

ロボット実験による Vicarious trial-and-error(VTE) の役割の解析

松田英子*¹ Julien Hubert*¹ 池上高志*¹
Eiko Matsuda Takashi Ikegami

*¹東京大学 総合文化研究科

Vicarious trial-and-error (VTE) is a type of conflict-like behavior, observed in route selection tasks [Tolman, 1939, Muenzinger and Flecher]. Studies of VTEs have shown a correlation between the number of VTEs exhibited by a system with its learning efficiency. At the onset of learning a task, the number of VTEs increases, and when the learning reaches its plateau, it decreases. From experiments of rats, VTEs have been reported in T-maze experiments of rats who were shown to be simulating their next decisions internally before acting [Johnson and Redish, 2007]. The question we explore in this paper concerns the role of VTE. Basing ourselves on a model developed by [Bovet and Pfeifer, 2005], we created a computer simulation and ran robotic experiments to compute the number of VTEs during the learning of a T-maze task.

1. 序論

[Tolman, 1939, Muenzinger, 1938] において、ラットが選択肢を選ぶ際、頭を左右に振って迷っているように見えることが観察された。この迷っているような振る舞いは「仮の試行錯誤 (vicarious trial-and-error; VTE)」と呼ばれる。VTE の回数は学習の度合いによって変化し、学習初期は多くの VTE が見られ、その後減ることがわかっている。

[Johnson and Redish, 2007] による海馬の場所細胞の活動を記録した実験から、VTE が観察される時、ラットの位置から数歩先の場所細胞が活動することがわかっている。これより、VTE を示している時に、ラットは頭の中で選択の結果を想像し、選択肢を吟味していると考えられている。Ikegami は頭の中で行動の結果をシミュレーションをすることは意識の原型となると論じており、VTE は意識の理解に重要な役割を果たすと述べている [Ikegami, 2007]。このように VTE は、意思決定や頭の中で思考することと深い関わりがあるとされているが、その役割は十分にわかっていない。本研究では、VTE の役割について理解するために、コンピュータシミュレーション上にロボットのモデルを作り、実験を行った。

2. 方法

2.1 実験環境

本研究の実験環境は [Bovet and Pfeifer, 2005] によって開発されたロボット実験に基づいたものである。ロボットには 4 種類のセンサー (カメラ・ひげセンサー・赤外線センサー・報酬への感受性) とモーターが装備され、これらの関係を学習することにより、図 1 に示す T 字型の迷路を解く。迷路の一方の端には報酬が置かれ、もう一方の端には罰が置かれている。曲がり角には、ひげセンサーで検知することの出来るキューが報酬と同じ側に置かれ、ロボットはキューと報酬の関係を学習することにより、迷路を解くことが出来る。5 回の試行毎に報酬と罰の位置が入れ替わり、ロボットは中央の通路のある決まった初期位置 (図 1 で $(x, y) = (29, 20)$) から試行を 100 回繰り返す。

2.2 ニューラルネットワーク

ロボットの行動はニューラルネットワークによって制御される。ニューラルネットワークは 5 つの感覚野によって構成され

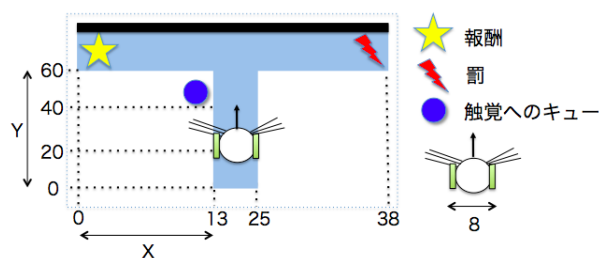


図 1: 実験環境の図。各試行毎にロボットは中央の通路のある決まった初期位置 $(x, y) = (29, 20)$ から試行を始め、100 回試行を繰り返す。曲がり角にはひげセンサーで検知できるキューが報酬と同じ方向に置かれ、ロボットは触覚へのキューと報酬との関係を学習することにより迷路を解く。

