Spatial and Temporal Hierarchy for Autonomous Navigation

Subjects: Computer Science, Artificial Intelligence Contributor: Daria de Tinguy , Toon Van de Maele , Tim Verbelen , Bart Dhoedt

Robust evidence suggests that humans explore their environment using a combination of topological landmarks and coarse-grained path integration. This approach relies on identifiable environmental features (topological landmarks) in tandem with estimations of distance and direction (coarse-grained path integration) to construct cognitive maps of the surroundings. This cognitive map is believed to exhibit a hierarchical structure, allowing efficient planning when solving complex navigation tasks.

active inference

autonomous navigation

spatial hierarchy

temporal hierarchy

1. Introduction

The development of autonomous systems that can navigate in their environment is a crucial step towards building intelligent agents that can interact with the real world. Just as animals possess the ability to navigate their surroundings, developing navigation skills in artificial agents has been a topic of great interest in the field of robotics and artificial intelligence ^{[1][2][3]}. This has led to the exploration of various approaches, including taking inspiration from animal navigation strategies (e.g., building cognitive maps ^[4]), as well as state-of-the-art techniques using neural networks ^[5]. However, despite significant advancements, there are still limitations in both non-neural-network- and neural-network-based navigation approaches ^{[2][3]}.

In the animal kingdom, cognitive mapping plays a crucial role in navigation. Cognitive maps allow animals to understand the spatial layout of their surroundings ^{[6][7][8]}, remember key locations, solve ambiguities from context ^[9], and plan efficient routes ^{[9][10]}. By leveraging cognitive mapping strategies, animals can successfully navigate complex environments, adapt to changes, and return to previously visited places.

In the field of robotics, traditional approaches have been explored to develop navigation systems. These approaches often rely on explicit mapping and planning techniques, such as grid-based ^{[11][12]} and/or topological maps ^{[13][14]}, to guide agent movement. While these methods have shown some success, they suffer from limitations in handling complex spatial relationships and dynamic environments as well as scalability issues as the environment grows larger ^{[2][3][15]}.

To overcome the limitations of these non-neural network approaches, recent advancements have focused on utilising neural networks for navigation ^{[5][16][17][18]}. Neural-network-based models, trained on large datasets, have shown promise in learning navigational policies directly from raw sensory input. These models can capture

complex spatial relationships and make decisions based on learned representations. However, the current neuralnetwork-based navigation approaches also face challenges, including the need for extensive training data, limitations in generalisation to unseen environments, distinguishing aliased areas, and the difficulty of handling dynamic and changing environments ^[2].

Active inference is a framework allowing agents to actively gather information through perception, select and execute actions in their environment, and learn from accumulated experiences ^{[19][20]}. World models, within this framework, form internal representations of the world. Agents endowed with a world model and engaged in active exploration continually update their internal understanding of the environment, empowering them to make well-informed decisions and predictions ^{[21][22]}. This principled approach enables continuous belief updates and active information gathering, facilitating effective navigation ^[20].

Noting that biological agents are building hierarchically structured models, researchers construct multi-level world models as hierarchical active inference. Hierarchical active inference warrants agents to utilise layers of world models, facilitating a higher level of spatial abstraction and temporal coarse-graining. It enables learning complex relationships in the environment and allows more efficient decision-making processes and robust navigation capabilities ^[23]. By incorporating hierarchical structures into active inference-based navigation systems, agents can effectively handle complex environments and perform tasks with greater adaptability ^[24].

2. Spatial and Temporal Hierarchy for Autonomous Navigation

Navigating complex environments is a fundamental challenge for both humans and artificial agents. To solve navigation, traditional approaches often address simultaneous localisation and mapping (SLAM) by building a metric (grid) map [11][12] and/or topological map of the environment [13][14]. Although there is progress in this area, Placed et al. ^[3] state that active SLAM may still fail to be fully autonomous in complex environments. The current approaches are also still lacking in distinct capabilities important for navigation, such as predicting the uncertainty over robot location, abstracting over features of the environment (e.g., having a semantic map instead of a precise 3D map), and reasoning in dynamic, changing spaces. The recent studies have explored the adoption of machine learning techniques to add autonomy and adaptive skills in order to learn how to handle new scenarios in realworld situations. Reinforcement learning (RL) typically relies on rewards to stimulate agents to navigate and explore. In contrast, the model breaks away from this convention, as it does not necessitate the explicit definition of a reward during agent training. Moreover, despite the success of recent machine learning, these techniques typically require a considerable amount of training data to build accurate environment models. This training data can be obtained from simulation $\frac{[25][26]}{25}$; provided by humans (either by labelling, as in the works in $\frac{[27][28]}{25}$ or by demonstration, as in ^[29]); or by gathering data in an experimental setting ^{[16][30][31]}. These methods all aim to predict the consequences of actions in the environment but typically generalise poorly across environments. As such, they require considerable human intervention when deployed in new settings ^[2]. The aim is to reduce both the human intervention and the quantity of data required for training by simultaneously familiarising the agent with the structure and dynamics found in its environment.

