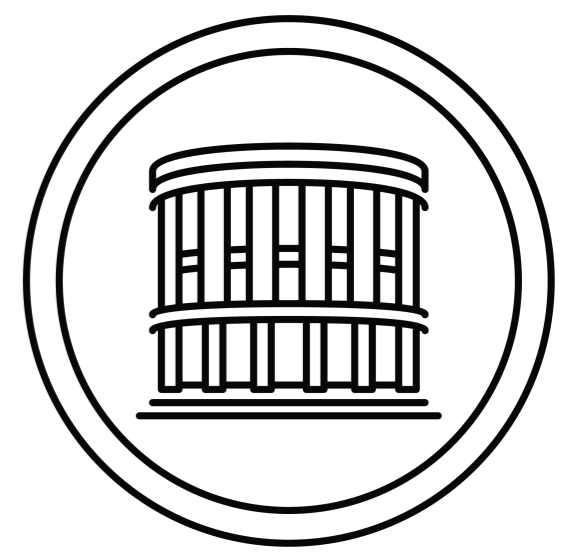


Exploring the Tolman-Eichenbaum Machine for Modeling Spatial Understanding in Robots

Michal Kostrhun, Igor Farkaš
 {kostrhun1, igor.farkas}@fmph.uniba.sk

DEPARTMENT OF APPLIED INFORMATICS, COMENIUS UNIVERSITY IN BRATISLAVA



ABSTRACT

Spatial understanding is a crucial cognitive component in all agents acting in 3D physical space, including people and robots. We investigate the Tolman-Eichenbaum Machine (TEM) for use in biologically inspired Simultaneous Localization and Mapping (SLAM) in robots. Establishing the preliminary context of neuroscience, we provide an approachable overview to the model and its capabilities. We compare it to other approaches used in previous bio-inspired SLAM systems and address their limitations. We summarize advantages and limitations of TEM and briefly outline our future work regarding its use in cognitive robotics.

BIOLOGICAL PRELIMINARIES

The hippocampus and entorhinal cortex (EC) are fundamental to forming cognitive maps in biological agents. Grid cells are present mostly in EC act as a hexagonal positioning system, while place cells in the hippocampus fire at specific locations. [2]

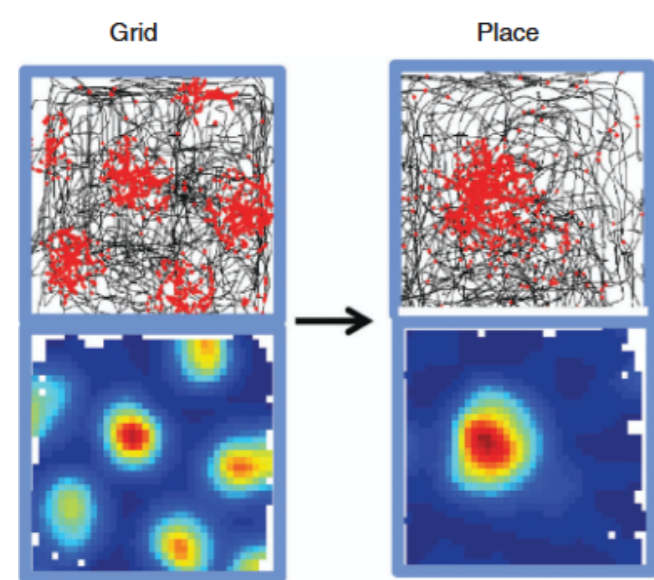


Figure 1: Grid and place cell firing patterns.

EMERGENT CELL TYPES

The TEM architecture leads to the emergence of diverse cell types empirically similar to those found in rodents, including object vector cells, landmark cells, and boundary cells. This makes TEM a promising candidate in bio-inspired SLAM modeling.

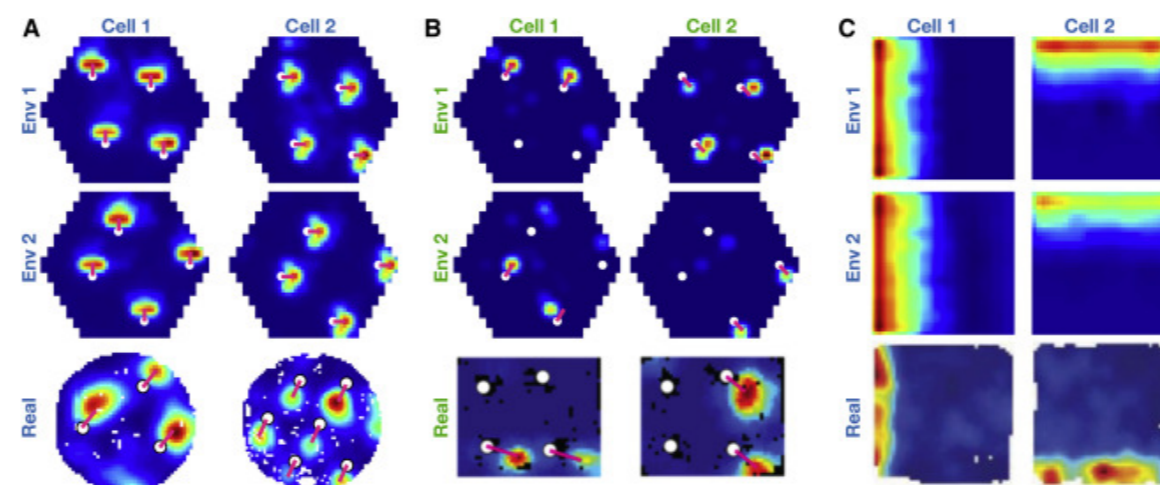


Figure 2: TEM cells compared to real biological data.

