Endotaxis: A Universal Algorithm for Mapping, Goal-Learning, and Navigation

Tony Zhang[1], Matthew Rosenberg[1], Pietro Perona[2], Markus Meister[1]
[1]Division of Biology and Biological Engineering
[2]Division of Engineering and Applied Science
California Institute of Technology
{tonyzhang, mhrosenberg, perona, meister}@caltech.edu

September 24, 2021

Abstract

An animal entering a new environment typically faces three challenges: explore the space for resources, memorize their locations, and navigate towards those targets as needed. Experimental work on exploration, mapping, and navigation has mostly focused on simple environments – such as an open arena, a pond [1], or a desert [2] – and much has been learned about neural signals in diverse brain areas under these conditions [3, 4]. However, many natural environments are highly constrained, such as a system of burrows, or of paths through the underbrush. More generally, many cognitive tasks are equally constrained, allowing only a small set of actions at any given stage in the process. Here we propose an algorithm that learns the structure of an arbitrary environment, discovers useful targets during exploration, and navigates back to those targets by the shortest path. It makes use of a behavioral module common to all motile animals, namely the ability to follow an odor to its source [5]. We show how the brain can learn to generate internal "virtual odors" that guide the animal to any location of interest. This *endotaxis* algorithm can be implemented with a simple 3-layer neural circuit using only biologically realistic structures and learning rules. Several neural components of this scheme are found in brains from insects to humans. Nature may have evolved a general mechanism for search and navigation on the ancient backbone of chemotaxis.

1 Introduction

Efficient navigation requires knowing the structure of the environment: which locations are connected to which others [6]. One would like to understand how the brain acquires that knowledge, what neural representation it adopts for the resulting map, how it tags significant locations in that map, and how that knowledge gets read out for decision-making during navigation. Here we propose a mechanism that solves all these problems and operates reliably in diverse and complex environments.

One algorithm for finding a valuable resource is common to all animals: chemotaxis. Every motile species has a way to track odors through the environment, either to find the source of the odor or to avoid it [5]. This ability is central to finding food, connecting with a mate, and avoiding predators. It is believed that brains originally evolved to organize the motor response in pursuit of chemical stimuli. Indeed some of the oldest regions of the mammalian brain, including the hippocampus, seem organized around an axis that processes smells [7, 8].

The specifics of chemotaxis, namely the methods for finding an odor and tracking it, vary by species, but the toolkit always includes a random trial-and-error scheme: Try various actions that you have available, then settle on the one that makes the odor stronger [5]. For example a rodent will weave its head side-to-side, sampling the local odor gradient, then move in the direction where the smell is stronger. Worms and maggots follow the same strategy. Dogs track a ground-borne odor trail by casting across it side-to-side. Flying insects perform similar casting flights. Bacteria randomly change

direction every now and then, and continue straight as long as the odor improves [9]. We propose that this universal behavioral module for chemotaxis can be harnessed to solve general problems of search and navigation in a complex environment.

For concreteness, consider a mouse exploring a labyrinth of tunnels (Fig 1A). The maze may contain a source of food that emits an odor (Fig 1A top). That odor will be strongest at the source and decline with distance along the tunnels of the maze. The mouse can navigate to the food location by simply following the odor gradient uphill. Suppose that the mouse discovers some other interesting locations that do not emit a smell, like a source of water, or the exit from the labyrinth (Fig 1A). It would be convenient if the mouse could tag such a location with an odorous material, so it may be found easily on future occasions. Ideally the mouse would carry with it multiple such odor tags, so it can mark different targets each with its specific recognizable odor (Fig 1A mid and bottom).

Here we show that such tagging does not need to be physical. Instead we propose a mechanism by which the mouse's brain may compute a "virtual odor" signal that declines with distance from a chosen target. That neural signal can be made available to the chemotaxis module as though it were a real odor, enabling navigation up the gradient towards the target. Because this goal signal is computed in the brain rather than sensed externally, we call this hypothetical process *endotaxis*.

2 A circuit to implement endotaxis

In Figure 1B we present a neural circuit model that implements three goals: mapping the connectivity of the environment; tagging of goal locations with a virtual odor; and navigation towards those goals. The model includes four types of neurons: feature cells, point cells, map cells, and goal cells.

Feature cells: These cells fire when the animal encounters an interesting feature that may form a target for future navigation. Each feature cell is selective for a specific kind of resource, for example water or food, by virtue of sensory pathways that respond to those stimuli.

Point cells: This layer of cells represents the animal's location.[1] Each neuron in this population has a small response field within the environment. The neuron fires when the animal enters that response field. We assume that these point cells exist from the outset as soon as the animal enters the environment. Each cell's response field is defined by some conjunction of external and internal sensory signals at that location.

