Navigating Conceptual Space;
A New Take on AGI

Per Roald Leikanger[(✉)]

UiT - The Arctic University of Norway, Tromsø, Norway
`Per.R.Leikanger@uit.no`

Abstract. Edward C. Tolman found reinforcement learning unsatisfactory for explaining intelligence and proposed a clear distinction between learning and behavior. Tolman's ideas on latent learning and cognitive maps eventually led to what is now known as conceptual space, a geometric representation where concepts and ideas can form points or shapes. Active navigation between ideas – reasoning – can be expressed directly as purposive navigation in conceptual space. Assimilating the theory of conceptual space from modern neuroscience, we propose autonomous navigation as a valid approach for emulated cognition. However, achieving autonomous navigation in high-dimensional Euclidean spaces is not trivial in technology. In this work, we explore whether neoRL navigation is up for the task; adopting Kaelbling's concerns for efficient robot navigation, we test whether the neoRL approach is general across navigational modalities, compositional across considerations of experience, and effective when learning in multiple Euclidean dimensions. We find neoRL learning to be more resemblant of biological learning than of RL in AI, and propose neoRL navigation of conceptual space as a plausible new path toward emulated cognition.

1 Introduction

Edward C. Tolman first proposed cognitive maps for explaining the mechanism behind rats taking shortcuts and what he referred to as latent learning [25]. Tolman was not satisfied with behaviorists' view that goals and purposes could be reduced to a hard-wired desire for reward [4]. Experiments showed that unrewarded rats would perform better than the fully rewarded group when later *motivated* by reward [26]. Arguing that a reinforcement signal was more important for behavior than for learning, Tolman proposed the existence of a cognitive model of the environment in the form of a *cognitive map*. The mechanisms behind neural representation of Euclidean space (NRES) has later been identified for a range of navigational modalities by electrophysical measurements [3]. Further, the NRES mechanism has been implied for navigating *conceptual space* [5], a Euclidean representation where betweenness and relative location makes sense for explaining concepts [7]. Results from theoretical neuroscience indicate NRES' role in social navigation [17], temporal representation [6], and reasoning [2]. Cognitive maps for representing thought have received much attention in

neurophysiology in the recent five years [2,5,17]. Navigating conceptual space as an analogy of thought could explain generalization and reasoning based on locality [7].
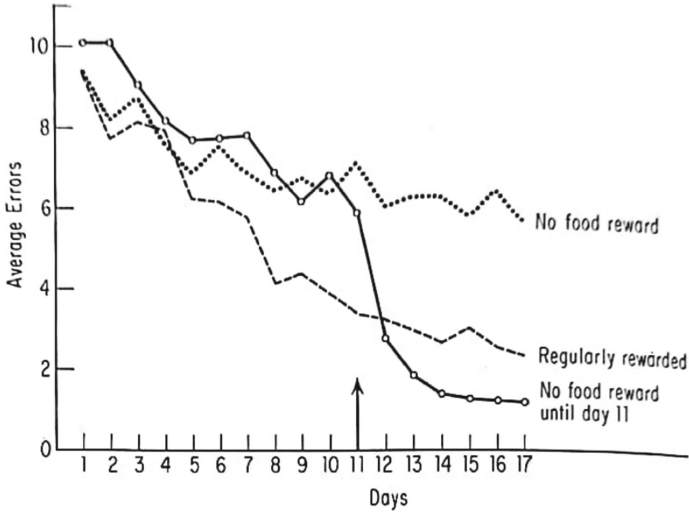

Fig. 1. Evidence for latent learning by Tolman and Honzik (1930). (after [26], from *Systems and theories of psychology* [4]).

