UC Merced

Proceedings of the Annual Meeting of the Cognitive Science Society

Title

Quick and Accurate Affordance Learning

Permalink

Journal

Authors

Scholz, Fedor
Ayari, Erik
Bertram, Johannes
et al.

Publication Date

Quick and Accurate Affordance Learning

Fedor Scholz (fedor.scholz@uni-tuebingen.de),
Erik Ayari (erik.ayari@student.uni-tuebingen.de),
Johannes Bertram (johannes.bertram@student.uni-tuebingen.de),
Martin V. Butz (martin.butz@uni-tuebingen.de)
Department of Computer Science and Department of Psychology, University of Tübingen
Sand 14, 72076 Tübingen, Germany

Abstract

Infants learn actively in their environments, shaping their own learning curricula. They learn about their environments' affordances, that is, how local circumstances determine how their behavior can affect the environment. Here we model this type of behavior by means of a deep learning architecture. The architecture mediates between global cognitive map exploration and local affordance learning. Inference processes actively move the simulated agent towards regions where they expect affordance-related knowledge gain. We contrast three measures of uncertainty to guide this exploration: predicted uncertainty of a model, standard deviation between the means of several models (SD), and the Jensen-Shannon Divergence (JSD) between several models. We show that the first measure gets fooled by aleatoric uncertainty inherent in the environment, while the two other measures focus learning on epistemic uncertainty. JSD exhibits the most balanced exploration strategy. From a computational perspective, our model suggests three key ingredients for coordinating the active generation of learning curricula: (1) Navigation behavior needs to be coordinated with local motor behavior for enabling active affordance learning. (2) Affordances need to be encoded locally for acquiring generalized knowledge. (3) Effective active affordance learning mechanisms should use density comparison techniques for estimating expected knowledge gain. Future work may seek collaborations with developmental psychology to model active play in children in more realistic scenarios.

Keywords: affordances; exploration; epistemic; aleatoric; uncertainty; neural networks

Introduction

Humans learn internal models of their environment in order to interact with it in a flexible, adaptive, and context-dependent manner (Butz & Kutter, 2016). Which models are suitable at a certain point in time depends on the current state of the environment: In order to be able to prepare a cup of tea, both a cup and tea must be available and within reach (Kuperberg, 2021). This observation is captured by the psychological concept of affordances as introduced by Gibson: affordances encode which behaviors are possible in a given world state (Gibson, 1986). While it may be said that, on the level of motor commands, any bodily action is executable at any point in time, its outcome clearly depends on the current context. For example, in the absence of a cup and tea, 'preparing a cup of tea' actions will at best result in pantomime. Therefore, we define affordances—slightly more general than Gibson (1986)—as any factors in the environment that locally influence the outcome or success of an agent's actions, that is, that afford particular interactions and prohibit others.

But what if an agent wants to learn how to prepare a cup of tea in a situation where neither is in reach? In order to actively search for a location where critical preconditions are met, an allocentric map that relates coordinates to locally available affordances is needed. Such a map enables an agent to search in allocentric space where to satisfy its personal motivations. While learning, it can use the map to actively explore affordances by maximizing expected information gain. By focusing on aspects that can be learned and disregarding aspects that cannot be learned, we emulate curiosity and boredom, and thereby let the agent create its own learning curriculum (Smith, Jayaraman, Clerkin, & Yu, 2018).

Neuroscientific evidence suggests that the brain indeed learns such cognitive maps. It was shown that such maps enable not only the navigation in allocentric, world-centered spaces but also the instantiation of local circumstances for reasoning and planning as well as for reflecting on the past and imagining potential futures (Bottini & Doeller, 2020; Buckner & Carroll, 2007; Tolman, 1948; O'Keefe & Nadel, 1978). To date it remains unclear, though, how such dual-use maps may be learned. Here we assume the availability of an allocentric cognitive map for navigation and for providing sensory cues about local circumstances. We fully focus on studying how to navigate the environment to effectively learn about the affordances the environment offers.

We showcase our reasoning in an artificial world with a simulated agent. The agent perceives its environment via sensors and is able to imagine navigating it utilizing the provided cognitive map. The environment it lives in is confined by borders and contains terrains that influence the behavioral dynamics of the agent in distinct manners: obstacles block the passage; force fields accelerate the agent in a certain direction; fog fields corrupt the sensory signals with noise, mimicking an aleatoric uncertainty region. Our study shows how navigation behavior may target affordance-respective knowledge gain, that affordances should be encoded egocentrically, and that expected knowledge gain may best be computed by means of information-theoretic belief density comparisons.

