The Role of Vicarious Trial-And-Error in a T-Maze Task

Eiko Matsuda, Julien Hubert and Takashi Ikegami
Graduate School of Arts and Sciences
The University of Tokyo
Komaba, Meguro-ku 3-8-1, Japan
Email: {eiko, jhubert, ikeg}@sacral.c.u-tokyo.ac.jp

***Abstract*—Vicarious trial-and-error(VTE) is a type of conflict-like behavior, observed in route selection tasks where rats have been observed evaluating their possibilities before moving toward one route [1]. Studies of VTE 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. The question we explore in this paper concerns the impact of VTE on the learning capability of an agent. Basing ourselves on a model developed by Bovet and Pfeifer (2005), we ran robotic experiments to compute the number of VTEs during the learning of a T-maze task. Our results first show that what has been found in rats can be replicated in artificial systems. In this work, by preventing the presence of the VTEs at the motor level, we discovered that their absence inhibited the correct acquisition of the task despite maintaining an accurate control of the robot. This implies that the small bodily oscillations are used actively by the neural controller to complete the T-Maze task.**

***Keywords*—robot, embodiment, vicarious trial-and-error, Hebbian learning**

I. Introduction

In his experiments, Tolman observed that rats are seemingly hesitating when they must choose between one of two rooms, one of which containing a reward while the other being empty [1]. The only cue differentiating the rooms is the color of their doors. A black door indicates the room provides a reward, and a white color indicates an empty room. To reach the reward, the rats must learn the relationship between the color of the door and the presence of the reward. During the learning phase, the rats have been seen moving their head from one door to another which is referred by Tolman as a conflict-like behavior named "vicarious trial-and-error (VTE)". In his experiments, Tolman noticed that the number of VTEs increases at the onset of the learning phase to start decreasing when the performance reaches its plateau. From that observation, VTE has been connected to learning efficiency.

Johnson and Redish reported the presence of VTEs in experiments on rats who were shown to be simulating their next decisions internally before acting [2]. Tarsitano found that, in a detour task, jumping spiders display two phases of action: the inspection phase, where spiders stop and inspect possible routes toward a target, and the locomotory phase, where spiders move toward a single direction. VTEs have been observed during the inspection phase. Tarsitano concluded that "one can speculate that it is a small but significant jump to use trial and error vicariously when choosing a goal to approach" [3]. From these researches, VTEs seem to have some essential role in internal reflection and decision making. However, the role of the VTEs has yet to be fully investigated.

The question we explore in this paper concerns the importance of VTEs for learning capability. Based on a model developed by Bovet and Pfeifer [4] for T-Maze learning experiments on robotic platforms, we found the presence of VTEs during the acquisition of the task [5]. Our results displayed the same pattern of increase followed by a decrease in the number of VTEs as observed in the rat. In those experiments we did not explore if the VTEs were an epiphenomenon due to the neural computation or if the oscillations observed were exploited by the controller to solve the task. By preventing the VTEs at the interface between the motors and the controller, we could show that the oscillations are necessary for the acquisition of the task.

II. Methodology

A. Experimental Setup

The environment is a T-Maze with one central arm and two side ones (see figure 1). A reward is located at the end of one arm, and a punishment is placed on the opposite one. The robot learns to reach the reward following a tactile cue placed at the end of the central arm, on the same side as the reward. The robot is modeled following the e-puck robot [6] and is equipped with 5 sensors (whiskers, omnidirectional camera, proximity sensors and reward sensitivity) and motors. The control of the robot is handled by two parameters: the forward velocity and the angular velocity. The former is set to a fixed value while the later is given by the controller of the robot described in the next section.

B. Neural Network

A neural network serves as the controller of the robot. It is composed of 5 interconnected groups of neurons(see figure 2): 4 groups receiving sensory information from the modalities of the robot and one group computing the motor commands necessary to solve the task. Each group is composed of sub-groups of neurons with different functionalities. Hebbian learning is applied on all connections in the network to enable it to learn the task. For additional details on the model, please refer to the original paper by Bovet and Pfeifer [4].

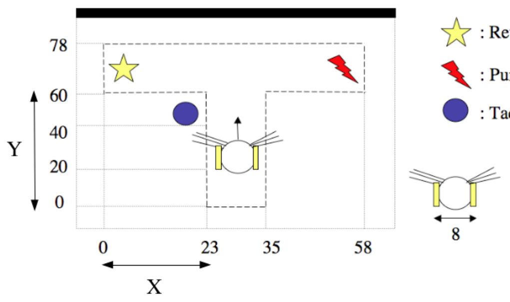

Fig. 1. T-Maze environment used for the experiment. At the beginning of each trial, the robot is placed on the central arm of the maze. The circle at the choice point represents the tactile cue, the star at one end of the maze indicates reward, and the lightning at the other end of the maze stands for punishment. The back wall is painted black and the other walls are white, which are detected by the robot's omnidirectional camera. Walls of the T-Maze are perceived by the robot's proximity sensors.