Autonomous navigation is difficult to reproduce in technology. Autonomous operation implies a decision *agent* capable of forming decisions based on own desires and experience. A well-renowned approach to establish experience-based behavior is reinforcement learning (RL) from AI. Via trial and error based on a scalar reward signal $\mathbb{R}$, a decision *agent* is capable of adapting behavior according to the accumulation of $\mathbb{R}$. Considering robot path planning as Euclidean navigation, we look toward robot learning for inspiration on autonomous navigation. However; whereas RL powered by deep function approximation has been demonstrated for playing board games at an expert level, requirements to sample efficiency combined with high Markov dimensionality in temporal systems makes deep RL difficult in navigation learning [10]. Leslie Kaelbling (2020) points out key challenges for efficient robot learning, apparently concerned with the current direction of deep RL. Navigation has to be efficient (require few interactions for learning new behaviors), general (applicable to situations outside one's direct experience), and compositional/incremental (compositional with earlier knowledge, incremental with earlier considerations). The current state-of-the-art deep RL for robotics struggles on all three points [9].

Inspired by neural navigation capabilities, Leikanger (2019) has developed an NRES-oriented RL (neoRL) architecture for online navigation [12]. Via orthogonal value functions (OVF) formed by off-policy learning toward each cell of an

NRES representation, the neoRL architecture allows for a distinction between learning and behavior. Inspired by animal psychology, the neoRL framework allows purposive behavior to form based on the desire for anticipated reward [13]. However, navigating a multi-dimensional conceptual space of unknown dimensionality in real-time would be impossible for any current learning algorithm. In this work, we adopt Kaelbling's three concerns when testing whether neoRL navigation allows for autonomous navigation in high-dimensional Euclidean space.

2 Theory

Central to all navigation is knowledge of one's current navigational state. Information about relative location, orientation, and heading to objects that can block or otherwise affect the path is crucial for efficient navigation. When such knowledge is represented as vectors relative to one's current configuration, neuroscientists refer to this representation as being *egocentric*. When represented relative to some external reference frame, coordinates are referred to as being *allocentric*. Vectors can be expressed as Cartesian coordinates, e.g. the vector $\vec{a} = [1.0, 3.0]$ represent a point or displacement in a plane, one unit size from the origin along the first dimension, and three units along a second dimension. Vectors can also be represented in polar coordinates $\vec{a} = [r, \varphi]$, a point with distance r from the origin in the allocentric direction φ. In order to apply RL for navigation, all this information must be represented according to the Markov property; each instance of agent state must contain enough data to define next-state distribution [20]. Combined with temporal dynamics, the number of such instances becomes prohibitively expensive for autonomous navigation by RL [10]. Neural state representation, on the other hand, appears to be fully distributed across individual neurons and parts of the hippocampal formation [18]. NRES coding for separate navigational modalities (as should be represented in separate Euclidean spaces) have been located in different structures in the hippocampal formation [3]. Navigational state representation for the only system capable of true autonomous navigation seems to be decomposed across multiple NRES modalities. This section introduces theory and considerations on how state is represented in the animal and the learning machine, an important inspiration for neoRL mechanisms for navigation and problem solving.

