

How sensory ecology affects the utility of planning

Ugurcan Mugan (umugan@u.northwestern.edu)

Northwestern University, Department of Biomedical Engineering, 2145 Sheridan Road
Evanston, IL 60208 USA

Malcolm A. MacIver (maciver@northwestern.edu)

Northwestern University, Department of Biomedical Engineering, Department of Mechanical Engineering
& Department of Neurobiology, 2145 Sheridan Road
Evanston, IL 60208 USA

Abstract

Prior to the vertebrate invasion of land, aquatic vision provided short range sensing with low contrast scenes. Once on land, aerial vision provided a 100-fold increase in range with high contrast scenes. This change in sensory ecology due to emergence onto land may have provided a selective advantage to those animals that were able to imagine alternative action sequences toward distant goals. To explore the relationship between sensory ecology and the utility of planning, we developed a simulation of predator-prey dynamics where we controlled visual range, planning depth, and environmental complexity. Simulations show that for prey with short visual range, increased planning results in a negligible change in survival rate with increased environmental complexity. However, at longer visual ranges, survival rate is strongly correlated with planning depth and environmental complexity, with peak survival rate occurring at high complexity and planning depth. These data suggest that planning is an adaptation to long range sensing enabled by terrestrial habitats 385 million years ago. Our results point to future research into the limitations on our temporal and spatial range of prospective cognition, a possible result of environments in which we have evolved, to raise awareness and create circumventions for looming existential threats.

Keywords: evolution; planning; partial observability; Markov decision processes; predator-prey interactions

Introduction

When a situation affords a long latency between stimulus and response, deliberative behavioral control can be used to generate behavior that is strategic, variable, and hard to predict by an adversary. As this latency decreases, reactive control takes over and generates responses that are fast, less variable, and easier to predict by an adversary (Catania, 2009). We can characterize deliberative behavioral control or planning as action choices that occur after internally simulating more than one action sequence and its respective consequences (Redish, 2016). Conversely, reactive control is a rapid stimulus-evoked response (Haggard, 2008; Jun, Longtin, & Maler, 2014). Our prior work on the vertebrate invasion of land and its sensory consequences revealed that just prior to moving onto land, eyes moved to the top of the head to



Figure 1: (A) Aquatic visual scenes have low complexity due to the high attenuation of light (MacIver et al., 2017). (B) Aerial visual scenes have higher complexity, with larger high frequency components (Balboa & Grzywacz, 2003). From Nilsson, 2017, a commentary on our MacIver et al., 2017 study.

look over the water surface and tripled in size. These morphological changes coupled with the higher transparency of air to light resulted in a nearly 100 fold increase in visual range over the aquatic condition (Fig. 1, MacIver, Schmitz, Mugan, Murphey, & Mobley, 2017). The increase in sensory range allowed animals to see distant potential drivers of behavior, such as predators or prey, and afforded long delays between stimulus and response.

The importance of the time-to-act metric can be understood when we consider dynamic targets, such as predators and prey. A fish with a sensory range of approximately two body lengths moving at a speed of one body length per second has little time after seeing a looming predator. As a testament to this diminutive spatiotemporal bubble, fish and amphibians feature a bilateral pair of dedicated giant fiber neurons in their brain, called Mauthner cells. This dedicated circuitry is responsible for creating fast, stereotypical responses to rapidly looming stimuli (Bhattacharyya, McLean, & MacIver, 2017). After amphibians, the Mauthner cell disappears among vertebrates, indicating that its advantage was lost once animals could see approaching threats at the much longer distances provided by aerial vision.

Increasing the time-to-act permits (but does not necessitate) the contemplation of multiple futures before acting. Such