(図 2(a)), そのうち 4 つの感覚野はロボットに装備された 4 種類のセンサー入力をそれぞれ受け取り、1 つの感覚野によってモーター出力が決定される。モーター出力は、それぞれの感覚野の相互作用によって計算され、相互作用の強さは Hebb 則によって学習される。感覚野間の関係はあらかじめ決まっておらず、ロボットは環境を動きまわることによって学習を行い、意味のある振る舞いを自己組織的に生成する。

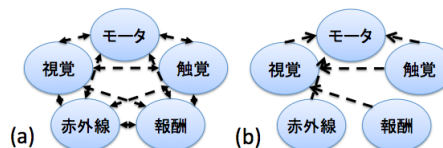


図 2: ロボットのニューラルネットワーク (a) 5 つの感覚野が全結合したモデル。 (b) 5 つの感覚野間に、課題を解くための必要最低限の結合のみを残したモデル。

[Bovet and Pfeifer, 2005] のモデルでは感覚野が全結合していたが、本研究では感覚野のつながり方が VTE に与える影響を見るために、新しいニューラルネットワークのモデルを構成した (図 2(b))。このモデルには、感覚野間に課題を解くための必要最低限の結合のみが残され、その他の結合はない。必要最低限の結合とは、[Bovet and Pfeifer, 2005] において特定の役

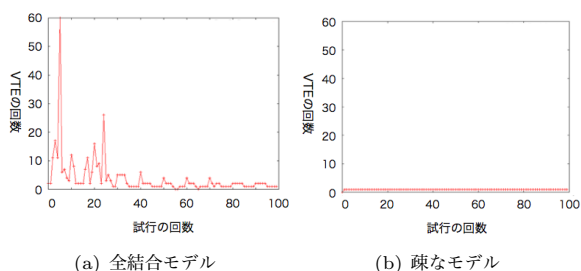


図 3: 100 回の試行における VTE の回数の変化。横軸: 試行の回数 縦軸: 各試行における VTE の回数 詳しくは本文参照。

割があると説明されている結合のことであり、これらの結合のみで課題を解くことが出来る。

2.3 遺伝的アルゴリズム

[Bovet and Pfeifer, 2005] のモデルには多くのパラメータが関わっており、パラメータの違いによって振る舞いが異なる。パラメータを調整し、振る舞いを最適化するために、遺伝的アルゴリズム (GA) の手法を用いた [Holland, 1975]。GA の評価関数は、ロボットが 100 回繰り返して迷路を解く間に報酬に到達する割合によって決めた。

3. 結果

GA によってパラメータを調整した結果、感覚野の結合が全結合のモデルと疎なモデルのどちらも課題を 100% の成功率で解くことが出来た。それぞれのモデルのロボットが課題を解く際、VTE の数を数えた。図 3(a) は全結合モデルのロボットが各試行で示した VTE である。これより、学習の初期において多くの VTE が見られ、学習が進むと VTE の回数が少なくなることが分かる。これは [Tolman, 1939] によるラットの実験において観察された VTE に似ている。一方、疎なモデルの VTE は回数が一定して少ない。どちらのモデルも 100% の成功率を示しているが、VTE に違いがあることから、VTE が直接課題を解く能力に影響しているわけではなく、冗長な感覚野の相互作用によって生成されることがわかる。

次に、ラットに似た VTE の変化を示すロボットは、ノイズに対して頑強であると仮説を立て、初期条件に摂動を加えることによる課題の成功率の変化を見た。GA によるパラメータ調整の過程において、ロボットは固定した初期位置 $(x, y) = (29, 20)$ から試行を始めたが、ここでは GA によって最適化されたロボットに異なる初期位置から迷路を解かせた。図 4 は 2 つのモデルの結果を表している。第一にわかることは、異なる初期位置から課題を始めると成功率が変化することである。次に言えることは、成功率の平均はほぼ同程度だが、全結合モデルでは分散が低い (データは [Matsuda, Hubert, and Ikegami, 2011] を参照のこと)。つまりこの結果は、全結合モデルが初期位置の違いによる影響を受けにくく、環境の変化に対して頑強であることを示唆している。