2.1 Neural Representation of Euclidean Space

The first identified NRES neuron was the *place cell* [16]; O'Keefe and Dostrovsky discovered that specific neurons in the hippocampus became active whenever the animal traversed a specific location in the test environment. Reflecting the allocentric position of the animal, the individual place cell could be thought of as a geometric *feature detector* on the animal's location; the place cell is active whenever the animal is located within the *receptive field* of the cell. Other NRES cells have later been identified, expressing information in various parameter spaces. Identified NRES modalities for navigation includes: one's allocentric location
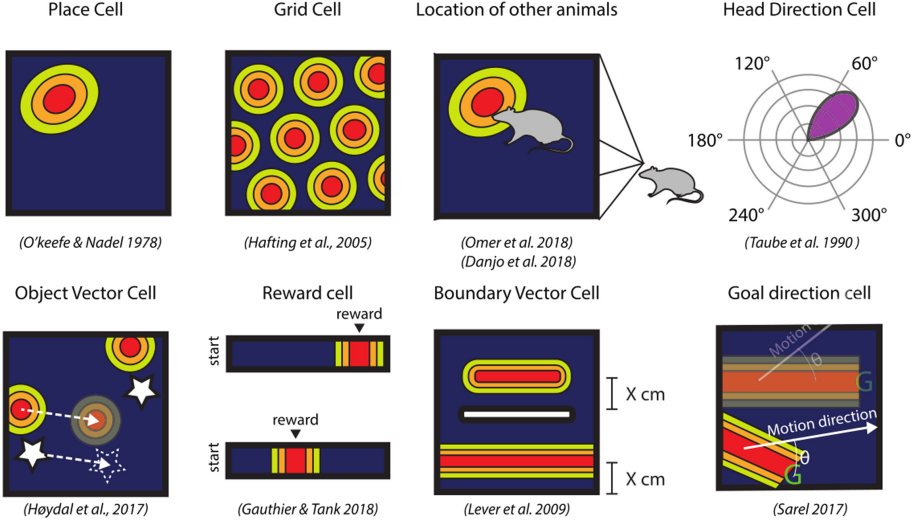
Fig. 2. Some identified NRES modalities of importance for navigation, with reference to the original publication. All NRES modalities could be important for autonomous spatial navigation. The place cell and the object vector cell will be of particular importance in the examples and experiments of this text. (Illustration adopted from [1]

[16], allocentric polar vector coordinates to external objects [8], and one's current heading [24]. A selection of relevant NRES modalities is listed in Table 1 or in Fig. 2. A more comprehensive study on NRES modalities in neurophysiology has been composed by Bugress and Bicanski [3].

Table 1. Neural representation for different Euclidean spaces of importance for navigation: Head-direction cells reflect the current allocentric (*ac.*) angle of the head (a scalar parameter). The place cell and border cell respond to a proximal allocentric location (2D). The remaining NRES reflect conditions represented in other Euclidean spaces – listed as NRES modalities.

	Location	Tuning	Direction	NRES modality	
Place cell	Ac.	[proximal] 2D	–	Current position	[16]
Border cell	Ac.	[proximal] 2D	–	Location of borders	[19]
Object vector cell	Polar c.	[spectrum] 2D	Ac.	Location of objects	[8]
Boundary vector cell	Polar c.	[spectrum] 2D	Ac.	Location of boundaries	[14]
Head-direction cell	–	[angular] 1D	Ac.	Head direction	[24]
Speed cell	–	[rate code] 1D	–	Current velocity	[11]

Neuroscientists assume that populations of NRES neurons map Euclidean vectors by neural patterns of activation. A simple mapping could be formed by

a population of NRES cells responding to mutually exclusive receptive fields. One could visualize this representation as a chessboard; exactly one cell (tile) would be satisfied for any point on the board. Referred to as *one-hot encoding*[1] in computing sciences, a mutually exclusive map structure is defined by the resolution and the geometric coverage of the map's tiles. This intuitive map is appropriate for demonstrative purposes: All examples and experiments in this text are encoded by a comprehensive one-hot mapping as illustrated in Fig. 3, where, e.g., $N13$ signifies a 13×13 tile set in $\Re^2$.

2.2   Autonomous Navigation by neoRL Agents

One can separate navigation into two distinct aspects; the desired location – the objective of the interaction – and how this objective can be reached. When both aspects are governed by one's own inclinations and *experience*, we refer to this as an autonomous operation. A most accomplished approach to experience-based behavior is RL from AI; a decision *agent* can be thought of as an algorithmic entity that learns how better to reach an objective by trial and error. The decision process of the agent can be summarized by 3 signals: the *state* of the system before the interaction, the *action* with which the agent interacts with the system, and a *reward* signal that reflects the success of the operation with regard to an objective. Experience can be expressed via the *value function*, reflecting the expected total reward from this state and forward under the current policy. Since behavior (policy) is based on the current value function, and the value function is defined under one policy, an alternating iterative improvement is required for learning. The resulting asymptotic progress is slow, requiring many interactions by RL learning. Although RL has proven effective for solving a range of algorithmic tasks, autonomous control for robotics remains a challenge [10]. Even RL powered by deep function approximation (deep RL) has limited applicability for online interaction learning in Euclidean spaces [9].