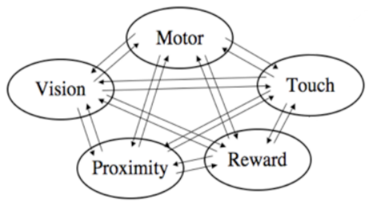

Fig. 2. Neural Model from Bovet and Pfeifer [4]. 5 populations of neurons process each modality of a robot in a T-Maze experiment.

C. Setup of the Genetic Algorithm

Bovet and Pfeifer's model relies on many parameters. Despite the authors not mentioning how to select those parameters, we found out that slight differences in their value can strongly influence the performance of the robot. To tune these parameters and optimize the performance of the controller, we employ a genetic algorithm (GA) with tournament selection, one point crossover and mutation [7]. The fitness function of the GA is the average success of the robot in reaching the goal for 100 repetitions started from a single initial position.

III. Results

We evolved five runs of GA. Two out of the five GA runs got the maximum fitness value (100% success). We selected the best individual from these evolved runs and counted the number of VTEs it displayed. Our methodology to count the number of VTE in a robot is similar to the one used by Tolman. One VTE is granted if, between two timesteps, its angular velocity changes its sign. In order to filter small oscillations around an angular velocity of 0, a VTE is only granted if the sign change is outside the range $[-0.3; 0.3]$.

In our previous study, we concluded that the pattern of VTEs observed during the acquisition of the T-Maze task was similar to the one observed in rats [1], [8]. That is the number of VTEs was increasing at the onset of the learning to decrease after reaching a plateau in its performance. This result is reproduced in figure 3.

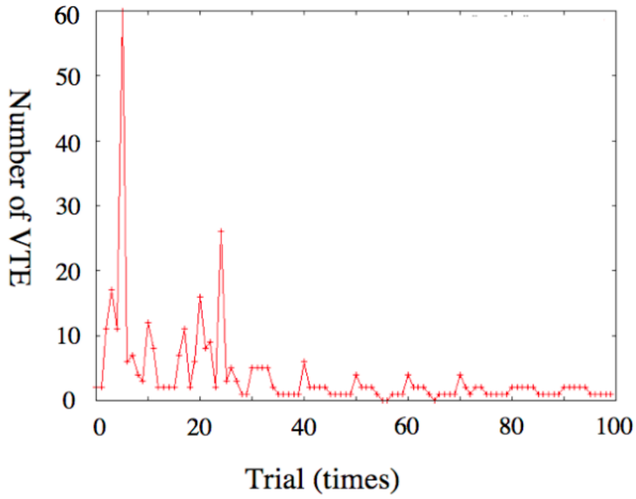

Fig. 3. Change in the number of VTE during learning. X axis stands for trial times. Y axis indicates the number of VTEs. This shows the number changes similarly to real rat's experiments.

In order to identify the role of those VTEs, we choose to perturb all the oscillations identified as VTEs in our previous result. Our hypothesis is that if the VTEs are an epiphenomenon not necessary to the task, perturbing them will not alter the performance of the robot. On the contrary, if the VTEs are being used to complete the task, we should see a variation in performance due to the perturbation.

In order to perturb the VTEs, we apply the same criterions used for counting their amount. Every time a VTE is detected, the angular velocity of the robot is reset to zero which forces it to maintain its current direction. As mentioned previously, we applied a threshold of $[-0.3; 0.3]$ within which an oscillation was not considered a VTE. Due to the subjective nature of this methodology, we analyzed the impact of perturbing the VTEs over a wider set of thresholds. The results of this analysis are show in figure 4.

We can see that there is no visible pattern into the change of performance but that every level of threshold is showing a reduced performance. The only exception is the threshold set at 1.0 which does not cancel any VTEs and displays non altered performance.

To ensure that the drop in performance is not caused by erratic movements of the robots due to the perturbations applied, we compared the trajectories of perturbed runs against a run with threshold 1.0, i.e. not perturbed. We concluded that the trajectory of the robot does not appear to have been altered. The only visible difference resides in the errors in the path taken by the robot. Despite no change in the position of the goal, the robot chooses the wrong arm to follow from one trial to the next. As an example of this type of error, the figure 5 shows 4 trials for 2 different thresholds: 1.0 on the left side and 0.3 on the right. In trials 1 and 2, both robot choose the

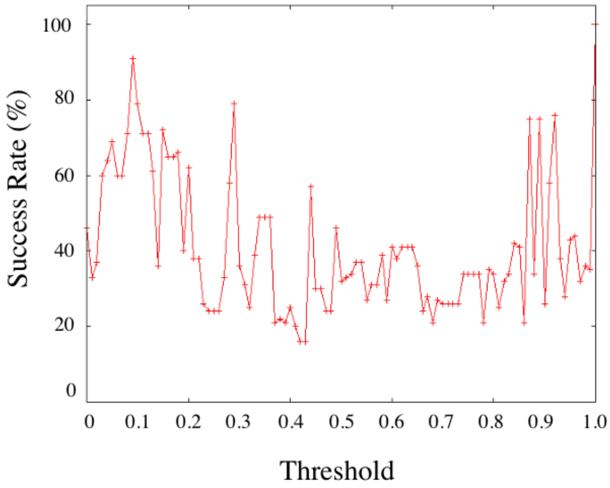

Fig. 4. Change in performance with blocked VTEs. The X axis represents the threshold. Y axis stands for the success rates of the T-Maze task.

same path. On trial 3, the perturbed one makes a mistake while no change in the goal position has occurred. On trial 4, both robots follow the same branch of the T-Maze again. This type of error implies that the learning capacities are suffering from the absence of VTEs.