Background

The problem setting we are concerned with can be described as a Markov Decision Process (MDP), where an agent receives observations from an environment and, based on that, executes actions that presumably lead to a desired state.

To this end, world models that encode visual information for planning were introduced before. Ha and Schmidhuber trained a vision model to produce codes that aid a controller in action selection (Ha & Schmidhuber, 2018). Since their vision model was trained as an autoencoder, though, it did not specifically produce codes that facilitate the controller's performance. Therefore, we do not regard the outputs of their vision model as *affordance* codes: They do not necessarily extract behavior-relevant information, but are produced to reconstruct the visual input. Qi, Mullapudi, Gupta, and Ramanan went one step further by training a neural network to encode behavior-relevant information (Qi et al., 2020). An agent was put into an environment to gather information about regions of harm and no harm. The experiences were backpropagated onto the input of the visual system, producing affordance maps. Subsequently, they trained a convolutional neural network to generate these maps from the visual input. Therefore, the architecture was not trained in an end-to-end fashion. As a result, the codes produced by the neural network were not optimized for behavioral control, which was anyway performed by a hard-coded A* algorithm.

In contrast to these studies, our work focuses on learning and exploration of a mapping from positions to affordances while a fixed, allocentric world model is provided, namely the world itself. This is in line with Epstein, Aroor, Evanusa, Sklar, and Parsons (2015), where the authors present an architecture that learns actual spatial affordances for navigation (Epstein et al., 2015). Their space of affordances, however, is limited to three predefined affordances specifically designed for navigation. Similarly, their action selection algorithm is based on handcrafted advisors and does not plan into the future.

Affordance architecture

The first end-to-end trained affordance architecture that produces codes that are explicitly optimized for behavioral control was introduced in Scholz, Gumbsch, Otte, and Butz (2022). In this case, the world model consists of a look-up map ω, an affordance model a_M, and a transition model t_M (see Figure 1). Given a position $\mathbf{p}_t$ in time step t, the hard-coded look-up map ω produces a sensory representation $\mathbf{v}_t$ of the environment at that position. The affordance model is a convolutional neural network (CNN) that computes a context code $\mathbf{c}_t$ based on $\mathbf{v}_t$. The transition model t_M—a multi-layer perceptron (MLP)—utilizes $\mathbf{c}_t$ as an additional input to predict the parameters of a probability distribution over positional changes $(\mu_{\Delta\tilde{\mathbf{p}}^{t+1}}, \sigma_{\Delta\tilde{\mathbf{p}}^{t+1}})$ given the last change in position $\Delta\mathbf{p}^t$ and the executed action $\mathbf{a}_t$.

Given sequences of position and action pairs, the model is trained end-to-end via backpropagation through time. The loss is given by the negative log-likelihood of the observed change in position $\Delta\mathbf{p}^{t+1}$ in the predicted distribution. This way, the affordance model a_M tends to produce context codes $\mathbf{c}_t$ that facilitate accurate predictions by the transition model t_M. The motor commands for data generation were selected randomly with a bias towards maintaining the same motor command for a few time steps. We select a bias that ensures a comprehensive coverage of the environment, emulating an approximately uniformly distributed exploration with regards to positions.

Given a state and a policy, i.e., a sequence of actions, the world model enables the agent to imagine how the interaction with its environment unfolds over time. In order to predict multiple time steps into the future, the predicted mean of the positional change is fed back into the model as the observed change in position. The agent then probes the look-up map ω with the anticipated position, allowing it to probe the environment for local sensory representations. We employ the cross-entropy method (Rubinstein, 1999), an evolutionary optimization algorithm, to infer the behavior that is expected to maximize reward and perform goal-directed control.

Exploration

The model was trained on previously generated sequences of observation and action pairs. As mentioned before, a heuristic was used that led to sequences which covered the whole environment. During the development of this heuristic it became apparent that it's exact implementation heavily influences the model's final performance and how fast it was able to learn. Suboptimal heuristics lead to the agent getting stuck in corners, thereby neglecting other parts of the environment.

