Navigating Conceptual Space; a New Take on Artificial General Intelligence

Per Roald Leikanger

September 7, 2021

Navigating Conceptual Space;
A new take on Artificial General Intelligence.

Per Roald Leikanger

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

Abstract. Edward C. Tolman revolutionized cognitive psychology by proposing a clear distinction between learning and performance in behaviorism. Tolman's ideas on latent learning and cognitive maps eventually led to what is now known as conceptual space; a paradigm where ideas are represented as locations in a high-dimensional Euclidean space. These insights are easily expanded to consider cognitive navigation between ideas – reasoning – as the basis of intelligence. Here, we explore whether conceptual navigation is plausible by the neoRL architecture, an RL architecture capable of having a distributed state representation as found in the hippocampus. Adopting Kaelbling's concerns for efficient robot learning to spatial navigation, we test whether neoRL is general across NRES 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 autonomous neoRL navigation of conceptual space as a plausible new path toward artificial general intelligence.

1 Introduction

The concept of a cognitive map was first introduced by Tolman as a model for explaining how spatial inferences and taking shortcuts [22]. Aiming to understand cognitive processes involved in behavior, Tolman was not satisfied with behaviorists' view that goals and purposes could be reduced to a hard-wired desire for reward [4]. Tolman observed that unrewarded rats could perform better when later *motivated* by reward [23]. 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*. Neural representation of Euclidean space (NRES) have later been identified by electrophysical measurements for a range of modalities [3]. Further, NRES has been implied for representing ideas as points in *conceptual space*, an Euclidean representation of ideas where betweenness and relative location makes sense when explaining concepts [7]. The involvement of NRES in cognition has been experimentally verified in an experiment where human subjects were asked to adapt an image to known concepts [5]. Results from theoretical neuroscience and psychology indicates NRES' role in social navigation [16], temporal representation [6], and reasoning [2]. This geometric view on cognition implies inference and generalization between ideas based on location; voluntary navigation of an emulated conceptual space could establish an interesting new take on artificial general intelligence (AGI).

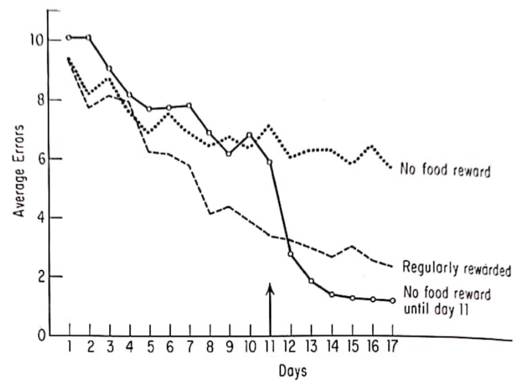

Fig. 1: Evidence for latent learning by Tolman and Honzik (1930).
(after [23], from *Systems and theories of psychology* [4]).

Autonomous navigation is difficult to reproduce in technology. Autonomous operation implies a decision *agent* capable of forming decisions based on own desires and experience. Experience-based behavior for digital decision algorithms is best approached by reinforcement learning (RL) from AI. Via trial and error with respect to a scalar reward signal $\mathbb{R}$, an RL decision *agent* is capable of adapting its behavior according to the accumulation of $\mathbb{R}$. A well-studied case for adaptive policies in Euclidean space is RL for robotics [10]. Considering robot path planning as an example of Euclidean navigation, we could look toward robot learning for inspiration on autonomous navigation. However; whereas deep RL 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 in 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). 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 in one NRES representation, the neoRL architecture makes a distinction between learning and behavior. Apparently inspired by Tolman, the neoRL framework allows for purposive behavior to form based on desires formed by anticipated reward [13]. However, navigating a multi-dimensional conceptual space of unknown dimensionality imposes high requirements on the agent. In this work, we adopt Kaelbling's three concerns for efficient robot learning to test whether neoRL navigation is plausible for a high-dimensional and multi-modal conceptual space.