Map cells: This layer of neurons learns the structure of the environment, namely how the various locations are connected in space. The map cells get excitatory input from point cells with low convergence: Each map cell should collect input from only one or a few point cells. These input synapses are static. The map cells also excite each other with all-to-all connections. These recurrent synapses are modifiable according to rules of Hebbian plasticity and, after learning, represent the topology of the environment.

Goal cells: These neurons mark the locations of special resources in the map of the environment. A goal cell for a specific feature receives excitatory input from the corresponding feature cell. It also receives Hebbian excitatory synapses from map cells. Those synapses are strengthened when the presynaptic map cell is active at the same time as the feature cell.

Each of the goal cells carries a virtual odor signal for its assigned feature. That signal increases systematically as the animal moves closer to the target feature. A mode switch selects one among many possible virtual odors (or real odors) to be routed to the chemotaxis module for odor tracking. [2] The animal then pursues its chemotaxis search strategy to maximize that odor, which leads it to the selected tagged feature.

2.1 Why does the circuit work?

The key insight is that the output of the goal cell declines systematically with the distance of the animal from that target. This relationship holds even if the environment is a complex graph with

[1] We avoid the term 'place cell' here because (1) that term has a technical meaning in the rodent hippocampus, whereas the arguments here extend to species that don't have a hippocampus; (2) all the cells in this network have a place field, but it is smallest for the point cells.

[2] That mode switch is controlled by the *murinculus*: a tiny mouse inside the mouse that tells the mouse what to do. We do not claim to know how that works.