We conclude that it would be helpful for the agent to actively explore its environment based on estimates of potential information gain. The agent should realize where its model is not able to generate accurate predictions and should thus explore those areas to improve its knowledge. The affordance maps from above allow it to plan considering environmental circumstances, thus enabling it to focus on affordance learning. To the best of our knowledge, such affordance-driven learning has not been explored before.

Active exploration can be guided by uncertainty, because reducing uncertainty translates into a more accurate world model. The mechanism should choose actions that produce high uncertainty in order to learn their effects in the current context and reduce uncertainty in the long run. Usually, not all uncertainty can be reduced though. As is often done in the machine learning community, we distinguish between two kinds of uncertainty: epistemic and aleatoric uncertainty (Der Kiureghian & Ditlevsen, 2009; Hüllermeier & Waegeman, 2021). Epistemic uncertainty is inherent in the model. It arises due to incompleteness or inaccuracy and can often be reduced by learning. In contrast, aleatoric uncertainty is inherent in the environment and cannot be reduced by learning. An example is the casting of a die, the outcome of which is practically unpredictable. To learn affordances quickly, the exploration mechanism should disregard aleatoric uncertainty, as there is nothing to be learned from it. Instead, it should choose actions that lead to high epistemic uncertainty. In this way, the agent will choose actions that have the highest learning potential. It is thus necessary to distinguish between aleatoric and epistemic uncertainties, to enable the active exploration of epistemic uncertainty while
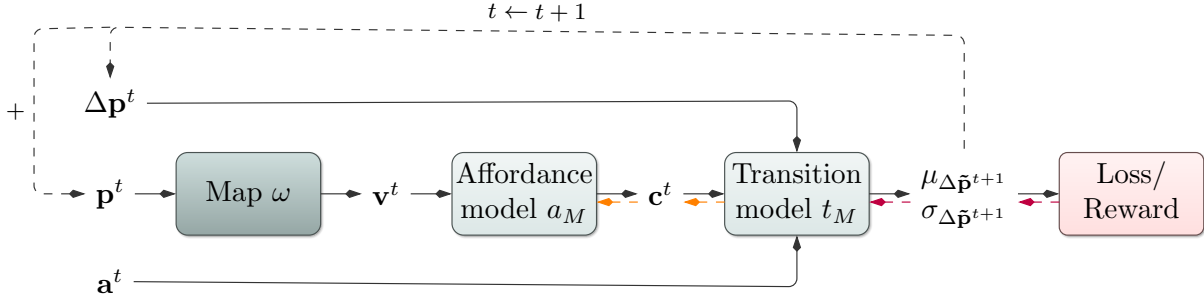
Figure 1: The overall architecture. The look-up map ω provides visual representations v_t of the environment at positions p_t. The affordance model a_M translates these representations into context codes c_t, which are utilized by the transition model t_M for the generation of predictions in the form of expected positional changes $(\mu_{\Delta \tilde{p}^{t+1}}, \sigma_{\Delta \tilde{p}^{t+1}})$. During training, the negative log-likelihood loss between predicted and observed Δp^{t+1} observations is backpropagated to t_M (red arrows) and further to a_M (orange arrows), training both subcomponents end-to-end. During control, potential behavioral interactions are evaluated via a reward function that combines estimates of epistemic knowledge gain with estimates of goal state proximity. In this paper we fully focus on affordance learning and thus on epistemic knowledge gain.

avoiding aleatoric uncertainty (Vlastelica, Blaes, Pinneri, & Martius, 2021).

Methods

The environment of our MDP is a 2-dimensional physics-based simulation. The agent is represented by a circular, inert vehicle that is able to glide around by sending motor commands to its four rocket jets, which cause accelerations in four diagonal directions. Observations consist of the last change in position $\Delta \mathbf{p}_t$ and a visual representation of the environment $\mathbf{v}_t$ at the currently considered position $\mathbf{p}_t$. Actions determine to which extent the four jets are activated. Their activities directly translate into accelerations.

In Scholz et al., 2022, the visual representation was given by a low-resolution image centered around the given position. Due to the discrete nature of pixels, this approach introduced uncertainty: the agent was not able to know exactly where, e.g., the borders of the environment were. Therefore, in this work, the visual representation is given by distances to surrounding entities in eight directions. Accordingly, we replace the CNN in the affordance model with an MLP.

