酬が得られなかった際には α^- で更新される。本論文では非対称な学習率 ($\alpha^+ \neq \alpha^-$) を持つモデルを DLR-Q モデルと呼ぶ。また、対象な学習率 ($\alpha^+ = \alpha^-$) を持つモデルは標準の Q 学習と一致し、標準 Q モデルと呼ぶ。

生物学的考察に基づき提案された DLR に対して Cazé は、2 本腕バンディット問題を用いて DLR の非対称な学習率に関する理論的解析を行った。その結果、正と負の学習率の比率 $x (= \alpha^+/\alpha^-)$ は、環境の報酬確率の分布によって適切に設定されると選択間の弁別性を上げ、獲得報酬の増加につながることを示した。特に、2 選択において選択間の価値の差分を最大化するような正と負の学習率の比率 x^* が以下の式 (3) から求められ、この式から生物は環境の報酬確率に対して低報酬環境で x を高め、高報酬環境で x を低めると仮説を立てている。

$$x^* = \frac{\sqrt{(1-p_0)(1-p_1)}}{\sqrt{p_0 p_1}} \quad (3)$$

3. 環境に応じた非対称な学習率の確認実験

Cazé 仮説を検証するために、Gershman はヒトを対象とした 2 本腕バンディットタスクの実験を行い、その行動データから強化学習を用いたモデルベース分析により DLR のパラメータ推定を行った。しかし、負の学習率の方が正の学習率よりも常に高いという結果となり、このヒトの研究はこの仮説を支持するものではなかった。

この Gershman の実験ではタスクの選択肢の数を 2 選択に制限するような実験条件が用いられており、この検証環境には問題があると考えられる。動物は前回の試行で報酬が得られた場合、その選択肢を続けて選択する一方、報酬が得られなかった場合、同じ選択肢を選ぶことを止め他の選択肢を選ぶ Win-Stay Lose-Shift (WSLS) と呼ばれる傾向があるため [20, 21, 22], 2 選択の場合は一方の選択肢から報酬が得られなかったことによる探索範囲はもう一方の選択肢に限定される。そのため、WSLS の報酬が得られなかったことによる切り替えと強化学習による選択肢の価値に基づいた探索を独立して調査することは困難である。さらに、動物には、不確定な 2 つの選択肢が提示されたときに交互に選択する交代性の応答を示すことが知られており [23, 24], タスクの初期段階の学習時に行動選択確率の変動がマスクされる可能性がある。また、行動選択主体にヒトが用いられたが、意思決定にタスク構造の把握などの高次元の推論などが含まれる可能性があり、強化学習を用いた行動データ解析では推論をモデル化できず解析は困難である。

そこで本研究では、5 本腕バンディットタスクを用いて、マウスの行動選択の変化を強化学習を用いたモデルベース分析実験を行った。2 選択の実験により正確な解析ができていないという問題に対して、実験の選択肢の数を 5 選択に増やすことで選択肢間の干渉を減らし、正確な解析を可能とする。

3.1 設定

本研究では Gershman の実験と同様に強化学習を用いたモデルベース分析によりマウスの行動解析を行い、Cazé 仮説を検証する。モデルベース分析とは、データからは直接観測できない内的な計算過程を推定する分析手法であり、先行研究ではヒトや動物の各選択に対応した行動価値が割り当てられると仮定して行動データの解析に使用されている [25, 26]。行動価値の学習は、DLR-Q モデルと標準 Q モデルに基づいて行われ、以下の softmax 方策に従うことで試行 t での選択確率の予測を与えた。

$$P(a_t) = \frac{\exp(Q_t(a_t)\beta)}{\sum_{k=0}^n \exp(Q_t(k)\beta)} \quad (4)$$

ここで n は穴の数、 β は逆温度パラメータであり、逆温度パラメータは行動価値と行動選択の関係を決定する、標準 Q モデルは α, β の 2 つ、DLR-Q モデルは $\alpha^+, \alpha^-, \beta$ の 3 つの自由パラメータを持つ。パラメータの最適化は python の scikit-learn[27] のベイズ

最適化を用いて、各マウスの実際の行動データと強化学習モデルから計算された行動選択確率の対数尤度を最大化した。

環境の報酬分布によって正と負の学習率のバランスが変化するかを確かめるために低報酬環境と高報酬環境を用意した。実験は、以下の 2 つの報酬確率条件で行った。

表 1 低報酬環境でのサブタスク設定 (%)