When designing an autonomous adaptable system, nature is a source of inspiration. Tolman's cognitive map theory ^[32] proposes that brains build a unified representation of the spatial environment to support memory and guide future actions. More recent studies postulate that humans create mental representations of spatial layouts to navigate ^[6], integrating routes and landmarks into cognitive maps ^[7]. Additionally, the research into neural mechanisms suggests that spatial memory is constructed in map-like representations fragmented into sub-maps with local reference frames ^[33]; meanwhile, hierarchical planning is processed in the human brain during navigation tasks ^[9]. The studies of Balaguer et al. ^[9] and Tomov et al. ^[10] show that hierarchical representations are essential for efficient planning for solving navigation tasks. Hierarchies provide a structured approach for agents to learn complex environments, breaking down planning into manageable levels of abstraction and enhancing navigation capabilities, both spatially (sub-maps) and temporally (time-scales). Thus, the model incorporates these elements as the foundation of its operation.

The concept of hierarchical models has gained interest in navigation research ^{[13][24]}. Hierarchical structures enable agents to learn complex relationships within the environment, leading to more efficient decision-making and enhancing adaptability in dynamic scenarios. There are two main types of hierarchy, both considered in the work: temporal—planning over a sequence of timesteps ^{[34][35][36][37]}—and spatial—planning over structures ^{[13][23][38][39]}.

In order to navigate without teaching the agent how to do so, researchers use the principled approach of active inference (AIF), a framework combining perception, action, and learning. It is a promising avenue for autonomous navigation ^[22]. By actively exploring the environment and formulating beliefs, agents can make informed decisions. Within this framework, world models play a pivotal role in creating internal representations of the environment and facilitating decision-making processes. A few models have proposed combining AIF and hierarchical models for navigation. Safron et al. ^[40] proposes a hierarchical model composed of two layers of complexity to learn the structure of the environment. The lowest level infers the state of each step while the higher level represents locations, created in a more coarse manner. Large, complex, aliased, and/or dynamic environments are challenges to this model. Nozari et al. ^[41] construct a hierarchical system by using a dynamic Bayesian network (DBN) over a naive and an expert agent, in which the naive agent learns temporal relationships, with the highest level capturing semantic information about the environment and low-level distributions capturing rough sensory information with their respective evolution through time. This system, however, requires expert data to be trained by imitation learning, which limits the performance of the model to that of the expert.

References

- 1. Schwartenbeck, P.; Passecker, J.; Hauser, T.U.; FitzGerald, T.H.; Kronbichler, M.; Friston, K.J. Computational mechanisms of curiosity and goal-directed exploration. eLife 2019, 8, e41703.
- 2. Levine, S.; Shah, D. Learning robotic navigation from experience: Principles, methods and recent results. Philos. Trans. R. Soc. B Biol. Sci. 2022, 378, 20210447.