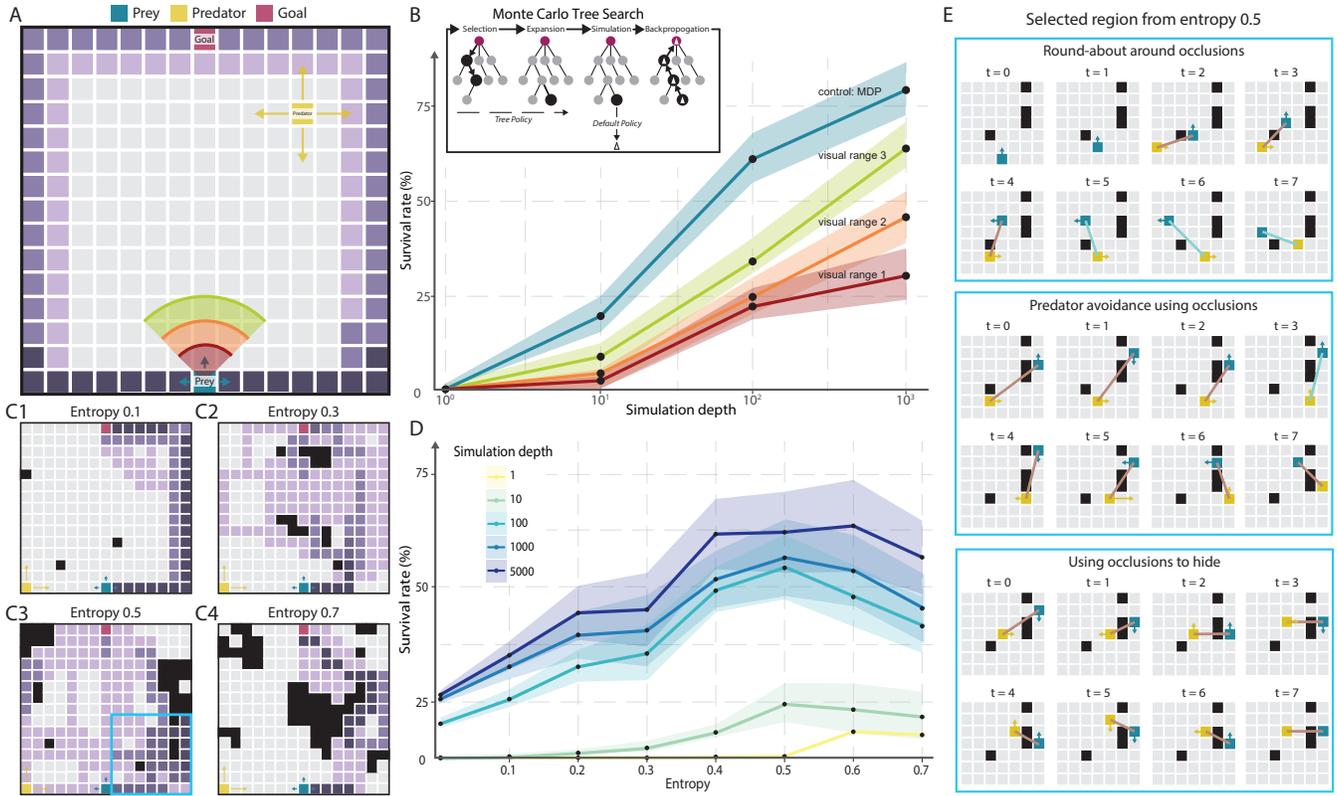


Figure 2: (A) Empty 15×15 gridworld with overlaid visual ranges tested (1, 2, or 3 grid cells ahead). The frequency of occupation heat map for the prey (purple) is the aggregation over trials including all possible predator initial locations. (B) Survival rate as a function of simulation depth for prey with different visual ranges acting in the empty gridworld shown in A. This is across randomly chosen predator locations ($n = 20$); fill shows SEM. Monte-Carlo tree search inset shows sequences of actions taken during 1 simulation depth (e.g. simulation depth of 100 is one hundred repeats of this process). Simulation depth 10^2 : survival rate for visual range 2 and 3 are significantly different (Mann-Whitney U test $p = 0.0041$) (MWU). Simulation depth 10^3 : survival rate for visual range 1 and 2 are significantly different (MWU, $p = 0.0086$) as it is for visual range 2 and 3 (MWU, $p = 0.00017$) (C) Examples of randomly generated gridworlds with varying entropy. Occlusions (black) act as walls that disrupt the predator’s line of sight. Frequency maps show paths taken by the prey that results in its survival for that specific world. (D) Survival rate as a function of entropy for different planning depths across random distributions of occlusions for a given entropy ($n = 20$ per depth); fill is SEM. (E) Three examples of behaviors seen in one environment with 0.5 entropy (zoom of blue rectangle in C3) during various trials. Brown line between predator and prey indicates that the predator cannot see the prey. Blue line between the predator and prey represents prey visibility. Videos of these three examples in addition to others is available online (Mugan & Maclver, 2018)