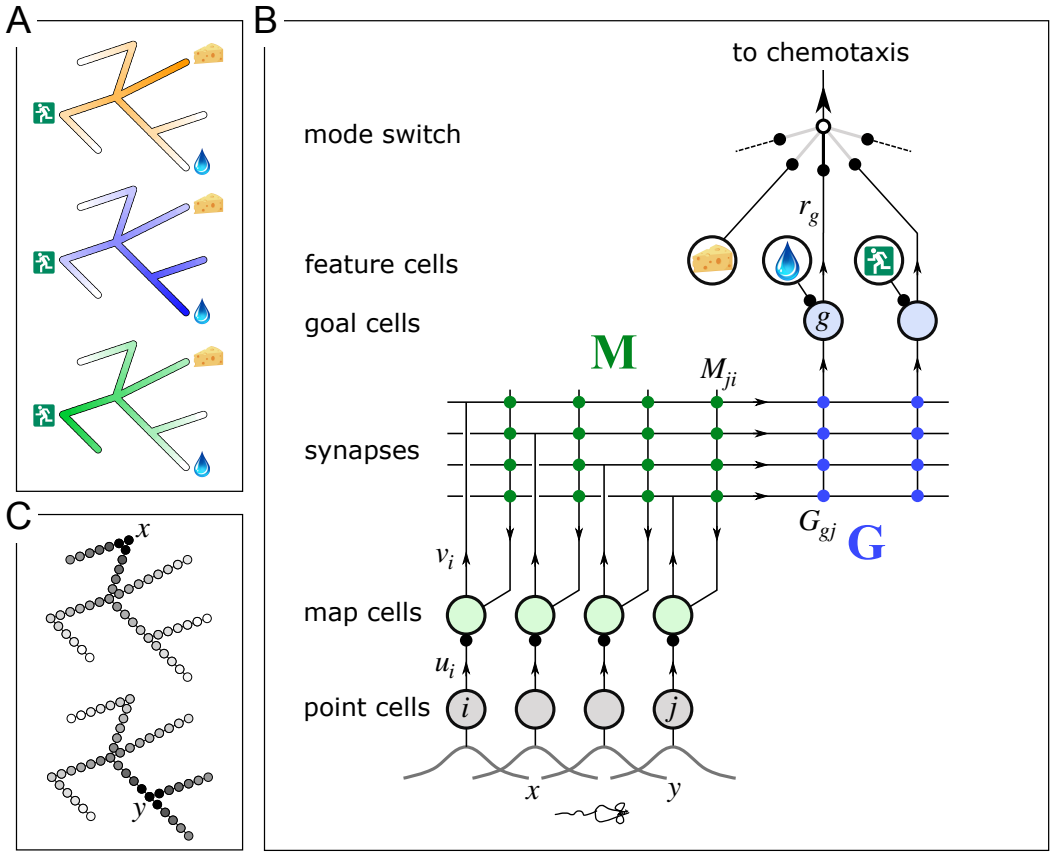

Figure 1: **A mechanism for endotaxis**. **A:** A constrained environment of nodes linked by straight corridors, with special locations offering food, water, and the exit. Top: A real odor emitted by the food source decreases with distance (shading). Middle: A virtual odor tagged to the water source. Bottom: A virtual odor tagged to the exit. **B:** A neural circuit to implement endotaxis. Open circles: four populations of neurons that represent "feature", "point", "map", and "goal". Arrows: signal flow. Solid circles: synapses. Point cells have small receptive fields localized in the environment and excite map cells. Map cells excite each other by recurrent Hebbian synapses and excite goal cells by another set of Hebbian synapses. A goal cell also receives sensory input from a feature cell indicating the presence of a resource, e.g. water or the exit. The feature cell for cheese responds to a real odor emitted by that target. A "mode" switch selects among various goal signals depending on the animal's need. They may be virtual odors (water, exit) or real odors (cheese). The resulting signal gets fed to the chemotaxis module for gradient ascent. Mathematical symbols used in the text: u_i is the output of a point cell at location i, v_i is the output of the corresponding map cell, $\mathbf{M}$ is the matrix of synaptic weights among map cells, $\mathbf{G}$ are the synaptic weights from the map cells onto goal cells, and r_g is the output of goal cell g. **C:** The output of map cells after the map has been learned; here the animal is located at points x (top) or y (bottom). Black means high activity. For illustration, each map cell is drawn at the center of its place field.

constrained connectivity. Here we explain how this comes about, with mathematical details in the supplement.

As the animal explores a new environment, when it moves from one location to an adjacent one, those two point cells briefly fire together. That leads to a Hebbian strengthening of the excitatory synapses between the two corresponding map cells. In this way the recurrent network of map cells learns the connectivity of the graph that describes the environment. To a first approximation, the matrix of synaptic connections among the map cells will converge to the correlation matrix of their inputs [10, 11], which in turn reflects the adjacency matrix of the graph (Eqn 22). Now the brain can use this adjacency information to find the shortest path to a target.

After this map learning, the output of the map network is a hump of activity, centered on the current location x of the animal and declining with distance along the various paths in the graph (Fig 1C top). If the animal moves to a different location y, the map output is another hump of activity, now centered on y (Fig 1C bottom). The overlap of the two hump-shaped profiles will be large if nodes x and y are close on the graph, and small if they are distant. Fundamentally the endotaxis network computes that overlap. How is it done?

Suppose the animal visits y and finds water there. Then the profile of map activity $v_i(y)$ gets stored in the synapses G_{gi} onto the goal cell g that responds to water (Fig 1B, Eqn 26). When the animal subsequently moves to a different location x, the goal cell g receives the current map output $v_i(x)$ filtered through the previously stored synaptic template $v_i(y)$. This is the desired measure of overlap (Eqn 27), and one can show mathematically that it declines exponentially with the shortest graph-distance between x and y (Eqn 28).

3 Performance of the endotaxis algorithm

Some important features of endotaxis can already be appreciated at this level of detail. First, the structure of the environment is acquired separately from the location of resources. The graph that connects different points in the environment is learned by the synapses in the map network. By contrast the location of special goals within that map is learned by the synapses onto the goal cells. The animal can explore and learn the environment regardless of the presence of threats or resources. Once a resource is found, its location can be tagged immediately within the existing map structure. If the distribution of resources changes, the knowledge of the connectivity map remains unaffected. Second, the endotaxis algorithm is "always on". There is no separation of learning and recall into different phases. Both the map network and the goal network get updated continuously based on the animal's trajectory through the environment, and the goal signals are always available for directed navigation via gradient ascent.

3.1 Simultaneous acquisition of map and targets during exploration

To illustrate these functions, and to explore capabilities that are less obvious from an analytical inspection, we simulated agents navigating by the endotaxis algorithm (Fig 1B) through a range of environments (Figs 2-3). In each case we assumed that there are point cells that fire at specific locations, owing to a match of their sensory receptive fields with features in the environment. The locations of these point cells define the nodes of the graph that the agent will learn. Both the map synapses and the goal synapses start out *tabula rasa* with zero synaptic strengths. This is because the animal has no notion of the topology of the environment (which location connects with which other location), and no information on the location of the resources. As the agent explores the environment, for example by a random walk, map synapses get updated based on the simultaneous firing of point cells corresponding to neighboring locations. We used a standard formulation of Hebbian learning, called Oja's rule, which has only two parameters. Similarly the synapses onto goal cells get updated based on the presynaptic map cell and the postsynaptic signal from feature cells. Map cells and goal cells were allowed to learn at different rates (see Section A for detail).

