

Toward Behavior-Based models of bat echolocation

1st Thinh H. Nguyen

Electrical Engineering and Computer Science
University of Cincinnati
Cincinnati, USA
nguye2t7@mail.uc.edu

2nd Dieter Vanderelst

Electrical Engineering and Computer Science, Biological Sciences
University of Cincinnati
Cincinnati, USA
vanderdt@ucmail.uc.edu

Abstract—We propose a Behavior-Based Robotic (BBR) architecture to model the cognitive controller of echolocation bats. The architecture used a neural network to perform high-level control by governing two sensorimotor loops. We trained our model in a simulated environment where the echoes returned from the environment were derived from real echoes collected by a physical sonar system. We trained our BBR architecture on a foraging task and tested the trained agent in different experiments. The agent demonstrated the ability to learn the foraging task on different maze geometries by avoiding obstacles and approaching food items. The agent also demonstrated robustness against considerable noise in actuation. This prototype demonstrated the feasibility of training a BBR model of complex bat echolocation tasks using a hybrid simulated environment.

Index Terms—bat echolocation; modeling; sonar; cognitive architecture; Real2Sim; artificial ethology

I. INTRODUCTION

Research on echolocating bats has shown that bio-sonar is capable of supporting swift flight through dense vegetation [1], navigation in changing environments [2], [3], object recognition [4], and airborne foraging [5]. Much of the research into bat echolocation attempted to isolate the acoustic cues (i.e., acoustic properties of the echoes) that support different sonar-based tasks and confirm the sufficiency of those cues in behavioral experiments [6], [7], simulations [8]–[10], or robotic studies [11]–[14]. For instance, Greif and colleagues [7] demonstrated that bats find water by recognizing the acoustic cues returned by smooth, horizontal surfaces. Several sensorimotor loops exploiting acoustic cues in task-specific motor control have been proposed. For example, prey capturing strategies have been evaluated by [11], [12], [15], and us [16]. In addition, we have proposed and tested sensorimotor loops for obstacle avoidance [17], [18], mapping [3], and gleaning prey from leaves [19].

Despite the progress in unraveling how bats exploit acoustic cues to address specific sonar-based subtasks, little work has been done on modeling how simple sensorimotor loops could be integrated to achieve more complex behavior. In previous work [3], we hypothesized that complex sonar-based behavior observed in bats could be modeled by a cognitive architecture mimicking the Behaviour-Based Robotics (BBR) approach

[20]. In brief, this robotic control approach aims for an agent's complex behavior to emerge from the interaction between multiple sensorimotor loops, with each loop dedicated to a simple subtask [21].

A BBR-inspired cognitive architecture is promising for modeling complex bat behavior for at least two reasons. First, this architecture assumes that agents do not (or only minimally) rely on internal representations [21]. This characteristic is an attractive feature for models of bat sonar. Indeed, deriving internal world representations from sonar data is notoriously difficult and unreliable (see [17], [18] for arguments). Therefore, models that do not require internal representations of the environment are promising candidates for explaining how bats exploit multiple acoustic cues in dealing with their complex environments. The second advantage of the BBR approach is responsiveness [21]. The low speed of sound limits the update rate in biosonar. However, using fast sensorimotor loops directly coupling sensory input to actuation should allow bats to respond appropriately under hard real-time conditions independent of the complexity of the environment.

While a BBR-inspired architecture seems promising to explain complex sonar-based behavior, our hypothesis [3] achieved remains untested. Postulating a Behavior-Based model of bat sonar requires (1) identifying acoustic cues for subtasks, (2) formulating sensorimotor loops that exploit these cues, and (3) stipulating an arbitration mechanism for activating specific loops. As stated above, previous work has identified several cues and sensorimotor loops for addressing specific sonar-based subtasks. However, proposing a model that arbitrates between multiple sensorimotor loops is not straightforward. The acoustics of the natural world is complex, making it often impossible to, a priori, specify how acoustic cues will change as the bat moves through space. Therefore, postulating the conditions under which specific loops should take control is challenging. Machine learning could be an alternative to manually designing [22] an arbitration model. A disadvantage to using machine learning is that this requires a large number of training iterations. Therefore, using machine learning often requires relegation to simulation. However, this would require simulating the complex acoustics of the world, which is not trivial. Even modeling the echoes returned by single objects (as we have before in a series of papers) is computationally very demanding [23]–[26]. This makes it hard (if not impossible) to provide the simulated bat with realistic

Research supported by NSF grant 2034885, IOS Division Of Integrative Organismal Systems. All code and data collected for this work can be found at github.com/huythinhnguyen/SSCI_2022_bat_foraging or by contacting the 1st author for the latest version.

echoes from complex environments.

This paper suggests a (hybrid) method for designing a model of sonar-based behavior by combining dedicated sensorimotor loops. As such, this paper is an initial attempt to try and explain more complex sonar-based behavior by hypothesizing a behavior-based cognitive architecture. We used reinforcement learning to train the arbitration between sensorimotor loops in a hybrid simulated environment. To ensure the acoustic validity of our simulated environment, we used acoustic data collected using a physical sonar sensor to supply the simulator with a realistic acoustic model of a complex environment. Using data captured using physical transducers, we could overcome the high computational cost of rendering echoic data from object geometry.

As an initial test of our hypothesis that a Behavior-Based architecture can support more complex sonar-based behavior, we selected a foraging task requiring (1) avoiding obstacles and (2) approaching food items. This task also represents an initial test of our hybrid modeling approach.

II. METHODS

We tasked an agent (simulated bat) with moving around an arena constructed of vegetation-like obstacles and foraging for objects labeled as food. We postulated that this foraging task could be completed using two sensorimotor loops. The first sensorimotor loop (avoidance) steered the agent away from obstacles. The second loop (approach) made the agent approach objects, i.e., food items. We used reinforcement learning to train the agent to appropriately activate each of these loops based on the current echoic input.