A neoRL agent, on the other hand, is composed by a set of sub-agents learning how to achieve different NRES cells for the corresponding NRES representation [12]. The whole set of learning processes constitutes the (latent) learning aspect of the agent; behavior can later be harvested as a weighted sum over the OVFs according to priorities [13]. Learning OVFs as general value functions [21] with $\mathbb{R}$ defined by NRES cell activation, the value function of the whole neoRL agent resembles Kurt Lewin's *fieldt theory* of learning [15]. Leikanger (2021) demonstrated how emulated NRES for agent state allows for autonomous navigation in a single Euclidean space [13], however, multi-modal navigation and combining experience across NRES modalities remains to be tested. As multi-modal NRES capabilities would bring neoRL state representation closer to navigational state representation in the brain, compositionality across NRES modalities would be important for making neoRL a plausible candidate for conceptual navigation.

---

[1] Note for computing scientists: NRES is not concerned with the Markov state. Any similarity to RL coarse coding and CMAC can therefore be considered to be an endorsement of these AI techniques, not grounds for direct comparison.
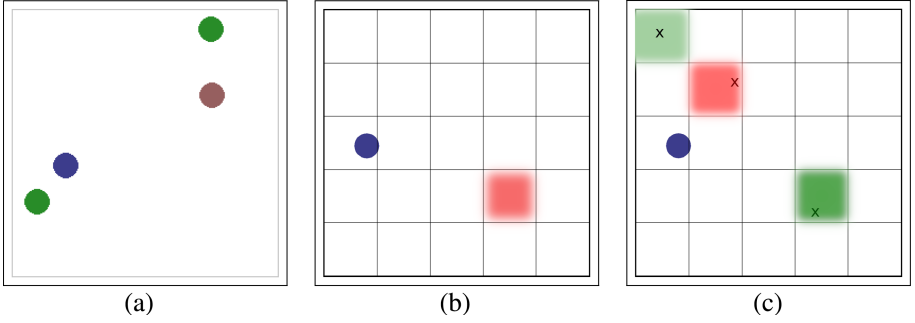
Fig. 3. (A) The allocentric WaterWorld environment: Blue entity is governed by inertia dynamics, with a desire for green ($\mathbb{R} = +1$) and aversion for red ($\mathbb{R} = -1$). (B) An $N5$ mapping of NRES: Each axis is divided into $N = 5$ equal intervals, resulting in $N^2 = 25$ NRES cells. An OVF represents the value function toward one NRES activation. (C) Learned NRES maps can form behaviors via anticipated reward: When an NRES tile contains an element associated with reward, the corresponding OVF is weighted accordingly. Anticipated rewards are illustrated using the same colors as in (A); one aversive NRES cell in red and two desirable NRES cells associated with various anticipation are represented in shades of green. (Color figure online)

3   Multi-modal neoRL Navigation

Adopting Kaelbling's three concerns for Euclidean navigation, we next explore how neoRL navigation scales with increasing (Euclidean) dimensionality. First, it is crucial that NRES-oriented navigation can operate based on different Euclidean spaces; with little knowledge of the form or meaning of conceptual spaces, neoRL must be capable of navigation by other information than location. Further, we are interested in how neoRL navigation scales with additional parameters or across multiple NRES modalities. Any exponential increase in training time with additional states would make conceptual navigation infeasible. NeoRL navigation must be *general* across NRES modalities, *compositional* across conceptual components, without any significant decline in learning *efficiency*. In this section, we explore neoRL capabilities for hi-dimensional navigation by experiments inspired by Kaelbling's concerns for efficient robot navigation.