A simple Gridworld environment (Fig 2) serves to observe the dynamics of learning in detail. There are three locations of interest: the entrance to the environment, experienced at the very start of exploration; a water source; and a food item. When the agent first enters the novel space, a feature neuron that responds to the entrance excites a goal cell, which leads to the potentiation of synapses onto that neuron. Effectively that tags the entrance, and from now on that goal cell encodes a virtual

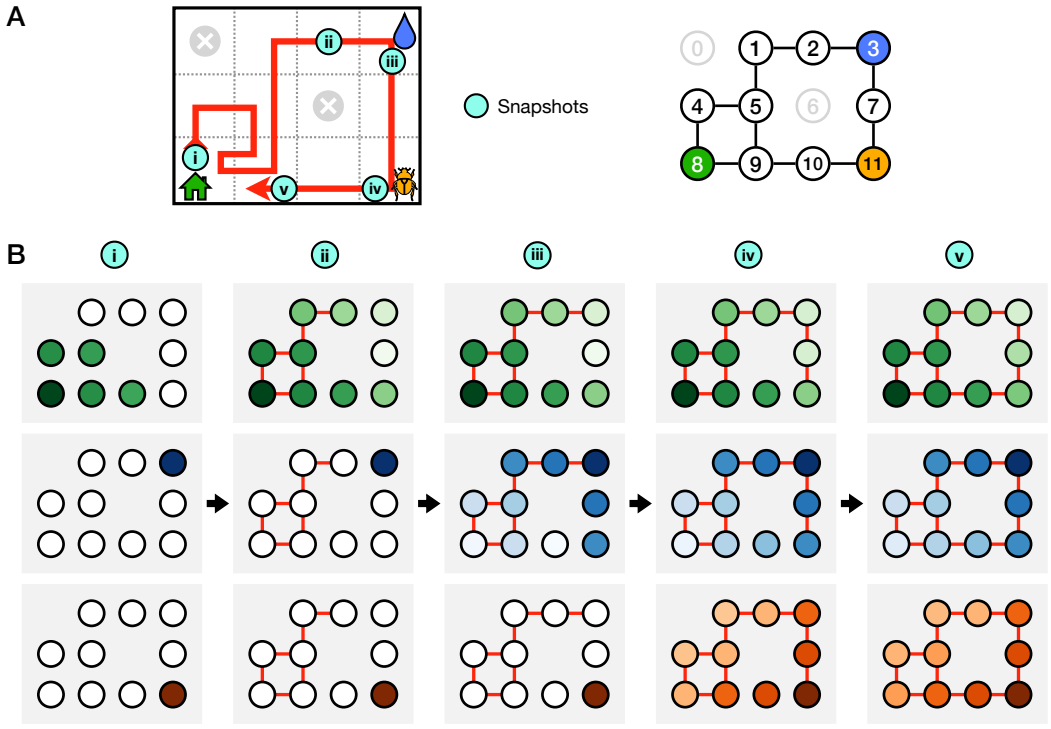

Figure 2: **The map and the targets are learned independently.** **(A)** Left: an agent explores a simple Gridworld with 3 salient goal locations following the red trajectory. Space is discretized into square tiles, each tile represented by one point cell. Circles with crosses represent obstacles, namely tiles that are not reachable. Right: graph of this environment, where each tile becomes a node, and edges represent traversable connections between tiles. **(B)** The response fields of three goal neurons for home (top), water (middle), and bug (bottom) at the 5 instants during the learning process (i-v). Red edges connect previously visited nodes. The response (log color scale) is plotted at each location where the agent could be placed. The agent starts random walking from the entrance (i) and gradually discovers the other two goal locations (water at time iii, bug at time iv). Upon discovery of a goal location, the corresponding goal cell's signal is immediately useful in all previously visited locations (iii, iv) as well as nodes that are ≤ 2 steps away. Any new locations visited subsequently and nodes ≤ 2 steps away are also recruited into the goal cell's response field (v).

"entrance odor" that declines with distance from the entrance. With every step the agent takes, the map network gets updated, and the range of the entrance odor spreads further (Fig 2B top). At all times the agent could decide to follow this virtual odor uphill to the entrance.

The water source starts out invisible from anywhere except its special location (Fig 2B mid i-ii). However, as soon as the agent reaches the water, the water goal cell gets integrated in the circuit through the potentiation of synapses from map cells. Because the map network is already established along the path that the agent took, that immediately creates a virtual "water odor" that spreads through the environment and declines with distance from the water location (Fig 2B mid iii).

As the agent explores the environment further, the virtual odors spread accordingly to the new locations visited (Fig 2B i-iv). After extensive exploration, the map and goal networks reach a steady state. Now the virtual odors are available at every point in the environment, and they decline monotonically with the shortest-path distance to the respective goal location (Fig 2B v). As one might expect, an agent endotaxing uphill on this virtual odor always reaches the goal location, and does so by the shortest possible path (Fig 3B-C i).