2 Theory

Central to all navigation is knowledge of one's current navigational state. Information about one's relative location, orientation and heading, and for 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, it is referred to as being *allocentric*. In Euclidean geometry, vectors can be represented in 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 be represented in polar coordinates $\vec{a} = [r, \varphi]$, representing the point by the distance r and the allocentric direction φ. In RL, all relevant information must be included in the monolithic Markov state; each instance of agent state must contain enough data to uniquely define a next state distribution [18]. Combined with temporal dynamics, the number of such instances becomes prohibitively expensive for autonomous navigation [10]. A more distributed approach is found in neural orientation, i.e., representation of navigational state in the only known system capable of autonomous navigation. Neural vector coding in spatial cognition have been reported for a wide range of parameters [3]. Navigational state appears to be represented across a number of independent NRES structures for different NRES modalities. This section introduces theory and considerations on how state is represented in the animal and the learning machine, working slowly toward the internal mechanism of neoRL navigation.

2.1 Neural Representation of Euclidean Space

The first identified NRES neuron was the *place cell* [15]; O'Keefe and Dostrovsky discovered that specific neurons in the hippocampus became active whenever the animal traversed specific location in the test environment. Reflecting the current location of the animal, individual place cells can be considered as geometric *feature detectors* on the animal's allocentric location; the place cell becomes active when 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 [15], allocentric polar vector coordinates to external objects [8], and the one's current heading [21]. Additional NRES modalities are listed in Table 1 or in fig.

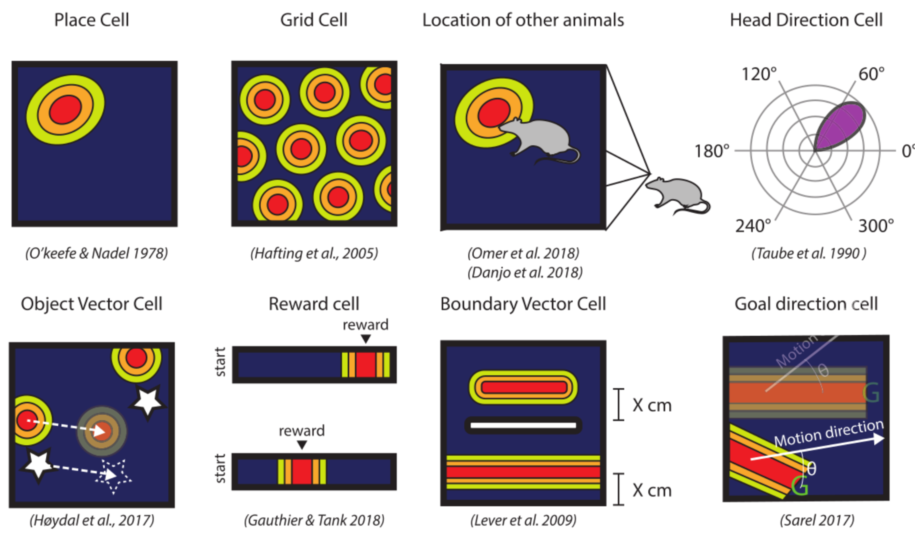

Fig. 2: Various identified NRES modalities of importance for navigation, with reference to the original publication. All NRES modalities could be important for autonomous spacial navigation. For the enclosed experiments, the reader only needs to understand the place cell and the object vector cell. (Illustration adopted from [1])

2. A comprehensive overview of NRES modalities identified in neurophysiology is composed by Bugress and Bicanski [3].

	Location	Tuning	Direction	NRES modality	
Place Cell	ac.	[proximal] 2D	-	Current position	[15]
Border Cell	ac.	[proximal] 2D	-	Location of borders	[17]
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	[21]
Speed Cell	-	[rate code] 1D	-	Current velocity	[11]

Table 1: Neural representation for different Euclidean spaces of importance for navigation: Head-direction cell reflects 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.