All experiments are conducted in the allocentric version of the WaterWorld environment [23], illustrated in Fig. 3A. An agent controls the movement of the self (blue dot), with a set of actions that accelerate the object in the four directions N, S, E, W. Three objects of interest move freely in a closed section of a Euclidean plane. When the agent encounters an object, it is replaced by a new object with a random color, location, and speed vector. Green objects are desirable with an accompanying reward $\mathbb{R} = +1.0$, and red objects should be avoided with $\mathbb{R} = -1.0$. No other rewards exist in these experiments, making $\mathbb{R}$ a decent measure of an agent's navigation capabilities. Note that the agent must catch the last green in a board full of red before the board can be reset and continue beyond (on average) 1.5 points per reset.

The PLE [22] version of WaterWorld reports the Cartesian coordinates of the agent and elements of interest; thinking of the Euclidean plane in Fig. 3A as representing location facilitates later discussion. A direct NRES encoding of this information will be referred to as place cell (PC) NRES modality in the remainder of this text. One can also compute a simple object vector cell (OVC) interpretation by vector subtraction:

$$\vec{o^i}_{\text{OVC}} = \vec{o^i}_{\text{PC}} - \vec{s}_{\text{PC}}$$

where $\vec{s}$ is the location of self and $\vec{o^i}$ is the location of object i in PC or OVC reference frame. Note that this OVC interpretation allows for a modality similar to OVC with the self in the center and allocentric direction to external objects, but not with polar coordinates as reported for OVC [8]. However, the two Cartesian representations of location still give different points of view due to different reference frames. Information is encoded in NRES maps as described in Sect. 2.1; the neoRL agent is organized across multiple NRES maps of different resolutions as described in [13]. Multi-res NRES modalities cover resolutions given by primes up to $N13$, i.e., with layers $N2, N3, N5, N7, N11,$ and $N13$. For more on multi-resolution neoRL agents and the mechanism behind policy from parallel NRES state spaces, see [13]. All execution runs smoothly on a single CPU core, and the agent starts with no priors other than described in this section. Referring to the NRES modalities as PC and OVC for WaterWorld is only syntactical to facilitate later discussions; 2D Euclidean coordinates are general and can represent any parameter pair.

Learning efficiency is compared by considering the transient proficiency of the agent as measured by the reward received by the agent during $0.2s$ intervals. Any end-of-episode reward is disabled in the WaterWorld settings; the only received reward is $\mathbb{R} = +1$ when encountering green elements and $\mathbb{R} = -1$ when encountering red elements. The simple reward structure makes accumulated $\mathbb{R}$ a direct measure of how well the agent performed during one run. However, observing the transient proficiency – real-time learning efficiency – of the agent requires further analysis: in all experiments, a per-interval average or received reward is computed over 100 independent runs with additional smoothing by a Butterworth low-pass filter. All runs are conducted in isolation; the agent is initiated before each run and deleted after the run – without any accumulation of experience between runs. The x-axis of every plot represents minutes since agent initiation. The y-axis represents proficiency as computed by the per-interval average of received reward, scaled to reflect $[\mathbb{R}/s]$. Proficiency thus measures how many more desirable (green) encounters happen per second than unwanted (red) ones.