imagined trajectories should take into account environmental factors such as topology and vegetation, which can occlude vision of the threat or opportunity, and the vision by the threat or opportunity in the case of sentient targets. The effective usage of these occlusions is dependent on: 1) knowing the target location with respect to the occlusions, and 2) being able to view the occlusions (perceived environmental complexity). The underwater optics of the turbid ancestral aquatic habitats causes both short range vision (Maclver et al., 2017) and low contrast (Balboa & Grzywacz, 2003), making observations have little directional variability (Fig. 1). Conversely, the low attenuation of light through air enables long range vision that is rich, making it possible for the animal to observe occlusions, such as vegetations. Such occlusions provide a visual barrier for the animal, making certain trajectories more favorable during ei-

ther pursuit or escape.

Here, we test the hypothesis that the selective advantage of planning with respect to moving targets is greatly enhanced with increased sensory range and environmental complexity.

Methods

In order to study the relationship between planning depth, visual range, and environmental complexity, we used partially observable Markov decision processes (POMDPs). By using POMDPs we were able to combine Markov decision process models of animal decisions during planning (Miller, Botvinick, & Brody, 2017; Daw, Gershman, Seymour, Dayan, & Dolan, 2011) with observations and tree-like planning systems that rely on model-based methods. These assume a previously learned model of the environment (Sutton & Barto, 2018).

We created a novel survival task by implementing a virtual prey and virtual predator both acting in a 15×15 virtual grid-world (Fig. 2A). Within this framework we equipped the prey with a predetermined visual cone that extends outward either 1, 2, or 3 cells and faces the direction of motion. The aim of the prey is to get to the goal position while being aggressively pursued by the predator, which is on average $1.5 \times$ faster. The prey uses the Monte-Carlo planning algorithm (POMCP: (Silver & Veness, 2010)), which combines a sample-based approach to belief state update and to the tree of decisions the prey has at each state (move left, right, forward and backward). For each simulation depth the tree is expanded by exactly one node.

To vary environmental complexity, we added randomly generated distributions of occlusions, calculated based on entropy, to our gridworlds (Fig. 2C). These occlusions obstruct the predator's line of sight, hiding the prey if an occlusion exists on the ray between the predator and the prey (e.g., Fig. 2E). We can analyze the complexity of the gridworld by transforming it into a graph, and applying graph theory concepts. With such a transformation, by definition a world with no occlusions is a highly connected graph. On the other hand, a world that is mostly occluded (high entropy) is highly disconnected. These two extreme examples both have low complexity (Bonchev & Buck, 2005), making the peak in complexity occur at midrange occlusion levels (around entropy=0.5).

Results

For all of the randomly generated gridworlds, each trial consisted of assigning a visual range and planning depth to the prey, and randomly selecting a predator start location. For a low entropy, low complexity world, survival rate increases proportionate to how much planning we allow the prey to do before making its next move (Fig. 2B). With no planning, survival rate is zero regardless of range. When a prey with small sensory range observes a predator, due to the differences in their respective speed and the predator's policy of aggressive pursuit, the prey cannot escape. Increasing planning depth for such prey increases survival rate only minimally. Conversely, for prey with long sensory ranges, survival rate rapidly increases with respect to planning depth.

We can look at all trials in which the prey succeeded in reaching the goal point to compute how often it occupied each grid cell. While this varies with predator initial position, linked cells with high occupancy frequency represent "success paths." The distribution of success paths for low entropy, low complexity gridworlds reveals that there are only a few distinct trajectories that the agent can take to increase its low probability of survival (Fig. 2A, C1). In these worlds planning only provides the agent with a simple policy: 1) if the predator is sensed, go to the wall furthest away; if not, then go to the closest wall; 2) follow wall to the goal. (Interestingly, these wall-following or "thigmotaxis" paths are the same that are followed by rodents in open field tests (Simon, Dupuis, & Costentin, 1994).) Clearly, simply following this policy (rather than planning) would lead to a similar frequency of occupation heat

map (Fig. 2A), but we have reason to believe that it would also result in similar survival rates. We will be testing this hypothesis in the near future.