All hyperparameters were optimized empirically, which led to the following configuration. The affordance model a_M is given by two linear layers with hidden sizes 64 and 32 and followed by ReLU activation functions. Its output, representing affordance codes, is produced by a linear layer that maps onto size 5 with the tanh activation function. The transition model t_M consists of a linear layer that maps onto size 64, followed by a ReLU activation function, followed by two parallel linear layers, one for predicting $\mu_{\Delta \tilde{p}^{t+1}}$ without an activation function and one for predicting $\sigma_{\Delta \tilde{p}^{t+1}}$ with the exponential activation function. We use Adam as our optimizer (Kingma & Ba, 2014). For each experiment, we train 5 different model instances based on different weight initializations.

The environment is confined by borders and contains ob-

stacles, both of which block the way. Other terrains in the environment locally alter the sensorimotor dynamics of the vehicle. Force fields accelerate the agent to the left or to the right and fog fields corrupt the observed position by Gaussian noise. Borders, obstacles, and force fields produce affordances that can be learned by the agent, resulting in epistemic uncertainty until they are learned. The uncertainty produced by fog fields cannot be reduced by learning and is therefore an instance of aleatoric uncertainty. We use two different environments in our experiments, one for training and one for validation (see Figure 2).
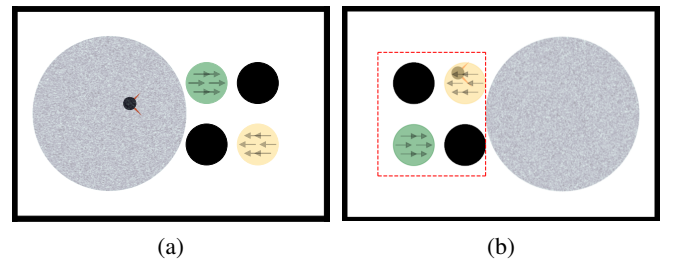

Figure 2: Environments used in our experiments. A small circular agent (black) navigates its environment using four diagonally attached rocket jets (orange). Fog fields are depicted in gray, obstacles in black. Force fields accelerate the agent to the right and left in green and yellow, respectively; (a) depicts the environment used during training and (b) depicts the environment used during validation. When focusing on affordance learning, we evaluate the model's performance only while the agent is within the red rectangle.

Uncertainty estimation

As shown in Figure 1, a single model instance predicts the parameters of a normal distribution $P = \mathcal{N}(\mu, \sigma^2)$ over positional changes. This way, the model performs uncertainty

estimation. However, it cannot distinguish between epistemic and aleatoric uncertainty. In order to be able to efficiently guide exploration towards learnable interactions, the two types need to be disentangled. We achieve this with ensembles of i models that predict probability distributions P_i. Since perfect models produce matching predictions, a measurement of an ensemble's members' disagreement can act as a proxy for epistemic uncertainty. We investigate the following uncertainty measures:

- Aleatoric uncertainty as the mean of the predicted standard deviations $AU = \mu(\sigma)$

- Epistemic uncertainty as the standard deviation of the predicted means $EU_{SD} = \sigma(\mu)$

- Epistemic uncertainty as the Jensen-Shannon divergence between predicted distributions $EU_{JSD} = \frac{1}{n} \sum_i D_{KL}(P_i \| M)$, where $M = \frac{1}{n} \sum_i P_i$ and D_{KL} denotes the Kullback-Leibler divergence

Using one of these uncertainty measures as the objective function during action selection allows the agent to perform exploration. The difference between EU_{SD} and EU_{JSD} is that the former does not consider the disagreement in predicted standard deviations. In contrast, JSD takes into account full distributions by extending the Kullback-Leibler divergence to potentially more than two distributions. The JSD possesses additional advantageous properties, namely being bounded and symmetric for all distributions (Briët & Harremoës, 2009). Therefore, we anticipate EU_{JSD} to provide the most precise measurements of epistemic uncertainty, leading to faster and more accurate affordance learning.

The training data is initialized with randomly generated sequences of observation and action pairs as in Scholz et al. (2022). If an uncertainty-based exploration mechanism is employed, we gradually replace a subset of the training data after each epoch by new sequences. These new sequences are generated by the agent itself via goal-directed control, based on behavior that is expected to maximize one of the above uncertainty measures. Validation is performed on a dataset that is generated with the same heuristic as the training data is initialized with.