3.1   NeoRL Navigation: NRES Generality

First, we examine the generality of the neoRL architecture by comparing navigational proficiency for an agent exposed to a PC modality to one exposed to an OVC modality. We are interested in the generality of neoRL navigation; can neoRL navigate the PC modality by different Euclidean information, and at what cost?
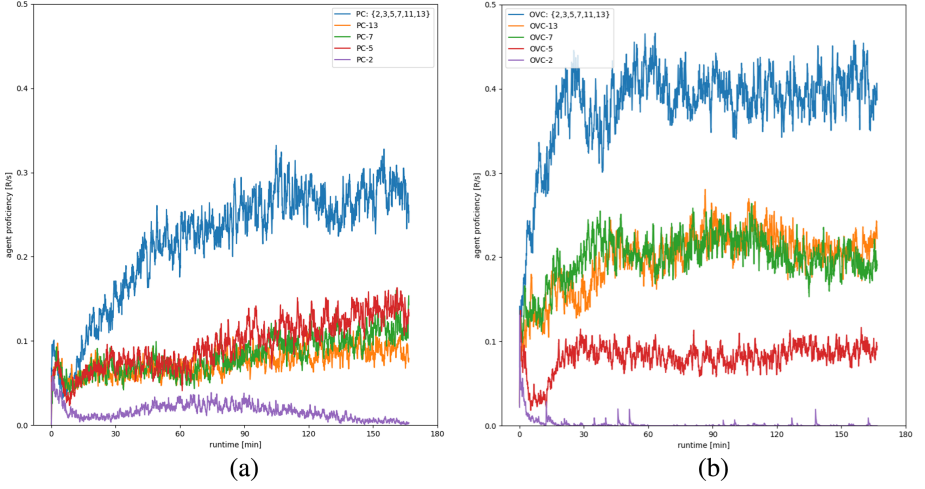
Fig. 4. The neoRL architecture is general across NRES modalities: (A) an original place cell (PC) NRES modality, implemented by applying NRES code directly on an allocentric location of the agent or elements of interest. (B) an emulated object vector cell (OVC) NRES modality, implemented by vector subtraction. OVC is centered on the self with an allocentric representation of other objects.

Results are presented in Fig. 4: agent proficiency from the original PC modality (Fig. 4A) can be compared with agent proficiency when navigating by the OVC modality (Fig. 4B). The immediate proficiency of several mono-resolution neoRL agents is plotted alongside the proficiency of a multi-resolution neoRL agent. There is no loss in sample efficiency when utilizing the OVC modality compared to PC modality. The multi-res neoRL agent performs better than mono-res neoRL agents for both the PC and the OVC modality. NeoRL navigation performs well across both aspects of experience, indicating that the neoRL architecture is general across navigational modalities.

3.2   NeoRL Navigation: NRES Compositionality

Secondly, we are interested in how neoRL scales with additional navigational information. Experiment 1 showed how a neoRL agent is capable of reactive navigation based on an auxiliary NRES modality. In this experiment, we explore the benefit of combining experience across more than one NRES modality. A multi-modal neoRL agent is exposed to both the PC and the OVC modality from experiment 1, effectively doubling the number of NRES states for the agent to consider. We are anxious about how well the neoRL architecture scales with the additional information, both for final proficiency and learning time.
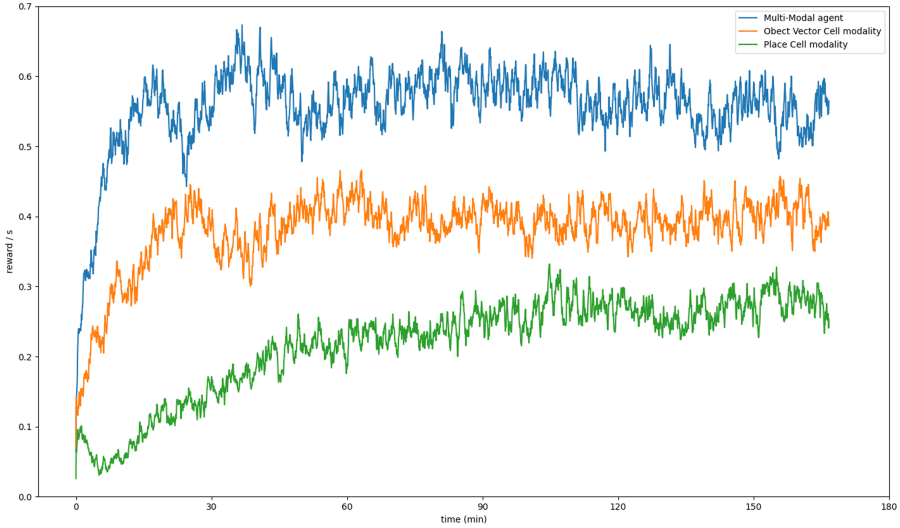
Fig. 5. Multi-modal neoRL navigation leads to higher proficiency and quicker learning than mono-modal agents, despite having twice as many NRES states.