We performed a similar simulation for a complex labyrinth used in a recent study of mouse navigation [12]. The topology of the maze was a binary tree with a single entrance, 63 T-junctions, and 64 end nodes (Fig 3A ii). A single source of water was located at one of the end nodes. In these experiments mice learned the shortest path to the water source after visiting it $\sim$10 times; they also performed

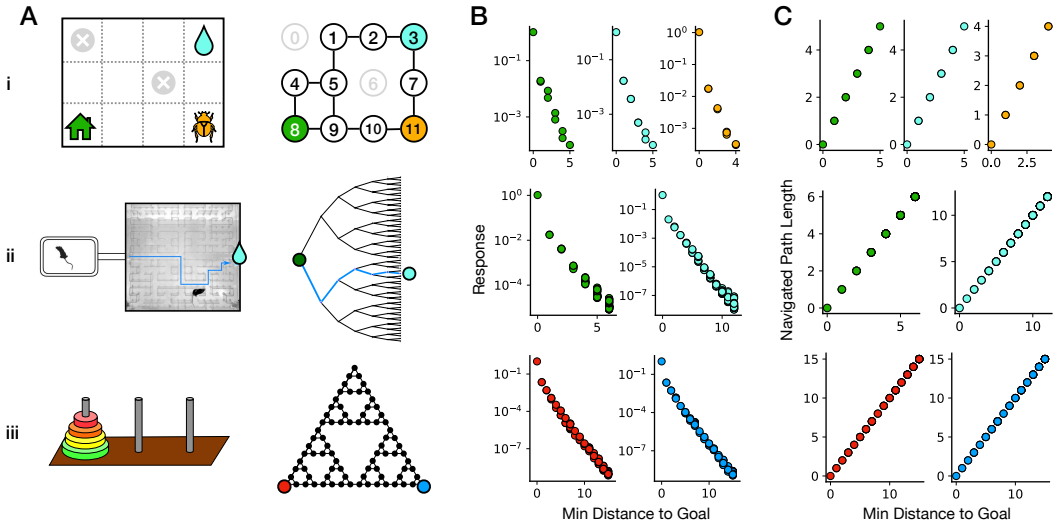

Figure 3: **Endotaxis can operate in environments with diverse topologies. (A)** Three tasks and their corresponding graph representations: i) Gridworld of Fig 2 with 3 goal nodes (home, water, and food). ii) A binary tree labyrinth used in mouse navigation experiments [12], with 2 goals (home and water). iii) Tower of Hanoi game, with 2 goals (the configurations of disks that solve the game). **(B)** The virtual odors after extensive exploration. For each goal neuron the response at every node is plotted against the shortest graph distance from the node to the goal. **(C)** Navigation by endotaxis: For every starting node in the environment this plots the number of steps to the goal against the shortest distance.

error-free paths back to entrance on the first attempt [12]. Again the simulated agent explored the labyrinth with a random walk. The virtual entrance odor allowed it to navigate back to the entrance from any point along the trajectory. The first visit to the water port established a goal cell with virtual water odor. After exploration had covered the entire labyrinth, both the entrance odor and the water odor were available at every location (Fig 3B ii), allowing for flawless navigation to the sources by endotaxis (Fig 3C ii).

It turns out that endotaxis is a useful strategy for cognitive tasks beyond spatial navigation. For instance, the game "Towers of Hanoi" represents a more complex environment (Fig 3A iii). Disks of different sizes are stacked on three pegs, with the constraint that no disk can rest on top a smaller one. The game is solved by rearranging the pile of disks from one peg to another. In any state of the game there are either 2 or 3 possible actions, and they form an interesting graph with many loops (Fig 3A iii). Again the simulated agent explored this graph by random walking. Once it encountered a solution, that state was tagged with a virtual odor. After enough exploration the virtual odor signal was available from every possible game state, and the agent could solve the game from any starting state in the shortest number of moves. This example illustrates that endotaxis can learn cognitive tasks that don't involve spatial movement. It merely requires the existence of neurons that recognize any given state of the game. To start with, the agent has no internal model of the game, so it must happen on the first solution by chance. However, when prompted to solve the problem again, the agent can use the learned virtual odor to complete the game in the fewest possible moves.

These simulations suggest that the endotaxis algorithm can function perfectly in environments of reasonable complexity, learning both the connectivity of the environment and the location of multiple resources within that map. How robust is that performance? First, the model did not require careful tuning of parameters. Instead, we found that endotaxis works over several log units of the two parameters in Oja's rule for synaptic plasticity (Fig 6). It fails in predictable fashion: For example if the agent takes longer to explore the environment than the time constant for synaptic change, then the map is always partially forgotten, and navigation to a target will fail. Second, we considered the effects of noise in neural signals, and found a gradual failure when the signal-to-noise value exceeded 1 (Fig 8).