Experiments

We first investigate how globally vs locally informative sensory information results in different generalization capabilities. Subsequently, we compare the different uncertainty measures with regards to their suitability for affordance learning.

Globally vs locally informative sensory information

The affordance model a_M allows the agent to perceive its environment. As input it receives distances to the nearest obstacle or terrain in each of eight directions. In our initial experiment, we compare an agent's affordance learning capabilities
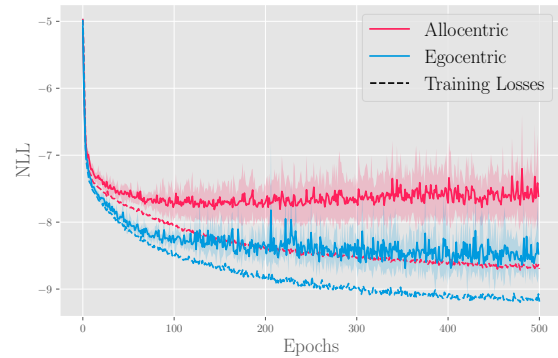

Figure 3: Losses for globally vs locally informative sensory information aggregated over 5 seeds. Shaded areas indicate standard deviations. Solid lines represent validation losses, dashed lines show training losses. The agent performs significantly better in the validation environment if equipped with distance sensors that are limited in range. With distance sensors that are not limited, slight overfitting is observed.

with globally informative sensory information vs locally informative sensory information. In the latter case, the distance sensors are limited in range, such that the agent is not able to perceive obstacles or terrains that are further away than a certain threshold. We choose the threshold such that the agent is always able to perceive obstacles or terrains that could influence its dynamics in the next time step. Globally informative sensory information is generated by sensors that are not limited in range. Here, the perceived distances to the borders encode the current position of the agent in the environment. For both bases, five model instances with differently initialized weights are trained on randomly generated sequences without any exploration taking place.

Results We find that the model generalizes significantly better to the validation environment if equipped with locally informative sensory information (see Figure 3). Globally informative sensory information allows the agent to learn the environment "by heart", relating absolute coordinates to affordances, and thereby harming generalization capabilities: The trained agent expects the obstacles and terrains to always be at certain absolute positions, an instance of overfitting. With locally informative sensory information, however, the agent is able to learn only egocentrically encoded knowledge. This kind of knowledge is applicable anywhere in the environment where the learned egocentric code applies. This is the case in our validation environment, where the obstacles and terrains have the same effects on the agent but are positioned differently in the environment.

In order to gain an understanding of the model's inner workings, we visualize the produced affordance codes as maps. To do so, we probe the environment at regularly distributed positions and feed the corresponding visual representations $\mathbf{v}_t$ into the affordance model. A principal component
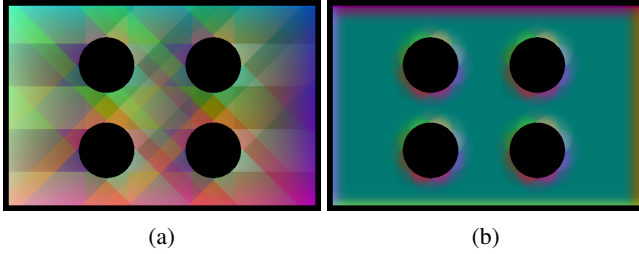
(a)   (b)

Figure 4: Affordance maps of an environment with four obstacles generated by a model with (a) globally informative sensory information vs (b) locally informative sensory information. The maps are produced by feeding visual representations of the environment at regularly distributed positions into the affordance model and mapping the produced context code onto RGB space via principal component analysis. True affordance maps, i.e., mappings from perceptual information to local behavioral constraints, only emerge in the latter case. Note how the obstacle's edges present the same local constraints as the borders, thus the matching borders show identical colors.

analysis reduces the dimensionality from 5 to 3, allowing us to visualize the encoded affordances by RGB values. These *affordance maps* represent local behavioral possibilities, such as whether it is possible to move to the right. Figure 4 shows the effects of the large-range distance sensors in a test environment with four obstacles: While globally informative sensory information yields distorted affordance maps, the more local sensory signals indicate great generalization abilities.

We have thus shown that it is rather advantageous to encode affordances in a local, egocentric manner. Therefore we use locally informative sensory information in the following experiments.