A population of NRES cells can map Euclidean coordinates to a representation suitable for neural computation. A simple mapping can be composed of NRES cells that respond to mutually exclusive parameter configurations. A group of NRES cells with non-overlapping receptive fields can be visualized as a chessboard; exactly one cell (tile) is satisfied for any parameter configuration on the board. This mapping is referred to as *one-hot encoding*[1] in computing sciences. The accuracy of this NRES map is defined by the map's resolution and the geometric coverage of the mapping; resolution is defined by the receptive fields of the involved neurons, whereas coverage is further defined by the number of NRES neurons involved in the mapping. In the remainder of this text, we only consider comprehensive one-hot encoding as illustrated in Figure 3. When discussing NRES resolution, we use the convention listed in the figure text, e.g., $N13$ signifying a $13x13$ tile set in $\Re^2$.

2.2 Autonomous Navigation by neoRL Agents

Navigation is characterized by two distinct features; the desired state — the objective with the interaction; and how to achieve this objective. When an operation is governed by own inclinations and *experience*, we refer to the agent's solution as an *autonomous* operation. A most accomplished approach to experience-based behavior is reinforcement learning; an RL decision *agent* is an algorithm 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 its 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 while learning in RL. This asymptotic progress is slow, requiring many interactions in RL. Although RL has proven effective for solving a range of 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].

With a set of sub-agents learning how to achieve different NRES cells in a single NRES representation, their experience can be combined by weighted sum [12]. Orthogonal value functions (OVFs) are learned as general value functions (GVFs) [19] with $\mathbb{R}$ defined by NRES cell activation. Via off-policy OVF learning with intrinsic reward and behavior formed by desire, Leikanger (2021) demonstrated how emulated NRES for agent state allows for autonomous navigation in Euclidean space [13]. NeoRL navigation improved significantly when agent value function was formed from multiple NRES resolutions. The neoRL value function is governed by the superposition principle, implying that value function could further be expanded across multiple NRES modalities. Combining experience across linearly independent parameters or NRES modalities remains to be tested; a multi-modal neoRL across NRES modalities would fully assimilate neural state representation in navigation – possibly acquiring some learning aspects from neural 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.

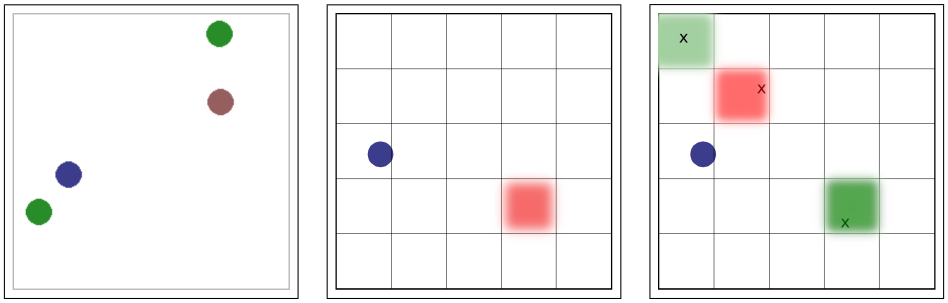

Fig. 3: (A) The allocentric WaterWorld environment: Blue entity is governed by inertia dynamics, with a desire for green ($\mathbb{R} = +1$) and aversion for red ($\mathbb{R} = -1$). (B) An $N5$ mapping of NRES: Each axis is divided into $N = 5$ equal intervals, resulting in $N^2 = 25$ NRES cells. An 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.

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 works across various 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 would make conceptual navigation infeasible. NeoRL navigation must be *general* across NRES modalities, *compositional* across conceptual components, without significant decline in learning *efficiency*. In this section, we explore neoRL capabilities for multi-modal navigation by experiments inspired by Kaelbling's concerns for efficient robot navigation.