	1	2	3	4	5
ALL	30	30	30	30	30
BIT	0	0	50	0	0
REV	30	30	0	30	30

表 2 高報酬環境でのサブタスク設定 (%)

	1	2	3	4	5
ALL	60	60	60	60	60
BIT	0	0	80	0	0
REV	60	60	0	60	60

実験は ALL, BIT, REV の 3 つのサブタスクで構成されており、それぞれのサブタスクは、決まった回数報酬を得ると次のサブタスクに移行する。学習率の推定を行うためにサブタスクの環境を遷移させることで再学習を行わせている。

3.2 結果と考察

低報酬環境、高報酬環境それぞれで 10 匹のマウスを用いて実験を行い、標準 Q モデルと DLR-Q モデルの強化学習を用いたモデルベース分析の結果を平均した。低報酬環境と高報酬環境における標準 Q モデルの赤池情報基準 (AIC) の全体平均は 3095.4 であり、DLR-Q モデルの AIC は 3023.5 であった。DLR-Q モデルにおける正負の学習率は、低報酬環境では $\alpha^+ = 0.0332(\pm 0.0203(SD))$, $\alpha^- = 0.0043(\pm 0.0019(SD))$, 高報酬環境では $\alpha^+ = 0.0188(\pm 0.0109(SD))$, $\alpha^- = 0.0131(\pm 0.0112(SD))$ と推定された。推定された α^+, α^- から計算される x は、低報酬環境 ($\pm 9.99(SD)$): 8.08, $n=10$) と高報酬環境 ($\pm 1.97(SD)$): 1.23, $n=10$; Welch の t-test, $p=0.0119$) では有意差があったが、 β は低報酬環境 ($\pm 6.06(SD)$): 1.80) と高報酬環境 ($\pm 5.85(SD)$): 3.65; $p=0.0112$) では有意差は見られなかった。以上のことから、低報酬環

境では x を高く、高報酬環境ではほぼ同じ値にすることで、行動データの再現性が向上した。

このモデルベース分析実験の結果である低報酬環境で x を高め、高報酬環境で x を低めるという傾向は Cazé の仮説と一致している。この結果より、動物であるマウスは学習に非対称性を持ち、環境の報酬確率に応じてその非対称な学習率のバランスを調整し獲得報酬を高める学習を行っていることが推察される。

4. 学習と確率加重関数の考察

マウスが非対称的な学習率のバランスを環境の報酬確率によって適応的に変化させていることが実験で示された。この非対称な学習を報酬確率によって適用的に変化させることによりプロスペクト理論の確率加重関数と同様の傾向が形成されることを示す。

4.1 非対称な学習による主観的確率

真の報酬確率を p とし無限回の試行を行うとすると、以下の式のようになる。

$$Q_\infty = Q_\infty + p\alpha^+(1 - Q_\infty) + (1 - p)\alpha^-(-Q_\infty) \quad (5)$$

この式を正と負の学習率の比である $x = \alpha^+/\alpha^-$ を用いて変換すると、以下の式が得られる。

$$Q_\infty = \frac{px}{px - p + 1} \quad (6)$$

図2は、 x を $1/5, 1, 5$ とした場合の p と Q_∞ の関係を示している。 α^- が α^+ に対して小さい ($x = 5$) とき Q_∞ は過大に見積もられ、 α^- が α^+ に対して大きい ($x = 1/5$) とき Q_∞ は過小に見積もられている。

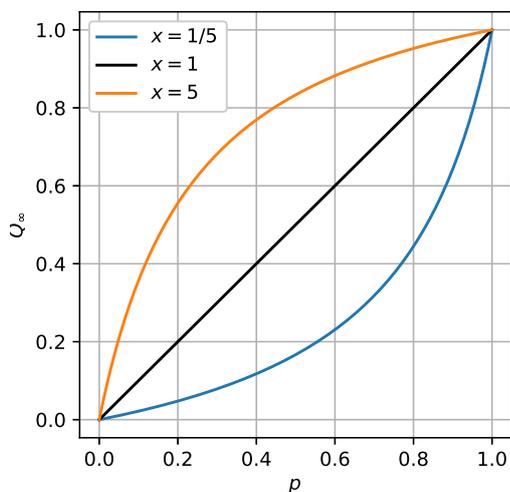


図2 x に対応する Q_∞

次に真の報酬確率 p_0 と p_1 の2択課題における最適な正と負の学習率の比率 x^* について、真の報酬確率と Q 値の関係を分析した。 x^* は2選択であるとき報酬確率 p_0, p_1 により式(3)で求められ、図3に2選択における最適な正と負の学習率の比率 x^* を示す。また、図4に x^* を元にした p_1 と Q_1 の関係を示す。

