Coordination of learning modules for competing navigation strategies into different mazes

Laurent Dollé [1], Mehdi Khamassi [1,3], Agnès Guillot [1] and Ricardo Chavarriaga [2]

[1] Université Pierre et Marie Curie - Paris 6, UMR7606, AnimatLab - LIP6, F-75005 Paris, France
[2] IDIAP Research Institute, CH-1920 Martigny, Switzerland
[3] Laboratoire de Physiologie de la Perception et de l'Action, UMR7152 CNRS, Collège de France, F-75005 Paris, France

Objectives

Improving autonomous navigation in a bioinspired robot

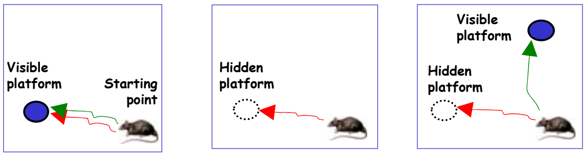
Selection of navigation strategies, especially between 'taxon' and 'locale' ones, in various environments

Taxon

- Visible goal
- Learning of associations Stimulus – Response ('cue-response')
- Egocentric reference

Locale

- Hidden goal
- Learning of associations Stimulus – Location – Response ('place-response')
- Allocentric reference



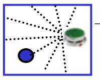
Examples in the Morris water maze

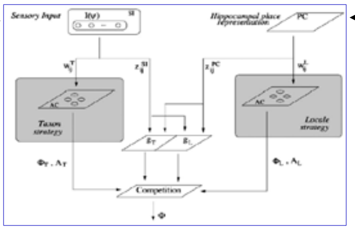
Strategies learned in parallel (Taxon and Locale paths); Selection of the *Locale* one when the platform is hidden; of the *Taxon* one when the platform is located at another position (Devan & White, 1999)

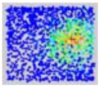
The Model (Chavarriaga et al., 2005)

Sensory inputs: 270° visual field (horizontal greyscale image; allocentric reference)

Sensory inputs: Place cells activation (metric information about walls, odometry; allocentric reference)







'Taxon' Expert (dorsolateral striatum)

'Locale' Expert (ventral-dorsomedial striatum)

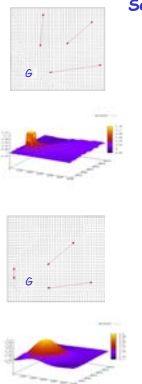
Motor outputs of Taxon and Locale experts: 36 action cells (10° direction each).
A direction is selected at each time-step by the gating network

The Taxon expert (resp. Locale) learns associations between sensory cells (resp. place cells) and action cells.
The gating network selects the most appropriate strategy according to the external input and the internal state of the system.
Both the experts and the gating network modify their parameters by means of reinforcement learning (Q-learning algorithm).

Reimplementation of the model and test in two different mazes

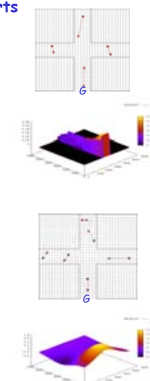
Separate training of experts

(Q-values, Navigation maps)



In the Plus-maze, the Taxon expert relies more on allocentric information than on visual perception.
Modification with an egocentric reference.

In the Plus-maze, the Locale expert relies more on the avoidance reflex than on the navigation map.
Modification with a segmentation of the state-space into sub-experts.

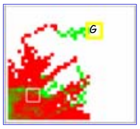


Main implementation differences:
Ad hoc place cells
Addition of an avoidance reflex

Separate training of experts:
10 experiments of 60 trials (visible goal for Taxon; hidden goal for Locale); different starting points.

Competition: same protocol as Chavarriaga's experiments (Devan & White, 1999). 10 experiments with*:*
4 trials each day, different starting points.
Days 1,2,4,5,8: Visible platform, assuming the learning of both Taxon and Locale strategies
Days 3,6,9: Hidden platform, assuming the learning of Locale strategy
Day 10: Competition – New position of the platform, assuming the selection of Taxon strategy instead of Locale one.