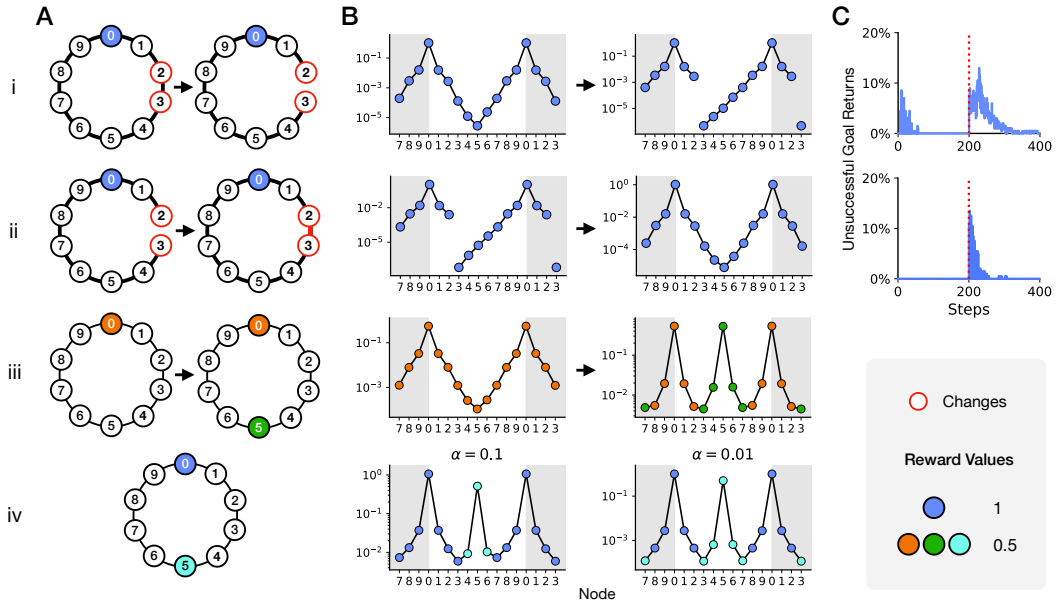

Figure 4: **Endotaxis adapts quickly to changes in the environment or the target locations.** **(A)** A ring environment modified by sudden appearance of a blockage (i), a shortcut (ii), an additional goal target (iii), or two targets with different reward size (iv). Graphs shown before and after modification. Shaded nodes are target locations. Labels identify nodes on the graph. **(B i-iii)** Response profile of the goal neuron after sufficient exploration, shown just before modification (left, after 200 random steps) and after adaptation to the change (right, after an additional 200 steps). Color of nodes indicates the target that the agent will reach by following the virtual odor starting from that node. Note the virtual odor peaks at either one or two targets depending on the environment, with a higher amplitude at the stronger target. **(B iv)** Varying α in Oja's Rule for map learning adjusts the tradeoff between distance and reward. With a large α the stronger target is favored from more starting nodes. **(C)** Fraction of errors in endotaxis from all possible starting nodes, as a function of time before and after the modification (dotted line).

4 Adaptation to change in the environment

An attractive feature of the endotaxis algorithm is that it separates learning the map from learning the target locations. In many real-world environments the topology of the map (how are locations connected?) is probably more stable than the targets (which locations are interesting?). Separating the two allows the agent to adjust to changes on both fronts using different rules and time-scales. We illustrate an example of each.

4.1 Change in connectivity

Suppose that the connectivity of the environment changes. For example, a shortcut appears between two locations that used to be separated, or a blockage separates two previously adjacent locations (Fig 4A i-ii). This alters the correlation in firing among the point cells during the agent's explorations, and over time that will reflect in the synapses of the map network. How will endotaxis adapt to such changes?

To explore these adjustments, we considered navigation on a ring-shaped maze with a single goal location (Fig 4A i). Note that the ring is the simplest graph that offers two routes to a target, and we will evaluate whether the algorithm finds the shorter one. A simulated agent explored the ring by stepping among locations in a random walk, and built the map cell network from that experience. After a period of $\sim$100 steps, navigation by endotaxis was perfect, in that the agent chose the shorter route to the goal from every start node (Fig 4B-C i). When we broke the ring by removing one link, endotaxis failed from some start nodes because it steered the agent towards the blocked path. However, after $\sim$200 steps of additional exploration navigation returned to perfect performance again

(Fig 4C i). Over this period the knowledge of the former link was erased from the map network (Fig 4B i), because the corresponding map synapses weakened while the link was not used.