The agent was equipped with a simulated sonar system modeled after bats' biosonar. The sonar system consisted of one ultrasonic emitter (mouth) and two microphones (ears). The simulated emitter corresponded to a narrowband (42 kHz) piezo ultrasonic emitter. In contrast to our simulated bat, most bats use frequency-modulated broadband emissions, particularly when operating in complex environments [27]. However, the acoustic properties of the simulated emitter were specified by how we collected echoes from physical objects, as described in section II-A. Using a narrowband emitter does not invalidate our demonstration as it provides our simulated bat with less information than real bats. As such, our bat can not exploit any information not available to real bats. The bat's ears' directionality was modeled based on the bat *Myotisotis microtis* (see Section II-A for details). The agent used a call rate of 50 Hz (which corresponds to bat call rates in complex environments, see [18] for references). The agent's controller used only the simulated echoes received by the two ears as sensory input.

A. Environment and echo simulation

We constructed a simulated environment to train and test the agent. We refer to this environment as the Echolocation-based Virtual Environment (EchoVR). As indicated, the echoes returned by this environment were derived from physical echoes collected from two reference objects. Many natural bat habitats

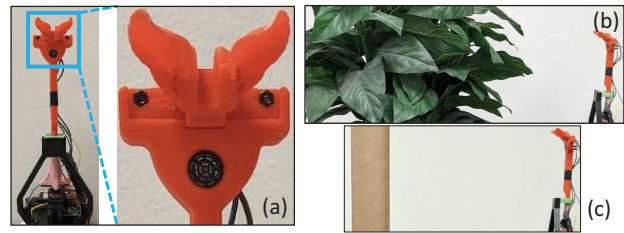


Fig. 1. (a) Custom sonar device. (b) Collect reference echoes of complex geometry object as obstacle: plant. (c) Collect reference echoes of simple geometry object as food: cardboard pole.

consist primarily of vegetation, with the complex geometry of unstructured reflectors returning many overlapping echoes for each call [28]. We used echoes from an artificial plant with densely packed leaves (the overall diameter of the plant is approximately 50 cm) to model the vegetation in the natural habitat. Echoes from simulated food items were based on echoes collected from a cardboard pole (10 cm in diameter). This object returned a single isolated echo, akin to simpler (than plants') echoes produced by airborne insect prey items [29].

To collect echoes from the artificial plant and the cardboard pole, we built a custom sonar device (see Figure 1) to collect reference echoes with a bat-like head-related transfer function (HRTF). Two omnidirectional microphones (Knowles, FG-23329-P07) were inserted into a pair of 3D-printed bat ears (*Myotisotis microtis* [23]). The ultrasonic emitter placed 2 cm below the ears produced a 1-2 ms pulse of 42 kHz. The microphones recorded echoes for both ears for a duration of 23 ms using a 300 kHz sampling rate (resulting in 7000 samples for each ear). For each reference object (i.e., plant and pole), we collected echoes at different distances, azimuths, and object orientations. We collected five measurements for each combination of these parameters. In this way, we constructed a bank of reference echoes allowing us to approximate the echo returned by both objects for different distances, azimuths, and orientations to the agent.

The bank of reference echoes allowed us to create complex simulated environments consisting of arbitrary numbers of plants and food items by combining plant and pole echoes to generate the echoes received by the agent in the simulated EchoVR. In simulating the echoes received by the agent, we set the field of view (FoV) for the agent to $\pm 45^\circ$ in azimuth and 2.5 m in range. Beyond this FoV, the echoes picked up by the physical sonar device are very weak (low signal-to-noise ratio). Hence, limiting the FoV of the simulated agent to these ranges is justified. This range is limited compared to the detection distance of bat sonar [30]. Therefore, our simulated FoV presents a worst-case scenario.

For each agent's pose in the arena, we listed all objects within the FoV, including their distances and azimuths to the agent. We selected reference echoes for each object in this list. In particular, we chose reference echoes collected at a distance closest (but smaller than) to the distance of the virtual

object. Next, we applied atmospheric and geometric spreading attenuation to correct for the difference between the reference echo's distance and the exact distance to the object in the EchoVR (see [16], [30] for the equations governing these phenomena). We linearly interpolated between two reference echoes nearest to the object's azimuth to correct the reference echoes for the object's exact azimuth. We summed all objects' resulting echoes (after appropriate time shifting) to obtain the complex simulated echoes received. We perform this operation once for each microphone on the physical sonar system, representing each ear. Thus, we obtained the echoes at the simulated agent's left and right ears.

B. Agent cognitive architecture

The agent's cognitive architecture consisted of three components: (1) a model of the bat's peripheral auditory system modeling the cochlear processing of simulated echoes as received by the ears; (2) a decision module that decides, for the current echoic input, which sensorimotor loop to activate based on the current sensory data; and (3) two sensorimotor loops that estimate acoustic cues from the current echoic sensory input and calculate the motor output (behavior) based on the cues. In the following sections, we discuss each of these components.

1) *Cochlear processing*: The simulated echoes returned by the EchoVR were subjected to a series of filtering operations mimicking the bat's cochlear processing, based on the model proposed by Wiegand [31]. The simulated echo at each ear was subjected to a gammatone filter, a half-wave rectifier, a 1 kHz lowpass filter, and a non-linearly compression filter using an exponent of 0.4. The filters essentially extracted the envelope of the echoes' waveforms. Next, the envelope for each ear was subsampled from 7000 samples to 50 samples by averaging every 125 samples and truncating to the 50 first points. We subsampled the envelopes to reduce the dimension of input fed to the decision module to lower the computational load. Moreover, previous work has shown that echoes can be substantially subsampled without losing much information [32]. Finally, we zeroed all the points below a fixed threshold based on the internal noise of our sonar transducer. Therefore, any values in the echo envelope above zero could be considered echoes returned by the environment. The filtering and compressing process is illustrated in Figure 2. The final envelopes from both ears were concatenated into a 100-point vector and fed into the decision module to select a sensorimotor loop. The same envelopes were also fed into the sensorimotor loops for calculating motor output.