All experiments are conducted in the allocentric version of the WaterWorld environment[2] [20], illustrated in fig. 3A. An agent controls the movement of the self (blue), with a set of actions that accelerate the object in the four directions N, S, E, W. Three objects of interest move freely in the closed Euclidean plane. When meeting with an object, it disappears, and a new object with a random color, location, and speed vector is initiated elsewhere. 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 good measure of an agent's navigation capabilities. Note how the agent must catch the last green in a board full of red before receiving more reward than an average

[2] The system that is interacted with by an RL agent is referred to as the *environment*.

of 1.5 points. All execution runs smoothly on a single CPU core, and the agent starts with no priors other than described in this section.

Observations from the WaterWorld environment could be considered as allocentric location coordinates. Thinking of the Euclidean plane in figure 3A as 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:

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

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 expressed through having different reference frames. All information is encoded in NRES maps as described in section 2.1. The neoRL agent is organized across multiple NRES maps of different resolutions as described in [13]. All multi-res agents are with resolutions covering all primes up to $N13$, i.e., with layers $N2, N3, N5, N7, N11, \text{and} N13$. For more on multi-resolution neoRL agents and the mechanism behind policy from parallel NRES state spaces, see [13]. Remember that the use of PC and OVC for WaterWorld is only syntactical to facilitate later discussion and how a 2D Euclidean plane could represent any 2D parameter configuration.

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, and the only received reward is $\mathbb{R} = +1$ when encountering green elements and $\mathbb{R} = -1$ when encountering red elements. The simple reward structure allows for direct measurement of how well the agent has learned during one run. However, observing the transient proficiency – real-time learning efficiency – of the agent requires some 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 the number of minutes since agent initiation. Proficiency is computed as the per time interval average of received reward, scaled to reflect $[\mathbb{R}/s]$. The y-axis can thus be interpreted as how many more green than red are captured per second.

3.1 NeoRL navigation: NRES generality

First, we test the generality of the neoRL architecture by comparing navigational proficiency for an agent exposed to the PC modality to one exposed to the OVC modality. We are interested in finding out whether neoRL navigation is generalizable across NRES modalities and how this would affect learning efficiency for the neoRL agent.

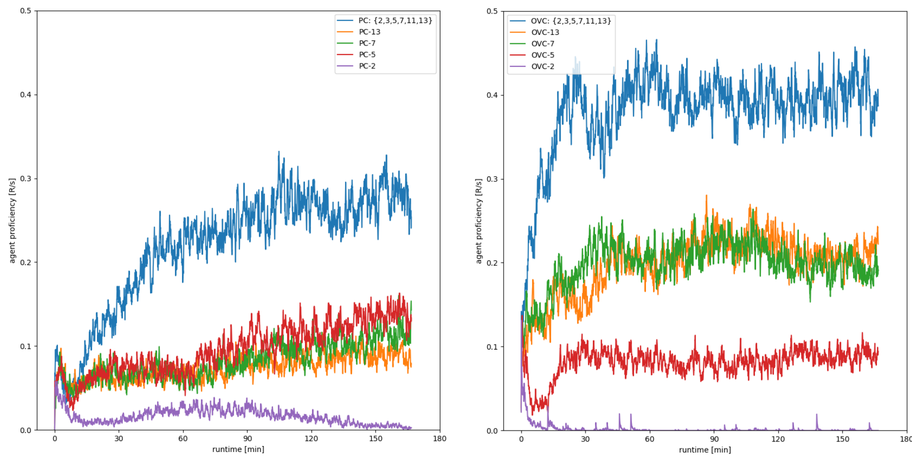

Fig. 4: The neoRL architecture is general across NRES modalities: (A) an original Place Cell (PC) NRES modality, implemented by applying NRES code directly on allocentric location of the agent or elements of interest. (B) an emulated Object Vector Cell (OVC) NRES modality, implemented by vector subtraction. The agent is centered with an allocentric representation of other objects – quite different from the PC modality.