Compare the proficiency of the neoRL agent when exposed to PC, OVC, and multi-modal information in Fig. 5. Combining information across multiple NRES modalities significantly improves navigational performance. The final proficiency of the multi-modal neoRL agent approaches $0.55[\mathbb{R}/s]$ while the PC neoRL agent barely reaches $0.29[\mathbb{R}/s]$. The multi-modal neoRL agent reaches final proficiency after 15 min, whereas the PC neoRL agent uses more than 160 min. In terms of learning efficiency, i.e., how fast the agent reaches final proficiency, and in terms of trained performance, the multi-modal neoRL agent performs better than both mono-modal neoRL agents.

4   Discussion

Contrary to RL in AI, neoRL navigation learns quicker, to higher proficiency, when more information is available to an agent. The neoRL agent is capable of multi-modal navigation, making multi-dimensional Euclidean navigation by a digital agent plausible.

Moving on from reinforcement learning and classical behaviorism, Tolman made a clear distinction between learning and performance after his latent learning experiments (see Fig. 1). Observing how an animal could learn facts about the world that could subsequently be used in a flexible manner, Tolman proposed what he called purposive behaviorism. When motivated by the promise of reward, the animal could utilize latent knowledge to form beneficial behavior toward that objective. Mechanisms underlying orientation have further been implied in cognition, a conceptual space where ideas are represented as points in a multi-dimensional Euclidean space. Technological advances have allowed new evidence

from modern neuroscience, supporting Tolman's hypotheses on cognitive maps' involvement in thought. Inferring that active navigation of such a space corresponds to reasoning and problem solving, we here propose autonomous navigation of conceptual space as an interesting new approach to artificial general intelligence. However, navigating conceptual space – with high dimensionality, an unknown form, and possibly an evolving number of Euclidean dimensions, is no trivial challenge for technology. Based on neural representation of space, the neoRL architecture is distributed and concurrent in learning, capable of separating between latent learning and purposive behavior, and a good candidate for emulated cognition by autonomous navigation of conceptual space.

Adopting Kaelbling's concerns for efficient robot learning to account for multi-modal navigation, we have methodically tested neoRL navigation in the WaterWorld environment. Firstly, it is crucial that neoRL navigation can operate in other Euclidean spaces than its primary navigation modality. Our first experiment verifies that the neoRL architecture is general across Euclidean spaces; a neoRL agent that navigates by the location modality is compared to one exposed to a relative-vector representation of external objects. Both NRES modalities perform admirably at this task, indicating that neoRL navigation is not restricted to one NRES modality. Secondly, we explore how neoRL navigation scales with additional NRES modalities; an agent based on both a place-cell and an object-vector-cell representation is compared to the two mono-modal neoRL agents from experiment 1. Navigation, both in training efficiency and final proficiency, improves significantly when more information is available to the agent. High-dimensional Euclidean navigation appears to be plausible with neoRL technology, formed by the basic principles from neuroscience and NRES.

In this work, we have collected evidence from theoretical neuroscience and the psychology of learning to propose a new direction toward emulated cognition. We have shown how online autonomous navigation is feasible by the neoRL architecture; still, the most interesting steps toward conceptual navigation in machines remain. What are the implications of autonomous navigation of conceptual space for AGI? Could latent spaces from deep networks be used for neoRL navigation? Should desires (elements of interest) propagate across NRES modalities based on associativity? Many important questions are yet to be asked. In showing that neoRL is up for the task of multi-modal navigation, we hereby propose a novel approach to AGI and present a plausible first step toward conceptual navigation in machines.