When we introduced a new shortcut between previously separated locations (Fig 4A ii), a similar change took place. For a brief period endotaxis was suboptimal, because the agent sometimes took the long route even though a shorter one was available (Fig 4C ii). However, that perturbation got incorporated into the map much more quickly than the broken link, after just a few tens of steps of exploration (compare Figs 4C i-ii). One can understand the asymmetry as follows: As the agent explores the environment, a newly available link is confirmed with certainty the first time it gets traveled. By contrast the loss of a link remains uncertain until the agent has not taken that route many times.

4.2 Appearance of new targets

Suppose the agent has discovered one location with a water resource. Some time later water also appears at a second location (Fig 4A iii). When the agent discovers that, the same water goal cell will get activated and therefore receive a potentiation of synapses active at that second location. Now the input network to that goal cell contains the sum of two templates, corresponding to the map outputs from the two target locations. As before, the current map output gets filtered through these synaptic weights to create the virtual odor. One might worry that this goal signal steers the agent to a location half-way between the two targets. Instead, simulations on the ring showed that the virtual odor peaks at both targets, and endotaxis takes the agent reliably to the nearest one (Fig 4B iii).

4.3 Choice between multiple targets

Suppose one of the targets offering the same resource is more valuable than the other, for example because it gives a larger reward (Fig 4A iv). In the endotaxis model (Fig 1B) the larger reward causes higher activity of the feature cell that responds to this resource, and thus stronger potentiation of the synapses onto the associated goal cell (Eqn 20). Thus the input template of the goal cell becomes a weighted sum of the map outputs from the two target locations, with greater weight for the location with higher reward. In simulations, the virtual odor still showed two peaks, but the stronger target had a greater region of attraction (Fig 4B iv left); for some starting locations the agent chose the longer route in favor of the larger reward, a sensible behavior.

What determines the trade-off between the longer distance and the greater reward? In the endotaxis model (Fig 1B) this is set by α_{M}, one of the two parameters of the synaptic learning rule in the map network (Eqn 19). A small α_{M} raises the cost of any additional step traveled and thus diminishes the importance of reward differences (Fig 4B iv right). By contrast a large α_{M} favors the larger reward regardless of distance traveled. One can show that the role of α_{M} is directly equivalent to the discount factor in reinforcement learning theory (Eqn 28).

In summary, endotaxis adapts readily to changes in the environment or in the availability of rewards. Furthermore, it implements a rational choice between multiple targets of the same kind, using a variable weighting of reward versus distance. None of these features required any custom tuning: They all follow directly from the basic formulation in Figure 1B.

5 Discussion

5.1 Summary of claims

We have presented a neural mechanism that can support learning, navigation, and problem solving in complex and changing environments. It is based on chemotaxis, namely the ability to follow an odor signal to its source, which is shared universally by most or all motile animals. The algorithm, called endotaxis, is formulated as a neural network that creates an internal "virtual odor" which the animal can follow to reach any chosen target location (Fig 1). When the agent begins to explore the environment, the network learns both the structure of the space, namely how various points are connected, and the location of valuable resources (Fig 2). After sufficient exploration the agent can then navigate back to those target locations from any point in the environment (Fig 3). The algorithm is *always on* and it adapts flexibly to changes in the structure of the environment or in the locations of targets (Fig 4). Furthermore, even in its simplest form, endotaxis can arbitrate among multiple

locations with the same resource, by trading off the promised reward against the distance traveled (Fig 4). Beyond spatial navigation, endotaxis can also learn the solution to purely cognitive tasks (Fig 3), or any problem defined by search on a graph. The neural network model that implements endotaxis has a close resemblance to known brain circuits. We propose that evolution may have built upon the ancient behavioral module for chemotaxis to enable much more general abilities for search and navigation, even in the absence of odor gradients. In the following sections we consider how these findings relate to some well-established phenomena and results on animal navigation.

5.2 Animal behavior

The millions of animal species no doubt use a wide range of mechanisms to get around their environment, and it is worth specifying which of those problems endotaxis might solve. First, the learning mechanism proposed here applies to complex environments, namely those in which discrete paths form sparse connections between points. For a bird, this is less of a concern, because it can get from every point to any other "as the crow flies". For a rodent and many other terrestrial animals, on the other hand, the paths they may follow are constrained by obstacles and by the need to remain under cover. In those conditions the brain cannot assume that the distance between points is given by euclidean geometry, or that beacons for a goal will be visible in a straight line from far away, or that a target can be reached by following a known heading. Second, we are focusing on the early experience with a new environment. Endotaxis can get an animal from zero knowledge to a cognitive map that allows reliable navigation towards goals encountered on a previous foray. It explains how an animal can return home from inside a complex environment on the first attempt [12], or navigate to a special location after encountering it just once (Figs 2,3). But it does not implement more advanced routines of spatial learning, such as stringing a habitual sequence of actions together into one, or internal deliberation to plan entire routes. Clearly, expert animals will make use of algorithms other than the beginner's choice proposed here.