Results are presented in figure 4: agent proficiency from the original PC modality (fig. 4A) can be compared with agent proficiency by the OVC modality (fig. 4B) The immediate proficiency of several mono-resolution neoRL agents are plotted alongside the proficiency of a multi-resolution neoRL agent. For both the PC and OVC modality, the multi-res neoRL agent performs significantly better when including information from multiple NRES maps with different resolution. There is no loss in sample efficiency when utilizing the OVC modality compared to PC modality; both the PC and the OVC neoRL agent instance perform well in the WaterWorld reactive navigation challenge. Results indicate that neoRL navigation is general for different aspects of experience; the neoRL architecture is flexible across experience modalities.

3.2 NeoRL navigation: NRES compositionality

Secondly, we are interested in how neoRL navigation scales with additional NRES modalities. Experiment 1 uncovered that neoRL agents are capable of reactive navigation by other information than allocentric location. The second experiment considers whether neoRL navigation is compositional across multiple NRES modalities; this experiment tests the effect of exposing one neoRL agent to both the PC and the OVF modality. We are concerned with how well the neoRL architecture scales with the additional information.

Results are presented in figure 5. Combining information across multiple NRES modalities significantly improves neoRL navigation compared to both

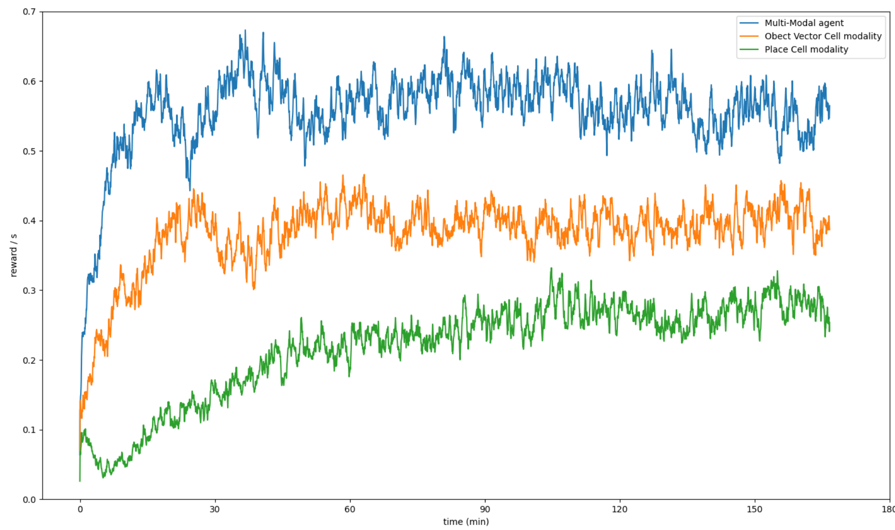

Fig. 5: Multi-modal neoRL navigation leads to higher proficiency and quicker learning than for mono-modal agents. These are important results toward the multi-dimensional navigation required for AGI by conceptual navigation.

mono-modal agents. The mono-modal performance of PC and OVC is shown in green and orange in figure 5. 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 mono-modal neoRL agents. The multi-modal neoRL agent reaches final proficiency after 10 minutes, whereas the PC neoRL agent uses 160 minutes. The final proficiency of the multi-modal neoRL agent approaches $0.55[\mathbb{R}/s]$ while the PC neoRL agent barely reaches $0.27[\mathbb{R}/s]$. NeoRL navigation can improve both learning speed and final proficiency when exposed to more information.