- Placed, J.A.; Strader, J.; Carrillo, H.; Atanasov, N.; Indelman, V.; Carlone, L.; Castellanos, J.A. A Survey on Active Simultaneous Localization and Mapping: State of the Art and New Frontiers. IEEE Trans. Robot. 2023, 39, 1686–1705.
- 4. George, D.; Rikhye, R.; Gothoskar, N.; Guntupalli, J.S.; Dedieu, A.; Lázaro-Gredilla, M. Clonestructured graph representations enable flexible learning and vicarious evaluation of cognitive maps. Nat. Commun. 2021, 12, 2392.
- 5. Hafner, D.; Pasukonis, J.; Ba, J.; Lillicrap, T. Mastering Diverse Domains through World Models. arXiv 2023, arXiv:2301.04104.
- 6. Epstein, R.; Patai, E.Z.; Julian, J.; Spiers, H. The cognitive map in humans: Spatial navigation and beyond. Nat. Neurosci. 2017, 20, 1504–1513.
- Foo, P.; Warren, W.; Duchon, A.; Tarr, M. Do Humans Integrate Routes Into a Cognitive Map? Map- Versus Landmark-Based Navigation of Novel Shortcuts. J. Exp. Psychology. Learn. Mem. Cogn. 2005, 31, 195–215.
- 8. Peer, M.; Brunec, I.K.; Newcombe, N.S.; Epstein, R.A. Structuring Knowledge with Cognitive Maps and Cognitive Graphs. Trends Cogn. Sci. 2021, 25, 37–54.
- 9. Balaguer, J.; Spiers, H.; Hassabis, D.; Summerfield, C. Neural Mechanisms of Hierarchical Planning in a Virtual Subway Network. Neuron 2016, 90, 893–903.
- 10. Tomov, M.S.; Yagati, S.; Kumar, A.; Yang, W.; Gershman, S.J. Discovery of Hierarchical Representations for Efficient Planning. bioRxiv 2018.
- Lakaemper, R.; Latecki, L.J.; Sun, X.; Wolter, D. Geometric Robot Mapping. In Discrete Geometry for Computer Imagery; Andres, E., Damiand, G., Lienhardt, P., Eds.; Springer: Berlin/Heidelberg, Germany, 2005; Volume 3429, pp. 11–22.
- Kamarudin, K.; Mamduh, S.M.; Shakaff, A.Y.M.; Zakaria, A. Performance Analysis of the Microsoft Kinect Sensor for 2D Simultaneous Localization and Mapping (SLAM) Techniques. Sensors 2014, 14, 23365–23387.
- Ge, S.S.; Zhang, Q.; Abraham, A.T.; Rebsamen, B. Simultaneous Path Planning and Topological Mapping (SP2ATM) for environment exploration and goal oriented navigation. Robot. Auton. Syst. 2011, 59, 228–242.
- Kim, S.H.; Kim, J.G.; Yang, T.K. Autonomous SLAM technique by integrating Grid and Topology map. In Proceedings of the 2008 International Conference on Smart Manufacturing Application, Goyangi, Republic of Korea, 9–11 April 2008; pp. 413–418.
- Li, Z.; Chen, G.; Peng, B.; Zhu, X. Robot Navigation Method based on Intelligent Evolution. In Proceedings of the 2018 IEEE 4th Information Technology and Mechatronics Engineering Conference (ITOEC), Chongqing, China, 14–16 December 2018; pp. 620–624.

- 16. Parisi, S.; Dean, V.; Pathak, D.; Gupta, A. Interesting Object, Curious Agent: Learning Task-Agnostic Exploration. arXiv 2021, arXiv:2111.13119.
- 17. Matsuo, Y.; LeCun, Y.; Sahani, M.; Precup, D.; Silver, D.; Sugiyama, M.; Uchibe, E.; Morimoto, J. Deep learning, reinforcement learning, and world models. Neural Netw. 2022, 152, 267–275.
- Gregor, K.; Jimenez Rezende, D.; Besse, F.; Wu, Y.; Merzic, H.; van den Oord, A. Shaping Belief States with Generative Environment Models for RL. In Proceedings of the Advances in Neural Information Processing Systems, Vancouver, BC, Canada, 8–14 December 2019; Wallach, H., Larochelle, H., Beygelzimer, A., d'Alché-Buc, F., Fox, E., Garnett, R., Eds.; Curran Associates, Inc.: Red Hook, NY, USA, 2019; Volume 32.
- 19. Friston, K. Life as we know it. J. R. Soc. Interface R. Soc. 2013, 10, 20130475.
- 20. Friston, K.; FitzGerald, T.; Rigoli, F.; Schwartenbeck, P.; Doherty, J.O.; Pezzulo, G. Active inference and learning. Neurosci. Biobehav. Rev. 2016, 68, 862–879.
- 21. Ha, D.; Schmidhuber, J. World Models. arXiv 2018, arXiv:1803.10122.
- 22. Friston, K.; Moran, R.J.; Nagai, Y.; Taniguchi, T.; Gomi, H.; Tenenbaum, J. World model learning and inference. Neural Netw. 2021, 144, 573–590.
- 23. Stoianov, I.; Maisto, D.; Pezzulo, G. The hippocampal formation as a hierarchical generative model supporting generative replay and continual learning. Prog. Neurobiol. 2022, 217, 102329.
- 24. Çatal, O.; Verbelen, T.; Van de Maele, T.; Dhoedt, B.; Safron, A. Robot navigation as hierarchical active inference. Neural Netw. 2021, 142, 192–204.
- 25. Sadeghi, F.; Levine, S. CAD2RL: Real Single-Image Flight without a Single Real Image. arXiv 2016, arXiv:1611.04201.
- 26. Müller, M.; Dosovitskiy, A.; Ghanem, B.; Koltun, V. Driving Policy Transfer via Modularity and Abstraction. arXiv 2018, arXiv:1804.09364.
- 27. Janai, J.; Güney, F.; Behl, A.; Geiger, A. Computer Vision for Autonomous Vehicles: Problems, Datasets and State-of-the-Art. arXiv 2017, arXiv:1704.05519.
- Feng, D.; Haase-Schütz, C.; Rosenbaum, L.; Hertlein, H.; Duffhauss, F.; Gläser, C.; Wiesbeck, W.; Dietmayer, K. Deep Multi-modal Object Detection and Semantic Segmentation for Autonomous Driving: Datasets, Methods, and Challenges. arXiv 2019, arXiv:1902.07830.
- 29. Silver, D.; Bagnell, J.A.; Stentz, A. Learning from Demonstration for Autonomous Navigation in Complex Unstructured Terrain. Int. J. Robot. Res. 2010, 29, 1565–1592.
- 30. Gupta, S.; Davidson, J.; Levine, S.; Sukthankar, R.; Malik, J. Cognitive Mapping and Planning for Visual Navigation. arXiv 2017, arXiv:1702.03920.