Within these low entropy worlds, observation distance becomes more important. If the prey can see the predator from further away it can change its direction earlier. Given these simple environment-induced dynamics, survival rate, independent of sensory range, is relatively low even at high planning depths. This implies that model-free reactive control may be more beneficial in such environments, especially when the prey sensory range is limited. Reactive control in this case would better utilize the resources that would otherwise be allocated to planning.

The previous portion of our study showed that high planning depths are only beneficial for animals with long visual ranges. Next, we examined the effects of planning depth on survival rate while environmental clutter (quantified by entropy) was varied, when visual range was kept long similar to the aerial vision condition. As expected, an increase in planning depth, independent of entropy, increases the survival rate. However, when planning depth is high, the survival rate increases until midrange values of entropy (0.4–0.6) and then decreases for higher entropy (> 0.6) (Fig. 2D). Considering the connectivity of these worlds, in light of the definition of network complexity mentioned above, these two ranges represent high and low complexity environments, respectively. Moreover, as planning depth increases, the incremental change in survival rate with respect to entropy also increases until entropy=0.5. These findings suggest that higher planning depths are required for the prey to strategically deploy occlusions to escape from the predator, such as through hiding or by engaging in diversionary tactics not unlike the broken-wing anti-predator tactic that birds use (examples of each are shown in Fig. 2E). While birds are diverting predators from fledglings, in cases like the "Round-about around occlusions" example of Fig. 2E, our agent can be interpreted as diverting the predator from cutting off potential escape paths.

Similar to our analysis of low entropy environments, we analyzed trajectory distributions across randomly generated worlds with different entropies. In contrast to gridworlds with low complexity, in gridworlds of high complexity, planning no longer cues one or two strategies that could be easily compiled to habit (such as "follow the wall"). The spread of the trajectories that lead to success in midrange entropy worlds—signaled by their diffuse occupancy frequency maps (Fig. 2C2, C3)—suggests that survival rate is highly dependent on the predator location and occlusion distribution. In these worlds, it becomes important for the agent to re-plan its trajectory after each step to take into account the changed environmental configuration (prey, predator and occlusion locations). At high planning depths this leads to the generation of complex behaviors (detailed above). It has been previously shown that planning enables animals to quickly update their policy (Daw, Niv, & Dayan, 2005). Our simulations indicate that the complexity of the terrestrial world greeting our fish ancestors, as

revealed by their vastly different sensory ecology, could be the origin of the need for such systems.

As entropy increases, the world becomes too occluded, which constricts the profusion of success trajectories at midrange entropy to one or two (Fig. 2C4). In these high entropy, low complexity environments, the importance of planning is diminished, possibly allowing for reactive strategies to succeed.

Conclusion

There is behavioral evidence for planning in birds (Clayton & Emery, 2015) and mammals (Redish, 2016). As yet, we are unaware of attempts to establish this capability in any vertebrates other than birds and mammals, although it is becoming clearer that other vertebrates have homologues of one of the key structures in planning, the hippocampus (Rodriguez et al., 2002; Elliott, Harvey-Girard, Giassi, & Maler, 2017). However, a recent survey of vertebrate brain size shows that the brain:body mass ratio increased by a factor of 10–40 times from fish/amphibians to birds/mammals (values for animals with body mass over 100 g) (Yu, Karbowski, Sachdev, & Feng, 2014). It seems likely that some portion of this size differential is due to the higher selective advantage of planning and consequent expansion of related neural structures known to be involved, such as the hippocampus and medial pre-frontal areas in mammals (Redish, 2016).

These early findings support the theory that a massive increase in visual range with the emergence of eyes above the water line diminished the need for reactive circuitry (such as the Mauthner system seen in fish and amphibians), and provided selective advantage to planning. If planning originated as an adaptive response to the change in sensory ecology with terrestriality, then our work opens up a new domain of research. In that domain, important questions would include how constriction on the temporal or spatial range of prospective cognition might be due to neural circuits carrying properties of our ancestral environment. There are a number of looming existential threats to our species, such as climate change, that may be partially attributable to spatiotemporal range limitations of human prospective cognition. Research on this problem could have practical utility in raising our awareness of these limitations and suggesting circumventions.