4 Discussion

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

Mechanisms underlying orientation have 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 to support Tolman's initial ideas on cognitive maps' involvement in thought. Inferring that active navigation of such a space corresponds to reasoning and problem solving, we have proposed autonomous navigation of conceptual space as a new take on artificial general intelligence. With a high dimensionality and unknown form, possibly with an evolving number of Euclidean dimensions, autonomous navigation becomes an interesting challenge. Adopting Kaelbling's three concerns for efficient robot learning to account for multi-modal navigation, we have considered NRES-oriented RL navigation for the task. Firstly, it is crucial that neoRL navigation works well for other Euclidean spaces than one's current location. The 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 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 place-cell and object-vector-cell representation NRES modality is compared to the two mono-modal neoRL agents. Navigation, both in terms of training efficiency and in terms of final proficiency, improves significantly when more information is available. It appears neoRL navigation (learning) improves with additional information instead of exploding with that cursed dimensionality, making multi-dimensional Euclidean navigation plausible for conceptual navigation.

Moving on from reinforcement learning and classical behaviorism, Tolman made a clear distinction between learning and performance based on results from his latent learning experiments (see Figure 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 a reward, the animal could utilize latent knowledge to form beneficial behavior toward that objective. A possible analogy could be how NRES-oriented RL allows for distinguishing the two aspects of an RL agent; Latent learning is expressed in the neoRL architecture through a set of off-policy OVF learners – learning the GVF for different NRES activation signals. Behavior is a result of the weighted sum over all OVF – scaled by reward expectancy. The resulting cognitive model is fully self-trained, reducing the challenge of autonomous navigation to being an online tuning of OVF weights. Purpose becomes an integral part of agent performance.

In this work, we have collected evidence from theoretical neuroscience and combined theory of learning with modern AI techniques to propose a new direction for AGI. We have shown how autonomous navigation is feasible by the neoRL architecture; Yet, the most interesting steps toward AGI by conceptual navigation remain. What are the implications of autonomous navigation of conceptual space? How is such a space affected by a changing environment? Would latent knowledge and adaptive conceptual space affect neoRL navigational performance? These and many more 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., Muller, T.H., Whittington, J.C., Mark, S., Baram, A.B., Stachenfeld, K.L., Kurth-Nelson, Z.: 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. Nature Reviews Neuroscience pp. 1–18 (2020)
4. Chaplin, J.P.: Systems and theories of psychology. Holt, Rinehart and Winston (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. Nature Reviews Neuroscience **15**(11), 732–744 (2014)
7. Gärdenfors, P.: Conceptual spaces: The geometry of thought. MIT press (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. The International Journal of Robotics Research **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. Cognitive and Computational Neuroscience (2019)
13. Leikanger, P.R.: Decomposing the prediction problem; autonomous navigation by neorl agents (2021)
14. Lever, C., Burton, S., Jeewajee, A., O'Keefe, J., Burgess, N.: Boundary vector cells in the subiculum of the hippocampal formation. Journal of Neuroscience **29**(31), 9771–9777 (2009)
15. O'Keefe, J., Dostrovsky, J.: The hippocampus as a spatial map: Preliminary evidence from unit activity in the freely-moving rat. Brain research (1971)
16. Schafer, M., Schiller, D.: Navigating social space. Neuron **100**(2), 476–489 (2018)
17. 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)
18. Sutton, R.S., Barto, A.G.: Reinforcement learning: An introduction. MIT press (2018)
19. Sutton, R.S., Modayil, J., Delp, M., Degris, T., Pilarski, P.M., White, A., Precup, D.: Horde: A scalable real-time architecture for learning knowledge from unsupervised sensorimotor interaction. In: The 10th International Conference on Autonomous Agents and Multiagent Systems-Volume 2. pp. 761–768. International Foundation for Autonomous Agents and Multiagent Systems (2011)
20. Tasfi, N.: Waterworld in pygame learning environment. https://github.com/ntasfi/PyGame-Learning-Environment, accessed: 2021-04-04
21. 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. Journal of Neuroscience **10**(2), 420–435 (1990)

22. Tolman, E.C.: Cognitive maps in rats and men. Psychological review **55**(4), 189 (1948)
23. 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)