2) *Decision module*: For the decision module of the model, we used a Strategy Selection Network (SEN), i.e., a fully connected neural network with 4 hidden layers, consisting of 128, 128, 128, and 64 Rectified Linear Units, respectively. The SEN took a 100-point vector as input containing the preprocessed echoes from both ears and provided output via two units corresponding to the two sensorimotor loops. Each output unit returned an estimate of the expected reward (Q-Value) for activating its corresponding sensorimotor loop.

3) *Sensorimotor loops*: The sensorimotor loop with the highest Q-Value as estimated by the SEN was activated at each point in time. The loops were implemented and (biologically plausibly) parameterized based on our previous work modeling prey capture [16], [19] and obstacle avoidance in bats [17], [18]. From this earlier work, we know that, in isolation, these sensorimotor loops can support prey capturing and obstacle avoidance.

Both loops use the same two acoustic cues derived from the echoic inputs to set the bat's linear and rotational velocity: (1) estimated distance d_e to the nearest object and (2) the interaural intensity difference (IID). Behavioral and neurophysiological evidence has established that the bat's auditory pathway extracts these cues [33]. In our implementation, the estimated distance d_e was derived from the simulated echoes' compressed envelopes by dividing the time-of-flight for the onset of the echoes (first above threshold sample) by twice the speed of sound. Next, from the onset of the echo envelope, we took an integrating window over 0.7 meters (0.4 ms) for each ear. This operation resulted in the intensity of the echo onset in each ear (W_{left} and W_{right}). Finally, the IID was calculated as follows,

$$\text{IID} = 10 \times \log_{10} \frac{W_{\text{left}}}{W_{\text{right}}} \quad (1)$$

Figure 3 shows the linear and angular velocity (v, ω) profiles applied under each of the two loops. The linear velocity for each sensorimotor loop is 2.5 m/s when $d_e > 1$ m and is $1.5 \times d_e + 1$ m/s when $d_e \leq 1$ m. However, the linear velocity is set to 0.05 m/s when $d_e < 0.5$ m for the avoidance loop. For the avoidance loop, $|\omega|$ is a function of d_e (see Figure 3b) with $\text{sign}(\omega) = \text{sign}(\text{IID})$. For the approach sensorimotor loop, $\omega = -\text{sign}(\text{IID}) \times 250^\circ/\text{s}$, with a magnitude cutoff set at $1000^\circ/\text{s}$ (see Figure 3c). In previous work, these values have been selected as corresponding to the velocities observed in bat flight (see [16], [17] for details).

C. Training the simulated bat

The SEN decision module was trained using deep reinforcement learning [34], an approach in which a deep neural network is trained by allowing the agent to learn from trials and errors by interacting with the environment to maximize rewards. The implementation of our deep reinforcement learning agent was based on the TF-agents package [35]. The training arena consisted of walls constructed of (simulated) plants spacing 0.4 meters apart. The walls formed a square of four 3-meter-wide tunnels with four 90° turn corners (see Figure 4a). At the beginning of each training episode, the agent was initialized at world coordinate $x = 5.5$ m, $y = -2.5$ m, facing north. The initial location was jittered by ± 25 cm and a jitter of $\pm 20^\circ$ was added to the facing direction to add some randomness to the initial pose. Food items spawned in the shaded areas indicated in Figure 4a. Food first spawned in area I. After being captured, the food item was removed from the map, and a new food item spawned in the next shaded area in the order of I, II, III, IV, and back to I. There was only

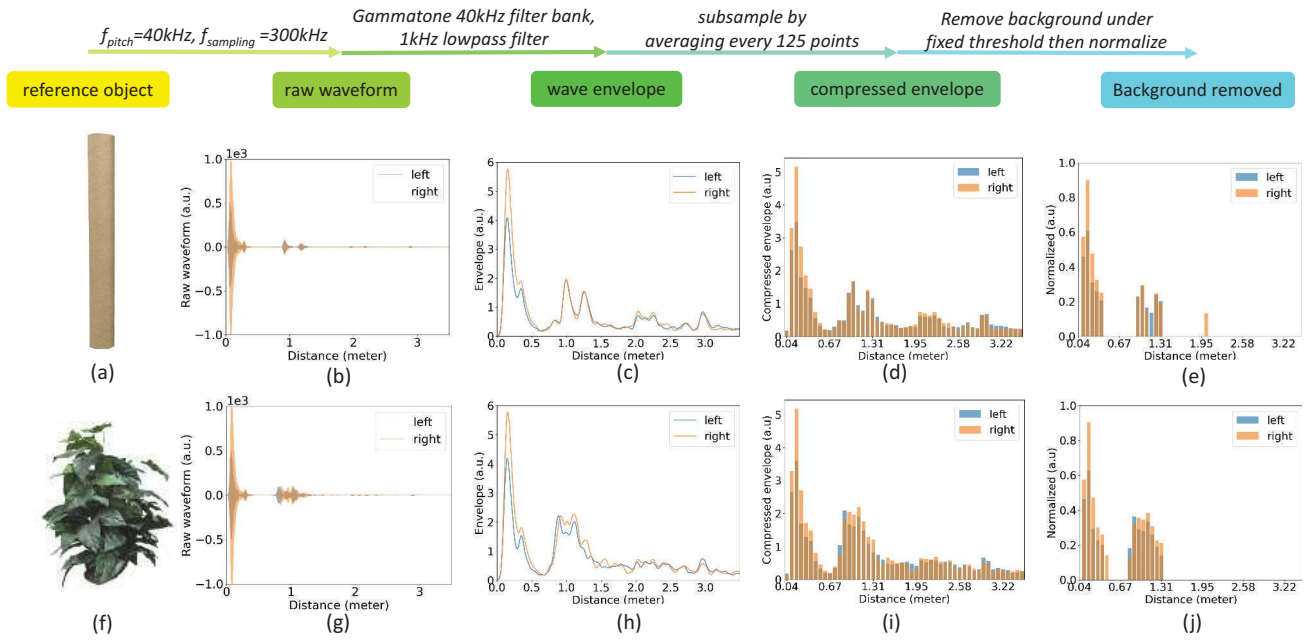


Fig. 2. Filtering and compressing process simulating the cochlear processing of echoes [31]. (a, f) physical reference objects (food: pole, obstacle: plant). (b, g) raw echoes waveform of the reference object collected by the physical sensor. (c, h) Envelopes of the echo waveform after a series of filtering. (d, i) compressed envelopes via subsampling. (e, j) Final envelopes with background noise removed by thresholding