Uncertainty-guided exploration

We now investigate how affordances can be explored more efficiently, focusing on locally informative sensory information only. Here, the agent replaces the 5% oldest training sequences with new sequences that are generated by the exploration mechanism in each epoch. We compare the different uncertainty measures $AU, EU_{SD},$ and EU_{JSD}. The uncertainty measures are used as the objective that is to be maximized by the behavior inference mechanism. We always compare performances to the *random* behavioral policy, where the agent is not able to explore its environment in an active manner but is trained on a static training set. We evaluate each condition using 5 different seeds, with each seed generating an ensemble of size 5.

Results First, we examine the velocities the agent exhibits for the different cases defined as the distance traveled between two consecutive time steps (see Figure 5). We find that the agent produces significantly higher velocities if no active exploration takes place, i.e., in the *random* condition. This poses a disadvantage for the other cases as high velocities
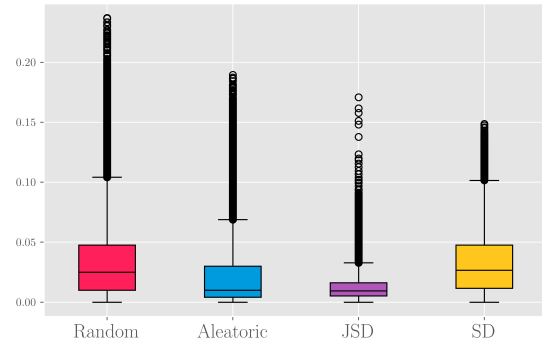

Figure 5: Boxplots of the velocities the agent exhibits during training with the different exploration mechanisms, taken over all epochs across the entire environment. No uncertainty estimate produces velocities as high as the *random* heuristic.
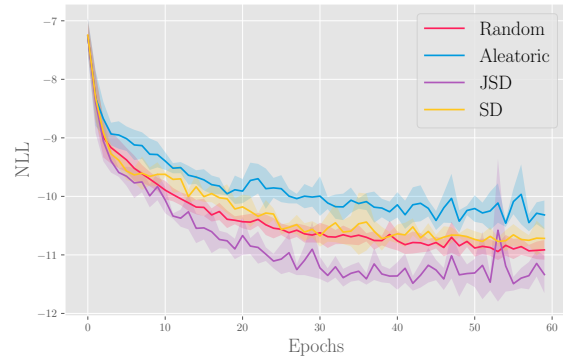

Figure 6: Losses for uncertainty-guided exploration aggregated over 5 seeds in the adapted, affordance-focussed validation set. Shaded areas indicate standard deviations. The best performance is achieved if the agent explores its environment based on the EU_{JSD} uncertainty measure. It also allows the agent to learn affordances the quickest.

that would be present in the validation data are never encountered during training. We therefore modify the validation set by adjusting the heuristic to produce lower velocities. Further, we focus on learnable affordances rather than on areas with high sensory uncertainty by restricting the validation set to data points where the agent is within the red rectangle in Figure 2.

The validation loss on the modified and restricted validation set is shown in Figure 6. An agent that focuses on aleatoric uncertainty during exploration indeed performs worst when confronted with the learnable affordances. EU_{JSD}-based behavioral inference learns the fastest and produces the lowest validation loss overall.

We also monitor the agent's positions and generate heatmaps to see which parts of the environment the agent explores the most with the different uncertainty estimates (see

Figure 7). Training based on sequences generated by the random heuristic results in relatively uniform coverage of the environment by design. We find that, of all approaches, exploration based on aleatoric uncertainty leads to the agent exploring the fog field the most. With EU_{SD}, the agent is more focused on other affordances than the fog, partially avoiding regions of high aleatoric uncertainty. The most profound effect, however, can be seen when EU_{JSD} is employed: the agent avoids the fog field to the greatest extent while exploring the rest of the environment relatively uniformly. We conclude that EU_{JSD} is the most suitable uncertainty measures for quick and accurate affordance learning.

Discussion

This study investigated how an agent can learn affordances in a quick and accurate manner. An artificial agent was put into a simulated environment and equipped with a cognitive map, which maps positions to sensory signals. The agent learned a predictive world model in the form of a neural network, predicting action consequences conditioned on local sensory information. The agent explored its environment and learned about affordances that inform it about environmental aspects that locally influence its behavior. Our results indicate that cognitive agents that actively learn about affordances should integrate three key ingredients in this process. First, search for novel affordances should be pursued within world-centered cognitive maps, which allow the activation of local views at particular positions within the map. Second, the learning of affordances should focus on local, body-relative sensory encodings. Third, divergence measures between a small collection of model-predictive densities are best-suited to identify regions that support further model learning.