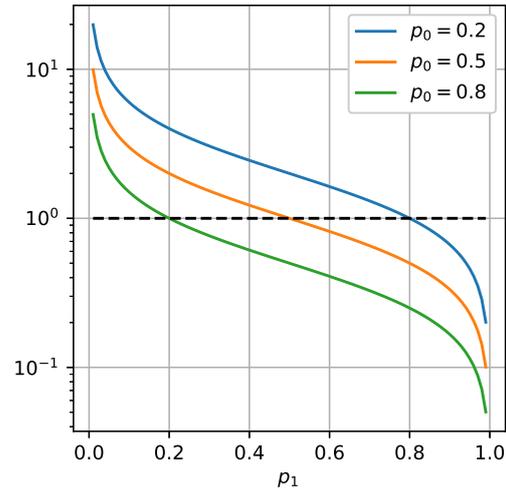


図3 2選択における最適な正と負の学習率の比率 x^*

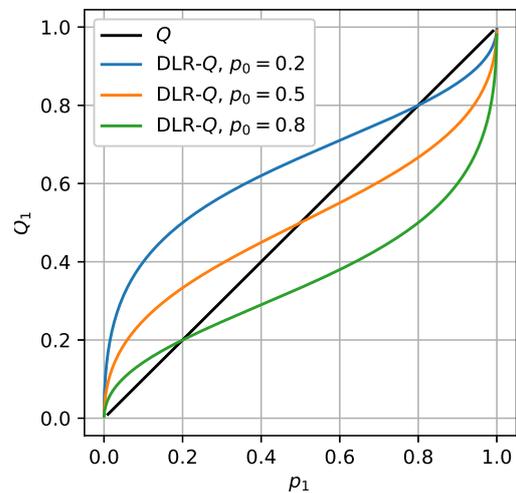


図4 報酬確率が低いときに過大評価し、高いときに過小評価する逆S字曲線

標準 Q モデル ($x = 1$) では、試行回数が無限と仮定した場合、行動価値 Q は真の報酬確率と同じになる。一方、学習率比 x を最適化すると、逆S字型の関係が導出される。なぜこのような逆S字型の関係になるかというと、環境の報酬確率の分布が低い時 x^* は高く

なり過大評価が起きる、環境の報酬確率の分布が高い時 x^* は低くなり過小評価が起きるといふ、環境の報酬確率の分布に応じたリスク態度の反射を行っているからである。

この逆 S 字はプロスペクト理論の確率加重関数と、主観的に小さな確率で起こる出来事を過大評価し、大きな確率で起こる出来事を過小評価するという点で一致している。このことから、ヒトの小さな確率を過大評価、大きな確率を過小評価する歪んだ主観的な確率は DLR による非対称な学習から形成されたと考察できる。

また、ヒトはプロスペクト理論の確率加重関数で説明されるように不確実性を伴う意思決定をするリスクを正確に把握することができず、誤った意思決定を促してしまうため非合理的に感じるが、2 選択の学習という側面から見てみると選択の弁別性を高めるために生じる合理的な主観的な評価と言える。

4.2 サンプルに基づく決定

本研究では 2 選択バンディットタスクで考えたが、実際の確率の過大・過小評価の起こる場面を想定すると、ある選択をするかしないかの選択をする場合が存在する。例えば、宝くじを買うか買わないかの選択が上げられる。この場合は宝くじを買って当たりが出る時だけの確率だけを経験的に知っており、宝くじを買わなかった際の当たりはない。そのためこの場合は 2 本腕で想定すると片方の確率だけが存在する状況となる。この場合は x^* は計算できず、DLR による主観的な評価が発生しないと考えられるが、Stewart らの確率加重関数の起源に対するアプローチの考えであるサンプルに基づく決定 [28] を元にとると説明がつく。サンプルに基づく決定とは、ヒトのある刺激に対する主観的な評価は、刺激間の順序の比較と、記憶頻度の累積から構成されるとする。つまり、評価対象となる刺激が評価者の記憶の中で関連する刺激と比較して何回より大きいと判断されたかによって、その刺激の主観的な評価が決まる。