IV. Conclusion

In this work, we analyzed the role of VTEs on the performance of a robot in a T-Maze task. We found that canceling them lead to a loss of performance without visible impact on the capacity of the controller to explore its environment. The robot rather seemed to have problems learning or remembering where the goal was located between trials. From those results, we can conclude that the bodily oscillations counted as VTEs are not an epiphenomenon but are actively used by the neural controller to facilitate the acquisition of the task.

Acknowledgment

Julien Hubert thanks the Monbukagakusho Scholarship from the Ministry of Education, Culture, Sports, Science and Technology of Japan.

References

[1] E. C. Tolman, “Prediction of vicarious trial and error by means of the schematic sowbug,” *Psychol. Rev.*, vol. 46, pp. 318–336, 1939.

[2] A. Johnson and A. D. Redish, “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*, vol. 27, pp. 12176–89, 2007.

[3] M. Tarsitano, “Route selection by a jumping spider (portia labiata) during the locomotory phase of a detour,” *Animal Behaviour*, vol. 72, pp. 1437–1442, 2006.

[4] S. Bovet and R. Pfeifer, “Emergence of delayed reward learning from sensorimotor coordination,” in *IEEE/RSJ International Conference on Intelligent Robots and Systems(IROS)*, pp. 841–846, Edmonton, 2005.

[5] E. Matsuda, J. Hubert, and T. Ikegami, “A robotic approach to understand the role of vicarious trial-and-error in a t-maze task,” in *Proceedings of the 15th Meeting of the Association for the Scientific Study of Consciousness*, 2011.

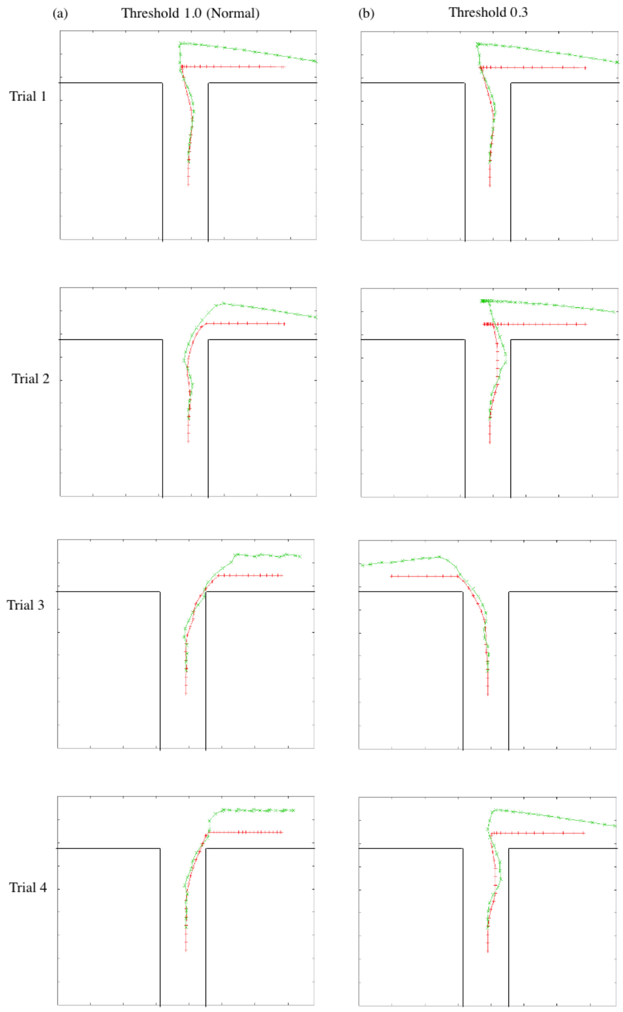

Fig. 5. Trajectory of a robot under 2 different thresholds. The red line indicates the location of the robot while the green line stands for the head direction of the robot. The black solid line indicates the boundary of the T-Maze. (a) The left column shows the behavior without perturbations. (b) The right column shows the behavior with a threshold of 0.3.

[6] F. Mondada, M. Bonani, X. Raemy, J. Pugh, C. Cianci, A. Klaptocz, S. Magnenat, J.-C. Zufferey, D. Floreano, and A. Martinoli, “The e-puck, a robot designed for education in engineering,” in *Proceedings of the 9th Conference on Autonomous Robot Systems and Competitions*, vol. 1, pp. 59–65, 2009.

[7] J. Holland, *Adaptation in Natural and Artificial Systems*. Ann Arbor : The University of Michigan Press, 1975.

[8] K. F. Muenzinger and F. Fletcher, “Motivation in learning: I. electric shock for correct response in the visual discrimination habit,” *J. Comp. Psychol.*, vol. 17, pp. 266–277, 1934.