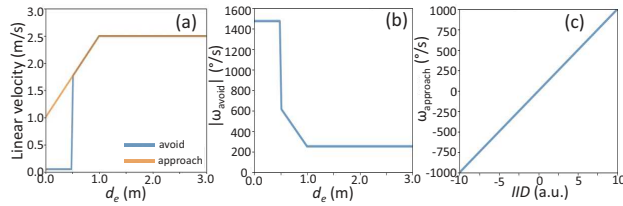


Fig. 3. Linear and angular velocity profile. (a) The linear velocity of both loops. (b) The magnitude of angular velocity for avoidance loop. (c) Angular velocity profile for approaching loop.

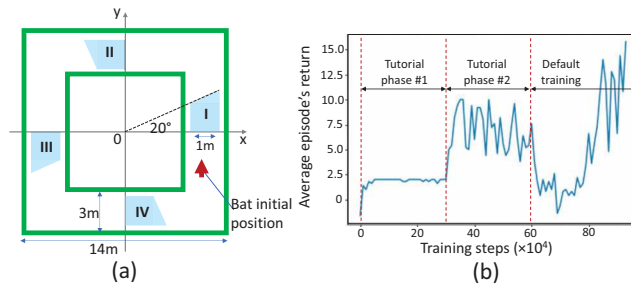


Fig. 4. (a) The default training setup for the foraging task. (b) Learning curve. Notes that pre-set maximum number of food in each training phase affected the maximum episode's returns the agent could reach.

one food item in the maze at any time, and the agent would hunt for food until it collected the maximum amount of food allowed. The episode ended when: (1) the agent collected the pre-set maximum number of food items (default was 11); (2) the agent hit a wall; or (3) the agent exceeded the time allowed to capture the current food item.

We used Deep Q-Network (DQN) [36] as the learning algorithm to train the SEN decision-making module. We gave the agent a +2 reward for each food item collected and -2 for hitting the wall. The discount factor was set to 0.999 to motivate the agent to seek long-term rewards. DQN uses a hybrid exploration-exploitation policy called ϵ -greedy, with ϵ being the probability that the agent chooses a random policy instead of the approximated optimal policy. It gradually switches to a deterministic policy by linearly reducing ϵ to zero over time. In our implementation, we set ϵ starting at 0.8 to allow a high level of exploration in the beginning. We trained the controller until an early-stopping condition of average episode reward exceeding 15 (more than eight food items collected) was reached.

At the beginning of this study, training resulted in the agent prioritizing the avoidance sensorimotor loop since the likelihood of hitting the plants was much higher than finding food. To address this, two tutorial phases with a duration of 300,000 steps each were introduced. In the first tutorial phase, we initialized the agent at a distance of 1.5 meters from the food spawning area and ended the episode when the first food was reached. This increased the likelihood of the agent encountering food so that it could come up with a

strategy to approach the food instead of constantly activating the avoidance sensorimotor loop. The second tutorial phase was similar to the first, but the agent was allowed to collect up to 6 food items. This allowed the agents to optimize for seeking more rewards after the first food item. Finally, the default training settings were restored after completing two tutorial phases.

III. RESULTS

The early-stopping condition was met after 930,000 steps, with an average episode's return of 15.8. Figure 4b shows the average episode returns as a function of the training steps. We noticed that the main reason for an episode to end was that the agent would encounter a food item that it failed to capture within the time limit. We designed a series of tests to study the agent's behavior in more detail. The setup and results of those tests are discussed in this section. (See [37] for a video illustrating these tests).

Figure 5a illustrates how the echoes of both ears changed through a single episode. Figure 5b also shows how the difference between the Q-values of the two sensorimotor loops changed. Here, we briefly point out some interesting aspects of the episode to explain the sensory data received by the agent. Around the middle of the episode (125th step), the agent headed straight toward a wall of plants before steering away. As Figure 5a shows, the echoes from the plants became increasingly loud and closer. Toward the end of the episode, as the agent approached the food item, the echoes of the food item (small unique pulse) stood out separately from other echoes. In Figure 5b, the Q-values of each sensorimotor loop, estimated by the SEN output layer, were similar, with a slight bias toward the avoidance loop. The SEN activates the sensorimotor loop with the large estimated Q-value. Therefore, the avoidance loop was activated over 70% of the time when cruising through the tunnel. As the agent homed in on the food item, the difference between the Q-value of the approach loop and which of the avoidance loop grew significantly (See the last 20 steps in Figure 5b).

A. Testing different maze geometries

We first tested the agent in the same arena it was trained in. Figure 6a shows the test setup. The agent was initialized at the center of the south tunnel facing eastward. The initial pose of the agent was added with the random jitters described in Section II-C. Food spawned in the shaded area. The test episode stopped when the agent either (1) successfully captured the food, (2) hit the wall, or (3) crossed the upper or lower boundary. The upper and lower boundaries were set not only to stop the episode from running forever in case the agent failed to capture the food item but also to help classify the outcome of an episode, as explained later in this section. Based on these stopping conditions, we categorize each episode into one of five categories: *success*, *miss*, *bail*, *wrong-way*, and *collision*.