4. 結論

本研究では「仮の試行錯誤 (VTE)」の役割を明らかにするために、[Bovet and Pfeifer, 2005] に基づいたロボット実験を行った。ニューラルネットワークにおける感覚野のつながり方を 2 通り用意し、学習の過程における VTE の回数の変化を観察した。結果として、感覚野同士が全結合しているモデルでは [Tolman, 1939] によってなされたラットの実験と同様 VTE の

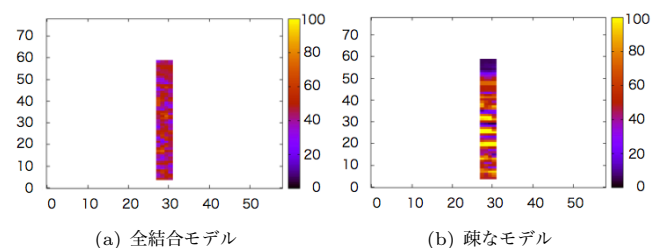


図 4: 各初期位置における課題の成功率。色のついた各点が初期位置を表しており、色はその初期位置から課題を 100 回繰り返した時の成功率。詳しくは本文参照。

回数が学習の初期に多く、その後少なくなることが見られた。一方、疎な結合のモデルでは、VTE の回数は少なく、ほぼ一定であった。両方のモデルが 100% の課題成功率を示すことから、VTE は課題を解く能力に直接影響があるのではなく、感覚野の冗長な結合によって引き起こされることが言える。また、全結合モデルでは、初期条件の変化に対して課題成功率があまり変化しないのに対し、疎なモデルでは大きく変化する。さらに解析を進めた結果、パラメータの違いによっては全結合モデルでも VTE を示さない個体が見つかり、この個体の課題成功率は、疎なモデルと同様初期位置の変化に鋭敏に影響する。これより、ラットと同様の VTE の変化によって、頑強性が獲得されることが示されている。

参考文献

- [Tolman, 1939] Tolman, E. C.: Prediction of vicarious trial and error by means of the schematic sowbug, *Psychol. Rev.*, 46:318 - 336 (1939).
- [Muenzinger, 1938] Muenzinger K.F.: Vicarious trial and error at a point of choice: a general survey of its relation to learning efficiency, *J. Genet. Psychol.*, 53:75 - 86 (1938).
- [Johnson and Redish, 2007] Johnson, A. and Redish, A. D.: Neural ensembles in ca3 transiently encode paths forward of the animal at a decision point, *The Journal of neuroscience : the official journal of the Society for Neuroscience*, 27:12176 - 89 (2007).
- [Bovet and Pfeifer, 2005] Bovet, S. and Pfeifer, R.: Emergence of delayed reward learning from sensorimotor coordination. In *IEEE/RSJ International Conference on Intelligent Robots and Systems(IROS)*, 841 - 846 (2005).
- [Hebb, 1949] Hebb, D. O.: *The Organization of Behavior*. Wiley, New York (1949).
- [Ikegami, 2007] Ikegami, T.: Simulating active perception and mental imagery with embodied chaotic itinerancy, *Journal of Consciousness Studies*, 14:111 - 125 (2007).
- [Holland, 1975] Holland, J.: *Adaptation in Natural and Artificial Systems*, Ann Arbor : The University of Michigan Press (1975).
- [Matsuda, Hubert, and Ikegami, 2011] Matsuda, E., Hubert, J., and Ikegami, T.: A Robotic Approach to Understand the Role of Vicarious Trial-and-Error in a T-Maze Task, *Proc. of ECAL*, 514 - 521 (2011).