We found that locality of perception is an important ingredient to allow for generalization, which is in accordance with the literature (Epstein et al., 2015). If sensors are too globally informative, then local, generalizable affordances are hard to learn. This is similar to infants which are born with low visual acuity (Smith et al., 2018), which may indeed be helpful to learn about global outlines and otherwise focus on local visual information, such as faces, hands, and objects. Note that locality does not necessarily need to be in space, but could also be in time or respective other conceptual spaces (Gärdenfors, 2014).

The learned affordances can be depicted by relating positions in the environment to affordance codes. The resulting affordance maps are related to cognitive maps and are possibly related to what the hippocampus is partially doing. Cognitive maps in the hippocampus appear to be essential for pursuing successful navigation and other environment-centered tasks. Meanwhile, the strong interconnectivity with neocortical areas indicates that particular local codes in the hippocampus may trigger views onto the local surrounding (Mallot, Roehrich, & Hardiess, 2014; Röhrich, Hardiess, & Mallot, 2014). Moreover, the mapping between allocentric positions and egocentric perceptions is probably mediated be-

(a) Random heuristic   (b) Aleatoric AU

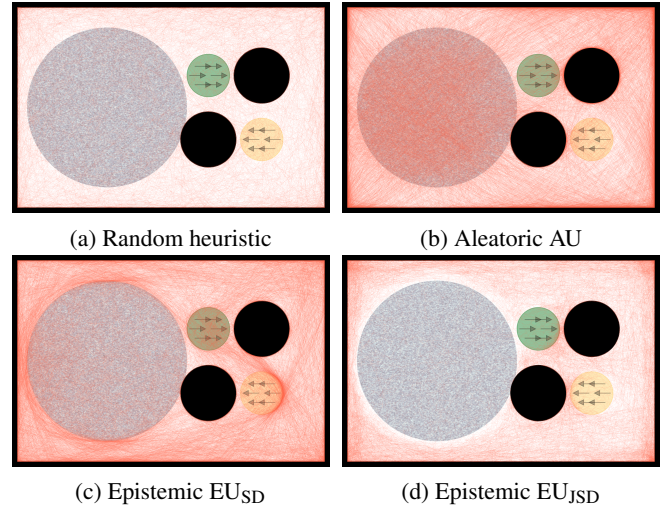(c) Epistemic EU_{SD} (d) Epistemic EU_{JSD}

Figure 7: Positional heatmaps based on all sequences the agent sees in the corresponding conditions during training. (a) The *random* heuristic sees all locations in an approximately uniformly distributed manner. As the training set does not change over the course of training, the heatmap is less dense. (b-d) When actively exploring the environment via the different uncertainty-based exploration strategies, the heat maps indicate most attractive sub-regions.

tween the hippocampal loop and the rest of the default mode network (Bottini & Doeller, 2020; Buckner, Andrews-Hanna, & Schacter, 2008; Zacks, 2020; Stawarczyk, Bezdek, & Zacks, 2021). It is this mapping that our agents explored and learned in this study.

In the future, we plan to extend our work to other problems, which do not necessarily need to be based on spatial navigation, such as when following a construction plan for building furniture from individual parts. Moreover, affordances should be directly based on objects, besides local circumstances, such as a kettle for boiling water for tea. With regards to developmental psychology, further explorations with respect to curriculum generation (Smith et al., 2018) might be interesting: How do children generate their curriculum in comparison to our approaches? How do they decide which experiences should never be forgotten, such as touching a hotplate? The way we replace sequences in the training set can certainly be improved. The challenge is to create a good set of training data and at the same time avoid forgetting of important experiences. Even though we did not observe catastrophic forgetting in the above experiments, it would be interesting to examine whether the learning process exhibits self-stabilizing behavior if disturbed by impactful experiences. In all of these cases, our study suggests that it will be advantageous to have local information available and to actively explore objects, entities, and locations in an epistemically-driven, active manner.

