



HAL
open science

Cooperation of S-R habit and cognitive learnings in goal-oriented navigation

Souheil Hanoune, Jean-Paul Banquet, Mathias Quoy, Philippe Gaussier

► **To cite this version:**

Souheil Hanoune, Jean-Paul Banquet, Mathias Quoy, Philippe Gaussier. Cooperation of S-R habit and cognitive learnings in goal-oriented navigation. SBDM 2014 Paris France, May 2014, Paris, France. pp.42-43. hal-01063964

HAL Id: hal-01063964

<https://hal.science/hal-01063964>

Submitted on 15 Sep 2014

HAL is a multi-disciplinary open access archive for the deposit and dissemination of scientific research documents, whether they are published or not. The documents may come from teaching and research institutions in France or abroad, or from public or private research centers.

L'archive ouverte pluridisciplinaire **HAL**, est destinée au dépôt et à la diffusion de documents scientifiques de niveau recherche, publiés ou non, émanant des établissements d'enseignement et de recherche français ou étrangers, des laboratoires publics ou privés.

Cooperation of S-R habit and cognitive learnings in goal-oriented navigation

Souheil Hanoune^{1 2}, Jean Paul Banquet², Mathias Quoy^{1 2}, Philippe Gaussier^{1 2}

¹University of Cergy Pontoise, France. ²ETIS lab, UMR 8051 CNRS, University of Cergy-Pontoise, ENSEA, France.
e-mail: {souheil.hanoune, quoy, gaussier}@ensea.fr

The Stimulus-Response (S-R) theory and Tolman’s Cognitive Theory of behavior control both issued from behaviorism in the early XXth century still provide a relevant general framework to account for animal reward-based adaptive behavior. In this paper, we propose a new paradigm for representing and implementing both the cognitive strategy and the S-R habit strategy within a unitary coding frame. Based on a parallel learning of both strategies, the model explains how the fast learning cognitive strategy can supervise and accelerate the slow learning S-R habit strategy. This parallel representation is inspired by the cortico-basal functional loops [Alexander et. al., 1986] and the cooperation between the cognitive associative loop, including the dorso-medial striatum and the mPF; and the sensory-motor loop, associated to the sensory motor cortex in relation with the dorso-lateral striatum.

The implementation of S-R habit strategy is different from a classical Q-learning and is based on the model of [Hirel et al., 2013], representing the functioning of the sensory-motor loop. This representation is based on a neuronal modelisation of the Q-learning algorithm. The states of the model are represented by the *transitions*. A transition is an association between two consecutive place-cells during the exploration of the environment, learned in the CA1-CA3 regions of the hippocampus. The cognitive strategy is based on a map representation of the environment namely the cognitive map [O’Keefe & Nadel, 1978]. Based on the association between learned transitions, the cognitive map allows the back-propagation of a reward within a tree, allowing the selection of the shortest path to the goal. While the cognitive map is quickly learned, the Q-values associated with the Q-learning are slower to acquire. On the other hand, the Q-learning tends to be more accurate than the cognitive map when fully learned.

The model exploits this speed difference in its parallel learning. The fast acquisition of the cognitive map allows the robot to quickly choose correct paths to the goal, and thus the time convergence of the Q-learning algorithm is optimized. The cooperation is based on the biasing of the selected transition by the cognitive map and the Q-learning in parallel (see Fig 1). In its early learning stage, the Q-learning biasing is too weak, and the cognitive map is dominant, inducing the supervision of the S-R habit by the cognitive strategy. In the later learning stages, the Q-learning is stronger and more precise. Cooperation of the cognitive strategy and S-R habit enables a faster S-R learning; as shown on Fig 2. It is also possible to study the lesioning of one or the other path and its effect on the behavior and the learning process.

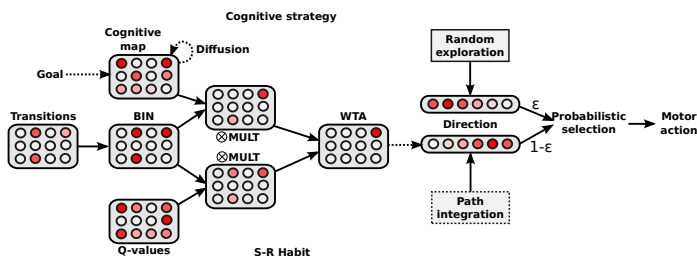


Figure 1: Cooperative architecture of the cognitive strategy and the S-R habit.

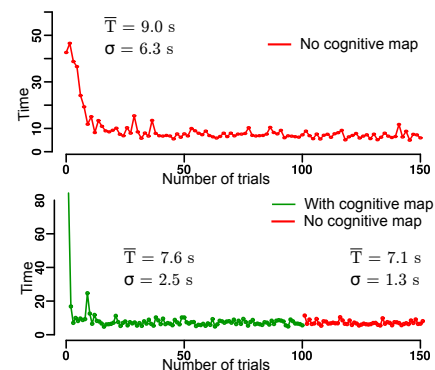


Figure 2: Example of convergence time of the Q-learning with and without the cognitive supervision.

References

- [Alexander et. al., 1986] Alexander, G E, DeLong, M R, Strick, P L (1986). Parallel organization of functionally segregated circuits linking basal ganglia and cortex. *Annual Rev of Neurosci*, 9(1):357–381.
- [Banquet et al., 2014] Banquet, J. P., Gaussier, P., Quoy, M., Revel, A., Burnod, Y. (2014). A hierarchy of associations in hippocampo-cortical systems: cognitive maps and navigation strategies. *Neural Comput.*, 17(6):1339–1384
- [Hirel et al., 2013] Hirel, J., Gaussier, P., Quoy, M., Banquet, J.-P., Save, E., Poucet, B. The Hippocampo-cortical Loop: Spatio-Temporal Learning and Goal-oriented Planning in Navigation (2013). *Neural Networks*, 43,:8–21
- [O’Keefe & Nadel, 1978] O’Keefe, J., Nadel, L. (1978). *The hippocampus as a cognitive map*. Oxford University Press.