Acknowledgments

This work was funded by NSF IOS-ORG 1456830.

References

Balboa, R. M., & Grzywacz, N. M. (2003). Power spectra and distribution of contrasts of natural images from different habitats. *Vision Res.*, *43*(24), 2527–2537.

Bhattacharyya, K., McLean, D. L., & Maclver, M. A. (2017). Visual Threat Assessment and Reticulospinal Encoding of Calibrated Responses in Larval Zebrafish. *Curr. Biol.*, *27*(18), 2751–2762.

Bonchev, D., & Buck, G. A. (2005). Quantitative measures of network complexity. In *Complexity in chemistry, biology, and ecology* (pp. 191–235). Springer.

Catania, K. C. (2009). Tentacled snakes turn c-starts to their advantage and predict future prey behavior. *P. Natl. Acad. Sci. USA*, *106*(27), 11183–7.

Clayton, N. S., & Emery, N. J. (2015). Avian models for human cognitive neuroscience: A proposal [Journal Article]. *Neuron*, *86*(6), 1330–1342.

Daw, N. D., Gershman, S. J., Seymour, B., Dayan, P., & Dolan, R. J. (2011). Model-based influences on humans' choices and striatal prediction errors. *Neuron*, *69*(6), 1204–1215.

Daw, N. D., Niv, Y., & Dayan, P. (2005). Uncertainty-based competition between prefrontal and dorsolateral striatal systems for behavioral control. *Nat. Neurosci.*, *8*(12), 1704–1711.

Elliott, S. B., Harvey-Girard, E., Giassi, A. C., & Maler, L. (2017). Hippocampal-like circuitry in the pallium of an electric fish: Possible substrates for recursive pattern separation and completion. *J. Comp. Neurol.*, *525*(1), 8–46.

Haggard, P. (2008). Human volition: towards a neuroscience of will. *Nat. Rev. Neurosci.*, *9*(12), 934–946.

Jun, J. J., Longtin, A., & Maler, L. (2014). Enhanced sensory sampling precedes self-initiated locomotion in an electric fish. *J. Exp. Biol.*, *217*(20), 3615–3628.

Maclver, M. A., Schmitz, L., Mugan, U., Murphey, T. D., & Mobley, C. D. (2017). Massive increase in visual range preceded the origin of terrestrial vertebrates. *P. Natl. Acad. Sci. USA*, *114*(12), E2375–E2384. Video explainer: <https://youtu.be/I19usgWHJLc>.

Miller, K. J., Botvinick, M. M., & Brody, C. D. (2017). Dorsal hippocampus contributes to model-based planning. *Nat. Neurosci.*, *20*(9), 1269–1276.

Mugan, U., & Maclver, M. A. (2018). *NxR Website : Simulation videos of agents with visual range planning in gridworlds of varying complexities*. <http://nxr.northwestern.edu/planning-vertebrates>.

Nilsson, D. E. (2017). Evolution: An Irresistibly Clear View of Land. *Curr. Biol.*, *27*(14), R715–R717.

Redish, A. D. (2016). Vicarious trial and error. *Nat. Rev. Neurosci.*, *17*(3), 147–159. doi: 10.1038/nrn.2015.30

Rodriguez, F., Lopez, J. C., Vargas, J. P., Gomez, Y., Broglio, C., & Salas, C. (2002). Conservation of spatial memory function in the pallial forebrain of reptiles and ray-finned fishes. *J. Neurosci.*, *22*(7), 2894–2903.

Silver, D., & Veness, J. (2010). Monte-carlo planning in large POMDPs. In *Adv. Neural Inf. Process. Syst.* (pp. 2164–2172).

Simon, P., Dupuis, R., & Costentin, J. (1994). Thigmotaxis as an index of anxiety in mice. Influence of dopaminergic transmissions. *Behav. Brain Res.*, *61*(1), 59–64.

Sutton, R., & Barto, A. G. (2018, May). *Reinforcement learning: An introduction (second edition)*.

Yu, Y., Karbowski, J., Sachdev, R. N., & Feng, J. (2014). Effect of temperature and glia in brain size enlargement and origin of allometric body-brain size scaling in vertebrates. *BMC Evol. Biol.*, *14*(1), 178.