Acknowledgements

References

Bottini, R., & Doeller, C. F. (2020). Knowledge across reference frames: Cognitive maps and image spaces. *Trends in Cognitive Sciences*, *24*(8), 606-619. doi: 10.1016/j.tics.2020.05.008

Briët, J., & Harremoës, P. (2009). Properties of classical and quantum jensen-shannon divergence. *Physical Review A*, *79*(5), 052311. doi: 10.1103/PhysRevA.79.052311

Buckner, R. L., Andrews-Hanna, J. R., & Schacter, D. L. (2008). The brain's default network. *Annals of the New York Academy of Sciences*, *1124*(1), 1–38. doi: 10.1196/annals.1440.011

Buckner, R. L., & Carroll, D. C. (2007). Self-projection and the brain. *Trends in Cognitive Sciences*, *11*, 49-57.

Butz, M. V., & Kutter, E. F. (2016). *How the mind comes into being: Introducing cognitive science from a functional and computational perspective*. Oxford University Press.

Der Kiureghian, A., & Ditlevsen, O. (2009). Aleatory or epistemic? does it matter? *Structural safety*, *31*(2), 105–112.

Epstein, S. L., Aroor, A., Evanusa, M., Sklar, E. I., & Parsons, S. (2015). Learning spatial models for navigation. In *Spatial information theory: 12th international conference, cosit 2015, santa fe, nm, usa, october 12-16, 2015, proceedings 12* (pp. 403–425).

Gärdenfors, P. (2014). *The geometry of meaning: Semantics based on conceptual spaces*. Cambridge, MA: MIT Press.

Gibson, J. J. (1986). *The ecological approach to visual perception* (Vol. 1). Psychology Press New York.

Ha, D., & Schmidhuber, J. (2018, Mar). World models. *arXiv preprint arXiv:1803.10122*. doi: 10.5281/zenodo.1207631

Hüllermeier, E., & Waegeman, W. (2021). Aleatoric and epistemic uncertainty in machine learning: an introduction to concepts and methods. *Machine Learning*, *110*(3), 457–506. doi: 10.1007/s10994-021-05946-3

Kingma, D. P., & Ba, J. (2014). Adam: A method for stochastic optimization. *arXiv preprint arXiv:1412.6980*.

Kuperberg, G. R. (2021). Tea with milk? a hierarchical generative framework of sequential event comprehension. *Topics in Cognitive Science*, *13*(1), 256–298.

Mallot, H. A., Roehrich, W. G., & Hardiess, G. (2014). A view-based account of spatial working and long-term memories: Model and predictions. *Proceedings of the 36th Annual Conference of the Cognitive Science Society*, 1911-1916. (Austin, TX: Cognitive Science Society)

O'Keefe, J., & Nadel, L. (1978). *The hippocampus as a cognitive map*. Oxford, UK: Clarendon Press.

Qi, W., Mullapudi, R. T., Gupta, S., & Ramanan, D. (2020). Learning to move with affordance maps. *arXiv preprint arXiv:2001.02364*.

Rubinstein, R. (1999). The cross-entropy method for combinatorial and continuous optimization. *Methodology and computing in applied probability*, *1*(2), 127–190.

Röhrich, W. G., Hardiess, G., & Mallot, H. A. (2014, 11). View-based organization and interplay of spatial working and long-term memories. *PLOS ONE*, *9*(11), 1-12. doi: 10.1371/journal.pone.0112793

Scholz, F., Gumbsch, C., Otte, S., & Butz, M. V. (2022). Inference of affordances and active motor control in simulated agents. *Frontiers in Neurorobotics*. doi: 10.3389/fnbot.2022.881673

Smith, L. B., Jayaraman, S., Clerkin, E., & Yu, C. (2018). The developing infant creates a curriculum for statistical learning. *Trends in cognitive sciences*, *22*(4), 325–336.

Stawarczyk, D., Bezdek, M. A., & Zacks, J. M. (2021). Event representations and predictive processing: The role of the midline default network core. *Topics in Cognitive Science*, *13*, 164-186. doi: 10.1111/tops.12450

Tolman, E. C. (1948). Cognitive maps in rats and men. *Psychological Review*, *55*, 189-208.

Vlastelica, M., Blaes, S., Pinneri, C., & Martius, G. (2021). Risk-averse zero-order trajectory optimization. In *5th annual conference on robot learning.*

Zacks, J. M. (2020). Event perception and memory. *Annual Review of Psychology*, *71*(1), 165-191. doi: 10.1146/annurev-psych-010419-051101