Success indicated the agent captured the food in the episode. *collision* was defined as the agent hitting an obstacle

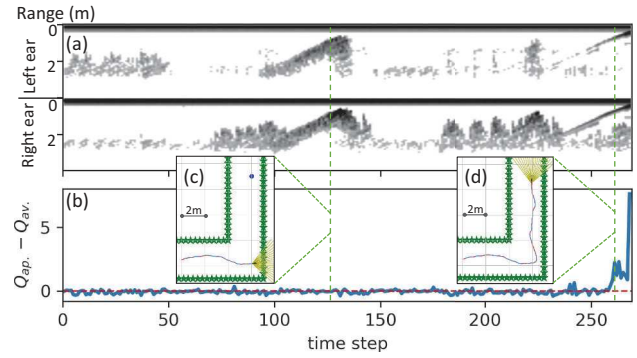


Fig. 5. (a). Visualization of echoic input the agent received during an episode. The horizontal axis represents the range of the sound source, while the vertical axis illustrates the steps. (Top: Left ear echoes, Bottom: Right ear echoes). (b) Difference between the Q-value of the approach loop ($Q_{ap.}$) and the Q-value of the avoid loop ($Q_{av.}$) over time steps of the same episode. (c,d) Top view snapshot of the episodes at the 125th step and the 260th step.

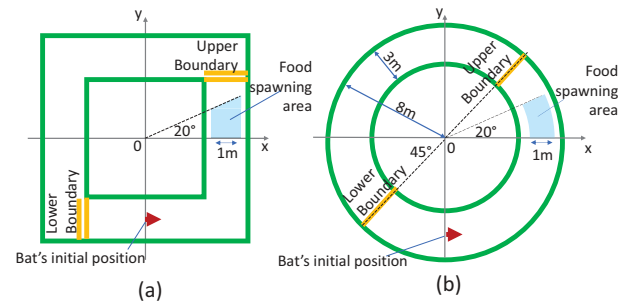


Fig. 6. (a) The test setup for the box maze. (b) The test setup for the donut maze.

(plant/wall). *miss* occurred when the agent crossed the upper boundary and never got close to the food item (i.e., within 0.5 m and $\pm 10^\circ$ azimuth). *Bail* occurred when the agent crossed either boundary after being close to the food item for at least a single step. *Wrong-way* was when the agent crossed the lower boundary and never got close to the food item. After running 1000 test episodes, 94.6% of all episodes ended in *success*, while none ended in *collision*. Furthermore, 4.3% ended in *miss*, 0.3% ended in *bail*, and 0.8% ended in *wrong-way* (see Figure 7).

To assess whether the trained controller generalizes to a different arena, we constructed a donut-shaped arena with an 8-meter radius whose tunnel has a width of 3-meter. The test setup is described in Figure 6b. For the donut maze, 92.9% of episodes ended in *success*. Even though the agent performed slightly worse than in the box maze (Figure 6a), the agent still performed well in the donut maze. The agent remained *collision-free* in this maze. The *miss*, *bail* and *wrong-way* rates were 6.4%, 0.5%, and 0.2%, respectively (Figure 7). This showed that the agent was not over-trained on the box maze.

For the first 100 steps of each episode in both arenas, the agent only traveled in an arena segment that did not contain

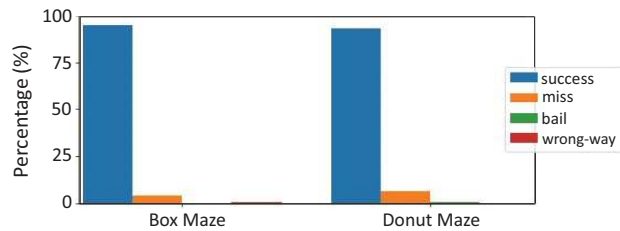


Fig. 7. Rates of different episode endings for tests on a different maze geometry.

a food item. In this arena section, the agent activated the *approach* loop about 23% of the time. Even though the best strategy for an empty arena segment might seem to activate the *avoid* loop 100% of the time, the agent still occasionally activated the *approach* loop. The proportion of steps in which the *approach* loop was activated increased as the agent neared the food item. Analyzing the last 20 steps of all *success* episodes, we found that the *approach* loop was activated approximately 65% of the time.

To better understand where the agent activated each sensorimotor loop, we reran 1000 successful iterations of the model with the agent starting at a fixed initial pose and a food item at world coordinate of $(x = 6m, y = 1m)$. Figure 8a,b shows the locations at which both loops were active throughout the 1000 episodes. This 8c,d separately shows the locations at which the difference between the Q-values for both loops was larger than 0.5. We isolated those instances as they can be considered locations at which the SEN's confidence in choosing a sensorimotor loop was high. Figure 8c suggests that the SEN became more confident in selecting the *approach* loop as the agent homed in on the food item. The SEN activated the *avoidance* loop with high confidence at locations close to the wall of obstacles (Figure 8d). The *avoidance* loop was also activated with confidence in a small region in front of the food item. We conjecture that this unexpected avoidance behavior might not have been unlearned as a successful approach is unavoidable at this close distance from the food item.

B. Trap avoidance test

Even though the agent achieved a high success rate, this does not show that the agent has learned to recognize the food items. Indeed, the controller might have figured out a strategy to isolate any object in the middle of the tunnel instead of learning to recognize a food item. We designed a trap avoidance test to examine whether the agent learned to discriminate between the plants and the targets. We used the same test setup for the box and donut arenas. However, the food item (which returned echoes derived from a cardboard pole) was replaced with a plant. If the agent avoids the isolated plants, we can conclude that the agent learned to discriminate between the plants and the targets.