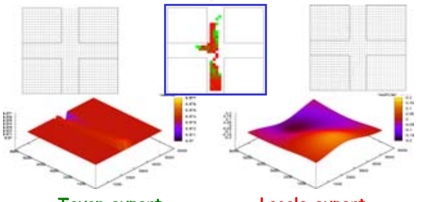
Competition – on Day 10



Morris water maze : 6/10 selections of Taxon expert for reaching the new goal
(old location: SW, new location: NE)
Example starting from S

Taxon and Locale paths

Plus-maze : only 3/10 selections of Taxon expert for reaching the new goal (old location: S, new location: N). Example starting from S

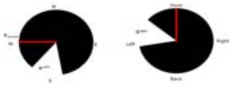


Taxon expert | Locale expert

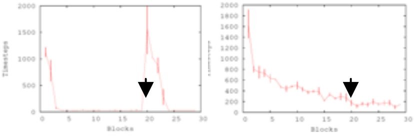
Modifications of the experts
Test in the Plus-maze

Separate training of experts

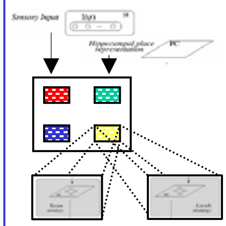
Egocentric reference for the Taxon expert : Better robustness to new goal position, but requires longer training (arrows: new goal position)



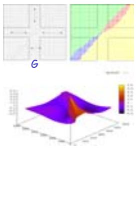
Allocentric | Egocentric



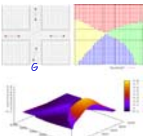
Segmentation into sub-experts for both Taxon and Locale strategies:
Efficient learning depends on the segmentation accuracy.



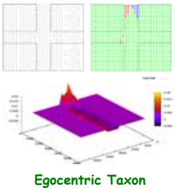
Segmentation of the sensory state-space by SOM (Kohonen maps) with 4 sub-experts
(cf. Khamassi et al., 2006)



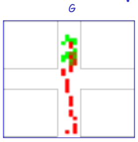
Examples of 'good' *(bottom)* and 'bad' *(left)* segmentations for Locale sub-experts



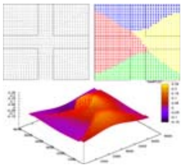
Competition – on Day 10



Egocentric Taxon sub-experts



Example of efficient selection
Taxon and Locale paths
(old goal S, new goal N)
starting from S



Allocentric Locale sub-experts

Still 3/10 selections of Taxon sub-experts for reaching the new goal, but the behaviours of Taxon and Locale sub-experts are more accurate.

Variability of the results due to insufficient training; due to 'bad' segmentations of the state-space; due to the inefficient Taxon strategy when the stimulus is out of the visual field.

Conclusion – Perspectives

Results obtained with the Plus-maze highlight some issues that were not apparent with the Morris water maze. Using an egocentric reference for the Taxon strategy, and a segmentation of the sensory state-space for both Taxon and Locale, improve the learning of navigation maps and Q-values. However, the variability of the results questions the relevance of the current selection criterion.

Even if this criterion seems appropriate in both environments, other mode of selection may be considered in the future: e.g., different weighting for Taxon, Locale (and Praxis, Guidance?) sub-experts in different parts of the environment or in different timing of the task (Packard & McGaugh, 2005).

References

Chavarriaga, R., Strösslin, T., Sheynikhovich, D. and Gerstner, W. (2005). A computational model of parallel navigation systems in rodents. *Neuroinformatics*, 3(3): 223-242

Devan, B. D. and N. M. White (1999). Parallel information processing in the dorsal striatum : Relation to hippocampal function. *Neural Computation*, 19(7): 2789–2798.

Khamassi, M., Martinet, L.-E., and Guillot, A. (2006). Combining self-organizing maps with mixture of experts: Application to an Actor-Critic of reinforcement learning in the basal ganglia. In Nolfi et al. (Eds), *SAB'06 Proceedings*, pp. 394-405.

Packard, M. G. and McGaugh, J. L. (1996). Inactivation of hippocampus or caudate nucleus with lidocaïne differentially affects expression of place and response learning. *Neurobiology of Learning and Memory*, 65(1): 65–72.

Acknowledgments

This research is granted by the European Project 'Integrating Cognition, Emotion and Autonomy' (ICEA).
The authors wish to thank Louis-Emmanuel Martinet for his useful contribution.

Contact: laurent.dolle@lip6.fr
Address: AnimatLab/LIP6,
8 rue du Capitaine Scott, 75015 Paris, France - Tel/Fax: +33 1 44 27 88 09