この比較のアイデアを参考にすると、ヒトはある刺激に対する複数の評価の仮説のようなものを持ち、その中で比較を行っている。そのため、宝くじの例のような 1 選択であろうとも、複数選択のバンディットが発生し、過大評価・過小評価された主観的な評価が構成されていると考えることができる。

4.3 魅力度

図 4 の結果は他方 p_0 の選択枝の報酬確率の推定に応じて、過大評価と過小評価の反転点がずれた。この p_1 の反転点は、他方 p_0 の確率が高い際は下へスライドし過小評価の傾向が強まり、他方 p_0 の確率が低い際は上へスライドし過大評価の傾向が強まる。Gonzalez らにおいて提案されている確率加重関数の式 (7) では魅力度 δ というパラメータによりこの反転点が制御されている [29]。

$$w(p) = \frac{\delta p^\gamma}{\delta p^\gamma + (1-p)^\gamma} \quad (7)$$

この魅力度は個人ごとによっての設問への魅力の高さを表しており、高い時は反転点が上へスライド、低い時は下へスライドしている。

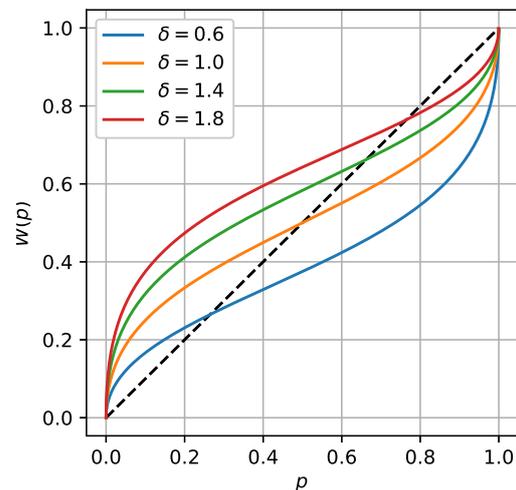


図 5 Gonzalez らの提案した確率加重関数における δ の変化

非対称な学習によって形成される逆 S 字曲線も魅力度と同様の傾向を持つ。図 4 において他方 p_0 の確率が高い時、 p_1 の魅力度は低く、他方 p_0 の確率が高い時、 p_1 の魅力度が高い。非対称な学習によって形成される逆 S 字曲線の交点のスライドは魅力度パラメータと同様の傾向を示し、このことからヒトの選択への魅力度とは他の選択との比較により変動するのではないかと推察できる。プロスペクト理論研究の心理的表現を非対称な学習によって説明ができる可能性があることを示した。

5. まとめ

本研究では、不確実性下における意思決定モデルであるプロスペクト理論では一般的に扱われていない学習に役割に焦点を当てた。マウスの実験から動物が環境の報酬確率の分布から非対称な学習率のバランスを調整するというリスク態度の反射効果を行っていることを示した。また、リスク態度の反射効果からプロスペクト理論の確率加重関数と同様の逆S字曲線が形成されることを示した。

今後、非対称な学習からの逆S字曲線は魅力度のような説明づけられる可能性があると考えているため解析を進める。また、本研究では動物であるマウスに対しての非対称な学習を行っているかの検証実験を行ったが、ヒトでは確認していない。そのため今後はヒトも非対称な学習を行っているかの検証実験が必要であると考えるため今後実験を行う予定である。