The *collision* rate was 18% for the box maze and 15.5% for the donut maze. The *Bail* rate was 74.8% for the box

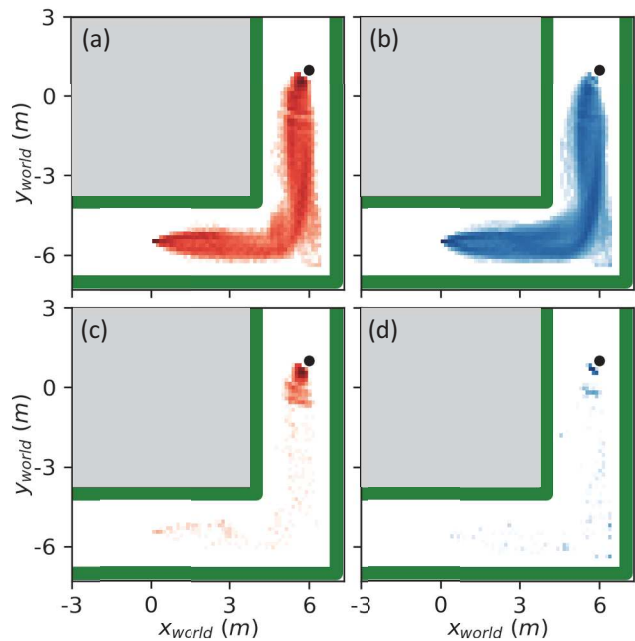


Fig. 8. Heat map of agent's locations in 1000 successful episodes with fixed initial pose and food location. (a) Locations where the agent activated the approach loop. (b) Locations where the agent activated the avoid loop. (c) Locations where the Q-value of the approach loop is 0.5 higher than the avoidance loop. (d) Locations where the Q-value of the avoidance loop is 0.5 higher than its counterpart.

maze and 77.0% for the donut maze. For the box maze, *miss* and *wrong-way* rates were 6.7% and 0.5%, respectively. For the donut maze, *miss* and *wrong-way* rates were 7.3% and 0.2%, respectively. We separated the collisions into two subcategories, *accidental* and *intentional*, based on the trap's location to the agent at the moment of collision. *Accidental collision* occurred when the trap was outside the $\pm 45^\circ$ FoV when hit. In this case, the bat could not detect the trap plant. In contrast, *intentional collision* occurred when the trap was within the $\pm 45^\circ$ view width when hit. *Accidental collisions* accounted for 17.1% of all episodes in the box maze and 15.1% in the donut maze. *Intentional collision* accounted for 0.9% of all episodes in the box maze and 0.4% in the donut maze (see Figure 9). Therefore, most collisions were in the *accidental* subcategory.

These results, especially the high *bail* rate, suggested that the agent could discriminate between plants and food items. However, it had to approach an isolated object closely before being able to distinguish food from obstacles. Hence, the agent often turns away from the trap when close to it. On many occasions, turning away at the last few steps put the trap in a blind spot of the agent at a close distance, causing the agent to collide with the plant accidentally.

C. Robustness against actuator noise

We assessed whether our model could deal with noise in the actuation of motor commands. This increases the biological plausibility of our model, as motor noise is inherent in

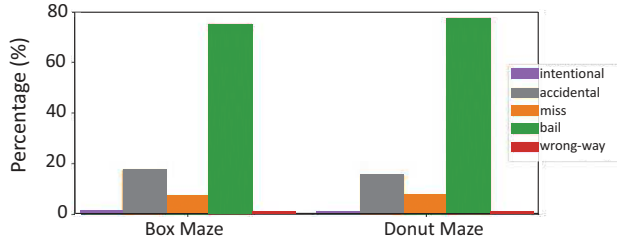


Fig. 9. Rates of different episode endings for trap test.

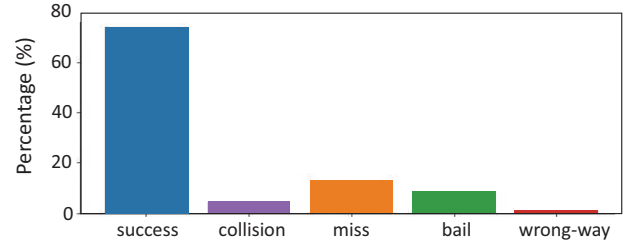


Fig. 10. Rates of the different episode endings for the Monte Carlo simulation where the velocity of each wheel was varied by randomly sampling the wheel's RPS and diameter.

biological systems. Moreover, in future work, we intend to test our trained model on a differential drive robot to evaluate whether our hybrid method results in a control model capable of operating in the real world. Before doing so, we wanted to assess how robust the trained agent would be against actuation noise. Our model already incorporated sensors' noise by using echoes collected using a physical system. Indeed, the signal-to-noise ratio of the sonar transducer used to collect the echoes is substantially lower than that of bat biosonar [32].

We performed a Monte Carlo simulation where rotations per second (RPS) and diameters of the two wheels were randomly sampled for each test episode. We set γ_L and γ_R as random variables scaling the left and right wheels' RPS. Likewise, the random variables scaling the diameters of left and right wheels are denoted as δ_L and δ_R . The effective velocities of two wheels $\bar{u}_{L,R}$ can be calculated as,

$$\bar{u}_{L,R} = \overline{\text{RPS}}_{L,R} \times \bar{D}_{L,R} = \gamma_{L,R} \delta_{L,R} \times u_{L,R} \quad (2)$$

where

$$\overline{\text{RPS}}_{L,R} = \gamma_{L,R} \times \text{RPS}_{L,R} \quad (3)$$

$$\bar{D}_{L,R} = \delta_{L,R} \times D_{L,R} \quad (4)$$

with $\text{RPS}_{L,R}$ the ideal rotational rate, $\overline{\text{RPS}}_{L,R}$ the actual (scaled) rotational rate, $D_{L,R}$ is ideal diameter, and $\bar{D}_{L,R}$ the actual (scaled) diameter of the left and right wheels.

At the beginning of each simulation episode, we sampled γ_L , γ_R , δ_L , and δ_R independently from normal distributions with means of 1 and standard deviations of 0.1 and 0.05 for $\gamma_{L,R}$ and $\delta_{L,R}$, respectively. It should be noted that these are very noisy conditions. Indeed, on average, the actual velocities of the wheels differed by over 10% from their ideal. In this simulation, we assumed that the distance between two wheels was 25 cm (this approximates the diameter of several commercially available robots).

We ran 3000 test episodes in the box maze. The *success* rate dropped to 73.7% while the collision rate increased from 0% to 4.3%. The rate of *miss*, *bail*, and *wrong-way* were 12.9%, 8.5%, and 0.6%, respectively (see Figure 10). We noticed that most collisions occur in episodes the right wheel was substantially faster than the left wheel. This caused the robot to over-steer when making a left turn at the southeast corner of the maze, resulting in a collision. However, the success rate was still relatively high in this simulation, which indicated

that the model could perform reasonably, even in the face of substantial actuator variation.