PREVIOUS BIO-INSPIRED SLAM

The main challenge with SLAM has always been the spatio-temporal complexity of the environment. Current visual SLAM methods fail to achieve real-time performance [5] and leveraging semantic information to develop semantic SLAM remains an active area of research [3].

Existing models like RatSLAM [1] and ViTa-SLAM [4] emulate hippocampal processes using bio-inspired architecture. RatSLAM utilizes a competitive network of “pose cells” merging position and orientation, while ViTa-SLAM employs a 3D Continuous Attractor Network to emulate grid cells using visuo-tactile fusion via biomimetic whiskers.

Both models are simplified abstractions that lack distinct biological cell populations. Key drawbacks include high computational costs for large networks, accumulated path integration drift, and a reliance on non-standard sensors, often prioritizing topological consistency over metric accuracy.

THE TOLMAN-EICHENBAUM MACHINE

The Tolman-Eichenbaum Machine (TEM) [6] is a unifying mechanistic framework for the hippocampal role in spatial and non-spatial tasks modeled as a graph where nodes are arbitrary sensory information and edges are transitions between them.

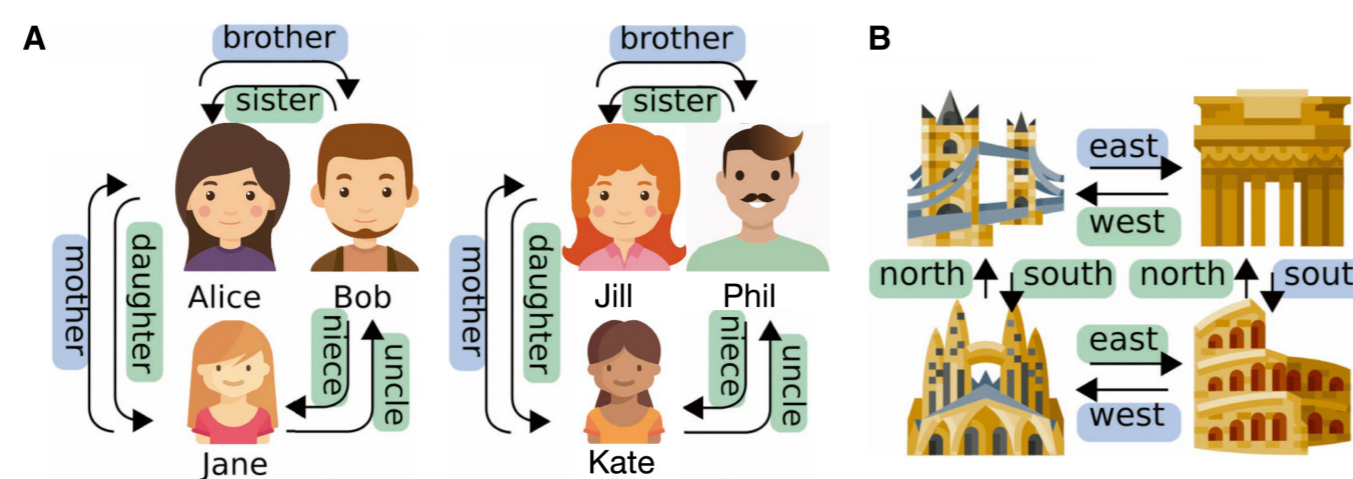


Figure 3: Examples of graph tasks in space and non-space.

TEM’s grid cells learn the structure of the environment during training beforehand. During operation, an agent is placed on the environment graph and TEM, running separately, receives the agent’s edge transitions and sensory information from each visited node. On each time step, TEM stores its representation of input sensory information (“what”) in conjunction with its location representation (“where”) inside its auto-associative network. This allows the future retrieval of “what” upon providing “where” (and vice versa) via attractor mechanics. TEM predicts the next “what” by inferring the next “where” from the desired edge transition, then retrieving the missing “what” of the conjunction with “where” using attractor mechanics.

LIMITATIONS & FUTURE WORK

To become a viable SLAM solution, TEM must overcome several challenges:

Continuous Space: Moving from discrete graphs to continuous 3D Euclidean space.

High-Dimensional Input: Integrating with visual encoders to process complex robotic sensory data (e.g., raw pixels, tactile feedback).

Computation: Addressing the high demands of iterative attractor networks for real-time performance.

In the future, we plan to implement TEM on the NICO humanoid robot, testing its mapping capabilities in simulated 3D environments and comparing its performance against state-of-the-art SLAM systems.

REFERENCES

- [1] Michael J Milford, Gordon F Wyeth, and David Prasser. “RatSLAM: a hippocampal model for simultaneous localization and mapping”. In: *IEEE International Conference on Robotics and Automation*. 2004.
- [2] May-Britt Moser, David C Rowland, and Edvard I Moser. “Place cells, grid cells, and memory”. In: *Cold Spring Harbor perspectives in biology* 7.2 (2015).
- [3] Saeid Nahavandi et al. “A comprehensive review on autonomous navigation”. In: *ACM Computing Surveys* 57.9 (2025).
- [4] Oliver Struckmeier et al. “ViTa-SLAM: A bio-inspired visuo-tactile slam for navigation while interacting with aliased environments”. arXiv:1906.06422. 2019.
- [5] Yuxiang Sun, Ming Liu, and Max Q.-H. Meng. “Improving RGB-D SLAM in dynamic environments: A motion removal approach”. In: *Robotics and Autonomous Systems* 89 (2017).
- [6] James CR Whittington et al. “The Tolman-Eichenbaum machine: unifying space and relational memory through generalization in the hippocampal formation”. In: *Cell* 183.5 (2020).

ACKNOWLEDGEMENT

This work was supported by Horizon Europe MSCA project no. 101072488.

M.K. acknowledges the grant from the Slovak Society for Cognitive Science.