References

1. Behrens, T.E., et al.: What is a cognitive map? organizing knowledge for flexible behavior. Neuron **100**(2), 490–509 (2018)
2. Bellmund, J.L., Gärdenfors, P., Moser, E.I., Doeller, C.F.: Navigating cognition: spatial codes for human thinking. Science **362**(6415) (2018)
3. Bicanski, A., Burgess, N.: Neuronal vector coding in spatial cognition. Nat. Rev. Neurosci. **21**, 1–18 (2020)

4. Chaplin, J.P.: Systems and Theories of Psychology. Holt, Rinehart and Winston, New York (1961)

5. Constantinescu, A.O., O'Reilly, J.X., Behrens, T.E.: Organizing conceptual knowledge in humans with a gridlike code. Science **352**(6292), 1464–1468 (2016)

6. Eichenbaum, H.: Time cells in the hippocampus: a new dimension for mapping memories. Nat. Rev. Neurosci. **15**(11), 732–744 (2014)

7. Gärdenfors, P.: Conceptual Spaces: The Geometry of Thought. MIT press, Cambridge (2000)

8. Høydal, Ø.A., Skytøen, E.R., Andersson, S.O., Moser, M.B., Moser, E.I.: Object-vector coding in the medial entorhinal cortex. Nature **568**(7752), 400–404 (2019)

9. Kaelbling, L.P.: The foundation of efficient robot learning. Science **369**(6506), 915–916 (2020)

10. Kober, J., Bagnell, J.A., Peters, J.: Reinforcement learning in robotics: a survey. Int. J. Rob. Res. **32**(11), 1238–1274 (2013)

11. Kropff, E., Carmichael, J.E., Moser, M.B., Moser, E.I.: Speed cells in the medial entorhinal cortex. Nature **523**(7561), 419–424 (2015)

12. Leikanger, P.R.: Modular RL for real-time learning. In: The 3rd Conference on Cognitive and Computational Neuroscience (2019)

13. Leikanger, P.R.: Decomposing the prediction problem; autonomous navigation by neoRL agents. In: ALIFE 2021: The 2021 Conference on Artificial Life (2021)

14. Lever, C., Burton, S., Jeewajee, A., O'Keefe, J., Burgess, N.: Boundary vector cells in the subiculum of the hippocampal formation. J. Neurosci. **29**(31), 9771–9777 (2009)

15. Lewin, K.: Field theory and learning (1942)

16. O'Keefe, J., Dostrovsky, J.: The hippocampus as a spatial map: preliminary evidence from unit activity in the freely-moving rat. Brain Res. **34**, 171–175 (1971)

17. Schafer, M., Schiller, D.: Navigating social space. Neuron **100**(2), 476–489 (2018)

18. Solstad, T.: Neural representations of Euclidean space. PhD thesis, Kavli Insitute of Systems Neuroscience/Center of Neural Computation (2009)

19. Solstad, T., Boccara, C.N., Kropff, E., Moser, M.B., Moser, E.I.: Representation of geometric borders in the entorhinal cortex. Science **322**(5909), 1865–1868 (2008)

20. Sutton, R.S., Barto, A.G.: Reinforcement Learning: An Introduction. MIT press, Cambridge (2018)

21. Sutton, R.S., et al.: Horde: a scalable real-time architecture for learning knowledge from unsupervised sensorimotor interaction. In: The 10th International Conference on Autonomous Agents and Multiagent Systems, vol. 2, pp. 761–768. International Foundation for Autonomous Agents and Multiagent Systems (2011)

22. Tasfi, N.: Pygame learning environment. https://github.com/ntasfi/PyGame-Learning-Environment, Accessed 01 Sept 2020

23. Tasfi, N.: Waterworld in pygame learning environment. https://github.com/ntasfi/PyGame-Learning-Environment, Accessed 04 Apr 2021

24. Taube, J.S., Muller, R.U., Ranck, J.B.: Head-direction cells recorded from the postsubiculum in freely moving rats. i. description and quantitative analysis. J. Neurosci. **10**(2), 420–435 (1990)

25. Tolman, E.C.: Cognitive maps in rats and men. Psychol. Rev. **55**(4), 189 (1948)

26. Tolman, E.C., Honzik, C.H.: Degrees of hunger, reward and non-reward, and maze learning in rats. University of California Publications in Psychology (1930)