IV. DISCUSSION

Brooks proposed the Behavior-Based architecture as an alternative to approaches that relied heavily on building internal world models and planning [21]. Analogously, we argued in several papers [3], [16]–[19] against the assumption that bats build a (3D) model of the environment using their sonar system and use this for planning actuation. We argued that a close link between sensing and acting (without the need for extensive representations or planning) would allow bats to respond in real-time to the complex and changing environments in which they operate. Moreover, not requiring a model of the environment sidesteps the problems associated with deducing the world's layout from underspecified echo signals.

In previous work, we proposed and tested sensorimotor loops in isolation [3], [16]–[19]. This paper is the first test of the idea that a combination of sensorimotor loops can be used to scaffold more complex sonar-based behaviors. As said in the introduction, we consider the current work preliminary. This paper combines only two sensorimotor loops previously tested and parameterized to model common bat behavior, i.e., approaching targets while avoiding obstacles. However, in modeling this, we developed a methodology that integrates simulation-based machine learning with acoustic data collected using real transducers. While simulation necessarily incurs a reality gap, building a simulated environment that returns echoes derived from measurements limits this gap. Therefore, we can now expand the current work to tackle more complex behaviors based on this work. For example, we plan on modeling the behavior of nectar-feeding bats [38]–[41]. This behavior requires the bat to locate the target flower (among other vegetation) and align itself with the flower's opening to 'dock' and access the nectar.

We consider the current work as an initial test of our overarching hypothesis concerning the cognitive architecture underlying sonar-based behavior in bats. The work here is also preliminary in one other respect. We have only preliminarily analyzed the control strategy that emerged from the training. Indeed, we expect but have not tested that analyzing the SEN could aid us in understanding the strategies bats might adopt in

tackling more complex tasks. In other words, by applying the methods of artificial ethology [42] (i.e., studying the behavior and controllers) of artificial agents, we hope to be able to draw conclusions about (1) the structure of more complex sonar-based tasks and (2) higher-level strategies that bats could use to tackle them. In turn, we hope this leads to hypotheses that can be tested in behavioral or neurophysiological studies on bats.