文献

- [1] Von Neumann, J., & Morgenstern, O. (1944). "Theory of games and economic behavior", Princeton University Press.
- [2] Allais, M. (1953). "Le Comportement de l'Homme Rationnel devant le Risque, Critique des Postulats et Axiomes de l'Ecole Americaine", *Econometrica*, 21 503-546.
- [3] Ellsberg, D. (1961). "Risk, Ambiguity, and the Savage Axioms", *Quarterly Journal of Economics*, 75, 643-669.
- [4] Kahneman, D., & Tversky, A. (1979). "Prospect Theory: An Analysis of Decision Under Risk", *Econometrica*, 47, 263-291.
- [5] Tversky, A., & Kahneman, D. (1992). "Advances in prospect theory: Cumulative representations of uncertainty", *Journal of Risk and Uncertainty*, 5, 297-323.
- [6] Hartinger, A. (1998). "Do subjective weighting functions of generalized expected value theories capture persisting properties of individual decision makers", *Universitat Regensburg, Germany, working paper*.
- [7] Cazé, R. D., & Van Der Meer, M. A. (2013). "Adaptive Properties Of Differential Learning Rates For Positive And Negative Outcomes", *Meer, Biological Cybernetics* 107, 711.
- [8] Gershman, S. J. (2015). "Do learning rates adapt to the distribution of rewards?" *Psychonomic Bulletin and Review*, 22, 1320-1327.
- [9] Thorndike, E. L. (1911). "The animal behavior series. Animal intelligence: Experimental studies" *Macmillan Press*.
- [10] R. S. Sutton, A. G. Barto (2018). "Reinforcement Learning: An Introduction (MIT press, Cambridge, MA, USA), second edn".
- [11] Neftci, E. O., & Averbeck, B. B. (2019). "Reinforcement learning in artificial and biological systems", *Nature Machine Intelligence*, 1, 133-143.
- [12] Schultz, W. (2015). "Neuronal Reward and Decision Signals: From Theories to Data", *Physiological reviews*, 95(3), 853-951.
- [13] Dabney, W. et al., (2020). "A distributional code for value in dopamine-based reinforcement learning", *Nature* 577, 671.
- [14] Samejima K, Ueda Y, Doya K, & Kimura M. (2005). "Representation of action-specific reward values in the striatum", *Science* 310: 1337-1340.
- [15] Ueda, Y. et al., (2017). "Distinct Functions of the Primate Putamen Direct and Indirect Pathways in Adaptive Outcome-Based Action Selection", *Frontiers in Neuroanatomy*, 11, 0.
- [16] Nonomura, S. et al., (2018). "Monitoring and Updating of Action Selection for Goal-Directed Behavior through the Striatal Direct and Indirect Pathways", *Neuron*, 99, 1302.
- [17] Frank, M. J. & Seeberger, L. C. & O'Reilly, R. C. (2004). "By carrot or by stick: cognitive reinforcement learning in parkinsonism", *Science*, 306, 1940.
- [18] Frank, M. J., Moustafa, A. A., Haughey, H. M. & Curran, T. Hutchison, K. E. 2007. "Genetic triple dissociation reveals multiple roles for dopamine in reinforcement learning", *PNAS*, 104, 16311.
- [19] Frank, M. J., Doll, B. B., Oas-Terpstra, J. & Moreno, F. (2019). "Prefrontal and striatal dopaminergic genes predict individual differences in exploration and exploitation", *Nat Neurosci* 12: 1062-1068
- [20] Lee, D., Conroy, M. L., McGreevy, B. P., & Barraclough, D. J. (2004). "Reinforcement learning and decision making in monkeys during a competitive game.", *Cognitive Brain Research*, 22, 45.
- [21] Skelin, I. et al., (2014). "The Memory Trace Supporting Lose-Shift Responding Decays Rapidly after Reward Omission and Is Distinct from Other Learning Mechanisms in Rats.", *European Journal of Neuroscience*, 39, 1655.
- [22] Amodeo, D. A., Jones, J. H., Sweeney, J. A. & Ragozzino, M. E. (2012). "mice on probabilistic reversal learning and stereotyped behaviors", *Behavioural Brain Research*, 227, 64.
- [23] Vecera, S. P., Rothbart, M. K. & Posner, M. I. (1991). "Development of spontaneous alternation in infancy", *Journal of Cognitive Neuroscience*, Vol. 3, No. 4, pp. 351-354.
- [24] Kraeuter, A. K., Guest, P. C. & Sarnyai, Z. (2019). "The Y-Maze for Assessment of Spatial Working and Reference Memory in Mice", *Methods in Molecular Biology*, Vol. 1916, pp. 105-111.
- [25] Ito, M. & Doya, K. (2009) "Validation of Decision-Making Models and Analysis of Decision Variables in the Rat Basal Ganglia", *The Journal of neuroscience : the official journal of the Society for Neuroscience*, 29, 9861.
- [26] Cinotti, F. et al. (2019). "Dopamine blockade impairs the exploration-exploitation trade-off in rats", *Scientific Reports* 9, 1-14.
- [27] Scikit-learn: Machine Learning in Python, Pedregosa et al., *JMLR* 12, pp. 2825-2830, 2011.
- [28] Stewart, N., Chater, N. & Brown, G. D., (2006). "Decision by sampling". *Cognitive Psychology*, 53, 1-26.
- [29] Wu, G., & Gonzalez, R. (1996). "Curvature of the probability weighting function", *Management Science*, 42, 1676-1690.