- 31. Chaplot, D.S.; Gandhi, D.; Gupta, S.; Gupta, A.; Salakhutdinov, R. Learning to Explore using Active Neural SLAM. arXiv 2020, arXiv:2004.05155.
- 32. Tolman, E.C. Cognitive maps in rats and men. Psychol. Rev. 1948, 554, 189–208.
- 33. Madl, T.; Franklin, S.; Chen, K.; Trappl, R.; Montaldi, D. Exploring the Structure of Spatial Representations. PLoS ONE 2016, 11, e0157343.
- 34. Zakharov, A.; Crosby, M.; Fountas, Z. Episodic Memory for Learning Subjective-Timescale Models. arXiv 2020, arXiv:2010.01430.
- 35. Shah, D.; Eysenbach, B.; Rhinehart, N.; Levine, S. RECON: Rapid Exploration for Open-World Navigation with Latent Goal Models. arXiv 2021, arXiv:2104.05859.
- 36. Çatal, O.; Wauthier, S.; De Boom, C.; Verbelen, T.; Dhoedt, B. Learning Generative State Space Models for Active Inference. Front. Comput. Neurosci. 2020, 14, 574372.
- Milford, M.; Jacobson, A.; Chen, Z.; Wyeth, G. RatSLAM: Using Models of Rodent Hippocampus for Robot Navigation and Beyond. In Robotics Research; Inaba, M., Corke, P., Eds.; Springer: Cham, Switzerland, 2016; pp. 467–485.
- 38. Neacsu, V.; Mirza, M.B.; Adams, R.A.; Friston, K.J. Structure learning enhances concept formation in synthetic Active Inference agents. PLoS ONE 2022, 17, e0277199.
- 39. Çatal, O.; Jansen, W.; Verbelen, T.; Dhoedt, B.; Steckel, J. LatentSLAM: Unsupervised multisensor representation learning for localization and mapping. arXiv 2021, arXiv:2105.03265.
- 40. Safron, A.; Çatal, O.; Verbelen, T. Generalized Simultaneous Localization and Mapping (G-SLAM) as unification framework for natural and artificial intelligences: Towards reverse engineering the hippocampal/entorhinal system and principles of high-level cognition. Front. Syst. Neurosci. 2022, 16, 787659.
- 41. Nozari, S.; Krayani, A.; Marin-Plaza, P.; Marcenaro, L.; Gómez, D.M.; Regazzoni, C. Active Inference Integrated With Imitation Learning for Autonomous Driving. IEEE Access 2022, 10, 49738–49756.

Retrieved from https://encyclopedia.pub/entry/history/show/122852