REFERENCES

- [1] M. W. Holderied, G. Jones, and O. von Helversen, "Flight and echolocation behaviour of whiskered bats commuting along a hedgerow: range-dependent sonar signal design, doppler tolerance and evidence for acoustic focussing," *Journal of Experimental Biology*, vol. 209, no. 10, pp. 1816–1826, 2006.
- [2] J. R. Barchi, J. M. Knowles, and J. A. Simmons, "Spatial memory and stereotypy of flight paths by big brown bats in cluttered surroundings," *The Journal of experimental biology*, vol. 216, no. 6, pp. 1053–1063, 2013.
- [3] D. Vanderelst and H. Peremans, "A computational model of mapping in echolocating bats," *Animal Behaviour*, vol. 131, pp. 73–88, 2017.
- [4] U. Firzlaff, M. Schuchmann, J. E. Grunwald, G. Schuller, and L. Wiegrebe, "Object-oriented echo perception and cortical representation in echolocating bats," *PLoS biology*, vol. 5, no. 5, p. e1007, 2007.
- [5] D. R. Griffin, *Listening in the dark; the acoustic orientation of bats and men*. Yale University Press, New Haven., 1958.
- [6] R. Simon, M. W. Holderied, and O. von Helversen, "Size discrimination of hollow hemispheres by echolocation in a nectar feeding bat," *Journal of experimental biology*, vol. 209, no. 18, pp. 3599–3609, 2006.
- [7] S. Greif and B. M. Siemers, "Innate recognition of water bodies in echolocating bats," *Nature communications*, vol. 1, p. 107, 2010.
- [8] D. Vanderelst, J. Reijnders, U. Firzlaff, and H. Peremans, "Dominant glint based prey localization in horseshoe bats: a possible strategy for noise rejection," *PLoS computational biology*, vol. 7, no. 12, p. e1002268, 2011.
- [9] J. Reijnders, D. Vanderelst, and H. Peremans, "Morphology-induced information transfer in bat sonar," *Physical review letters*, vol. 105, no. 14, p. 148701, 2010.
- [10] D. Vanderelst, J. Reijnders, Steckel, J., and H. Peremans, "Information generated by the moving pinnae of rhinolophus rouxi: tuning of the morphology at different harmonics," *PLOS ONE*, vol. 6, p. e20627, 2011.
- [11] R. Kuc, "Sensorimotor model of bat echolocation and prey capture," *The Journal of the Acoustical Society of America*, vol. 96, no. 4, pp. 1965–1978, 1994.
- [12] V. Walker, H. Peremans, and J. Hallam, "One tone, two ears, three dimensions: A robotic investigation of pinnae movements used by rhinolophid and hipposiderid bats," *The Journal of the Acoustical Society of America*, vol. 104, no. 1, pp. 569–579, 1998.
- [13] F. Schillebeeckx and H. Peremans, "Biomimetic sonar: 3d-localization of multiple reflectors," in *Intelligent Robots and Systems (IROS), 2010 IEEE/RSJ International Conference on*. IEEE, 2010, pp. 3079–3084.
- [14] J. Steckel and H. Peremans, "Batslam: Simultaneous localization and mapping using biomimetic sonar," *PLOS ONE*, vol. 8, no. 1, p. e54076, Jan. 2013.
- [15] B. Barshan and R. Kuc, "Bat-like sonar system strategies for mobile robots," in *Systems, Man, and Cybernetics, 1991. 'Decision Aiding for Complex Systems, Conference Proceedings., 1991 IEEE International Conference on*, Oct 1991, pp. 905–910 vol.2.
- [16] D. Vanderelst and H. Peremans, "Modeling bat prey capture in echolocating bats: The feasibility of reactive pursuit," *Journal of theoretical biology*, vol. 456, pp. 305–314, 2018.
- [17] D. Vanderelst, M. W. Holderied, and H. Peremans, "Sensorimotor model of obstacle avoidance in echolocating bats," *PLoS computational biology*, vol. 11, no. 10, p. e1004484, 2015.
- [18] C. B. Mansour, E. Koreman, J. Steckel, H. Peremans, and D. Vanderelst, "Avoidance of non-localizable obstacles in echolocating bats: A robotic model," *PLoS Computational Biology*, vol. 15, no. 12, 2019.
- [19] T. Nguyen, D. Vanderelst, and H. Peremans, "Sensorimotor behavior under informational constraints: a robotic model of prey localization in the bat *micronycteris microtis*," in *ALIFE 2021: The 2021 Conference on Artificial Life*. MIT Press, 2021.
- [20] R. Brooks, "A robust layered control system for a mobile robot," *IEEE Journal on Robotics and Automation*, vol. 2, 1986.
- [21] R. A. Brooks, *Cambrian intelligence: The early history of the new AI*. MIT press, 1999.
- [22] G. Schouten and J. Steckel, "A biomimetic radar system for autonomous navigation," *IEEE Transactions on Robotics*, vol. 35, no. 3, pp. 539–548, 2019.
- [23] D. Vanderelst, F. De Mey, H. Peremans, I. Geipel, E. Kalko, and U. Firzlaff, "What noseleaves do for fm bats depends on their degree of sensorial specialization," *PLOS ONE*, vol. 5, no. 8, p. e11893, 2010.
- [24] F. De Mey, J. Reijnders, H. Peremans, M. Otani, and U. Firzlaff, "Simulated head related transfer function of the phyllostomid bat *phyllostomus discolor*," *The Journal of the Acoustical Society of America*, vol. 124, no. 4, pp. 2123–2132, 2008.
- [25] D. Vanderelst, F. De Mey, and H. Peremans, "Simulating the morphological feasibility of adaptive beamforming in bats," in *From Animals to Animals 11*. Springer, 2010, pp. 136–145.
- [26] D. Vanderelst, R. Jonas, and P. Herbert, "The furrows of rhinolophidae revisited," *Journal of The Royal Society Interface*, vol. 9, no. 70, pp. 1100–1103, 2012.
- [27] A. Surlykke and C. F. Moss, "Echolocation behavior of big brown bats, *ptesicus fuscus*, in the field and the laboratory," *The Journal of the Acoustical Society of America*, vol. 108, no. 5, pp. 2419–2429, 2000.
- [28] Y. Yovel, M. O. Franz, P. Stilz, and H.-U. Schnitzler, "Complex echo classification by echo-locating bats: a review," *Journal of Comparative Physiology A*, vol. 197, no. 5, pp. 475–490, 2011.
- [29] K. D. Roeder, "Echoes of ultrasonic pulses from flying moths," *The biological bulletin*, vol. 124, no. 2, pp. 200–210, 1963.
- [30] W. P. Stilz and H. U. Schnitzler, "Estimation of the acoustic range of bat echolocation for extended targets," *The Journal of the Acoustical Society of America*, vol. 132, no. 3, pp. 1765–1775, 2012.
- [31] L. Wiegrebe, "An autocorrelation model of bat sonar," *Biological cybernetics*, vol. 98, no. 6, pp. 587–595, 2008.
- [32] V. Adarsh, "Place recognition using batlike sonar: compressability of templates," *In preparation*, 2020.
- [33] G. Neuweiler, *The biology of bats*. Oxford University Press, 2000.
- [34] Y. Li, "Deep reinforcement learning: An overview," *arXiv preprint arXiv:1701.07274*, 2017.
- [35] S. Guadarrama, A. Korattikara, O. Ramirez, P. Castro, E. Holly, S. Fishman, K. Wang, E. Gonina, N. Wu, E. Kokopoulou, L. Sbaiz, J. Smith, G. Bartók, J. Berent, C. Harris, V. Vanhoucke, and E. Brevdo, "TF-Agents: A library for reinforcement learning in tensorflow," <https://github.com/tensorflow/agents>, 2018, [Online; accessed 25-June-2019]. [Online]. Available: <https://github.com/tensorflow/agents>
- [36] V. Mnih, K. Kavukcuoglu, D. Silver, A. A. Rusu, J. Veness, M. G. Bellemare, A. Graves, M. Riedmiller, A. K. Fidjeland, G. Ostrovski et al., "Human-level control through deep reinforcement learning," *nature*, vol. 518, no. 7540, pp. 529–533, 2015.
- [37] T. H. Nguyen, "Behavior based foraging agent demo," 2021. [Online]. Available: <https://youtu.be/IQgzsbVNa5o?t=3>
- [38] D. von Helversen and O. von Helversen, "Acoustic guide in bat-pollinated flower," *Nature; London*, vol. 398, no. 6730, pp. 759–760, Apr. 1999.
- [39] D. V. Helversen and O. V. Helversen, "Object recognition by echolocation: A nectar-feeding bat exploiting the flowers of a rain forest vine," *Journal of Comparative Physiology. A, Neuroethology, Sensory, Neural, and Behavioral Physiology*, vol. 189, no. 5, pp. 327–36, 2003.
- [40] R. Simon, M. W. Holderied, C. U. Koch, and O. von Helversen, "Floral Acoustics: Conspicuous Echoes of a Dish-Shaped Leaf Attract Bat Pollinators," *Science*, vol. 333, no. 6042, pp. 631–633, 2011.
- [41] T. P. Gonzalez-Terrazas, J. C. Koblit, T. H. Fleming, R. A. Medellín, E. K. Kalko, H. U. Schnitzler, and M. Tschapka, "How nectar-feeding bats localize their food: Echolocation behavior of *Leptonycteris verbabuena* approaching cactus flowers," *PLoS ONE*, vol. 11, no. 9, pp. 1–18, 2016.
- [42] O. Holland and D. McFarland, *Artificial ethology*. Oxford University Press on Demand, 2001.