A key characteristic of endotaxis, distinct from other forms of navigation, is the reliance on trial-and-error. The agent does not deliberate to plan the shortest path to the goal. Instead, it finds the shortest path by locally sampling the real-world actions available at its current point, and choosing the one that maximizes the virtual odor signal. In fact, there is strong evidence that animals navigate by real-world trial-and-error, at least in the early phase of learning [13]. Rats and mice often stop at an intersection, bend their body halfway along each direction, then choose one corridor to proceed. Sometimes they walk a few steps down a corridor, then reverse and try another one. These actions – called "vicarious trial and error" – look eerily like sniffing out an odor gradient, but they occur even in absence of any olfactory cues. Lashley [14], in his first scientific paper on visual discrimination in the rat, reported that rats at a decision point often hesitate "with a swaying back and forth between the passages". Similar behaviors occur in arthropods [15] and humans [16] when poised at a decision point. We suggest that the animal does indeed sample a gradient, not of an odor, but of an internally generated virtual odor that reflects the proximity to the goal. The animal uses the same policy of spatial sampling that it would apply to a real odor signal, consistent with the idea that endotaxis is built on the ancient behavioral module for chemotaxis.

Frequently a rodent stopped at a maze junction merely turns its head side-to-side, rather than walking down a corridor to sample the gradient. Within the endotaxis model, this could be explained if some of the point cells in the lowest layer (Fig 1B) are selective for head direction or for the view down a specific corridor. During navigation, activation of that "direction cell" systematically precedes activation of point cells further down that corridor. Therefore the direction cell gets integrated into the map network. From then on, when the animal turns in that direction, this action takes a step along the graph of the environment without requiring a walk in ultimately fruitless directions. In this way the agent can sample the goal gradient while minimizing energy expenditure.

The vicarious trial and error movements are commonplace early on during navigation in a new environment. Later on the animal performs them more rarely and instead moves smoothly through multiple intersections in a row [13]. This may reflect a transition between different modes of navigation, from the early endotaxis, where every action gets evaluated on its real-world merit, to a mode where many actions are strung together into behavioral motifs. At a late stage of learning the agent may also develop an internal forward model for the effects of its own actions, which would allow for prospective planning of an entire route. An interesting direction for future research is to

seek a neuromorphic circuit model for such action planning; perhaps it can be built naturally on top of the endotaxis circuit.

While rodents engaged in early navigation act as though they are sniffing out a virtual odor, we would dearly like to know whether the experience *feels like* sniffing to them. The prospects for having that conversation in the near future are dim, but in the meantime we can talk to humans about the topic. Human language has an intriguing set of metaphors for decision making under uncertainty: "this doesn't smell right", "sniff out a solution", "that idea stinks", "smells fishy to me", "the sweet smell of success". All these sayings apply in situations where we don't yet understand the rules but are just feeling our way into a problem. Going beyond mere correlation, there is also a causal link: Fishy smells can change people's decisions on matters entirely unrelated to fish [17]. In the endotaxis model (Fig 1B) this might happen if the mode switch is leaky, allowing real smells to interfere with virtual odors. Perhaps this partial synesthesia between smells and decisions results from the evolutionary repurposing of an ancient behavioral module that was intended for olfactory search.

5.3 Brain circuits

The proposed circuitry (Fig 1) relates closely to some real existing neural networks: the so-called cerebellum-like circuits. They include the insect mushroom body, the mammalian cerebellum, and a host of related structures in non-mammalian vertebrates [18, 19]. The distinguishing features are: A large population of neurons with selective responses (e.g. Kenyon cells, cerebellar granule cells), massive convergence from that population onto a smaller set of output neurons (e.g. Mushroom body output neurons, Purkinje cells), and synaptic plasticity at the output neurons gated by signals from the animal's experience (e.g. dopaminergic inputs to mushroom body, climbing fiber input to cerebellum). It is thought that this plasticity creates an adaptive filter by which the output neurons learn to predict the behavioral consequences of the animal's actions [18, 20]. This is what the goal cells do in the endotaxis model.

The analogy to the insect mushroom body invites a broader interpretation of what purpose that structure serves. In the conventional picture the mushroom body helps with odor discrimination and forms memories of discrete odors that are associated with salient experience [21]. Subsequently the animal can seek or avoid those odors. But insects can also use odors as landmarks in the environment. In this more general form of navigation, the odor is not a goal in itself, but serves to mark a route towards some entirely different goal [22, 23]. In ants and bees, the mushroom body receives massive visual input, and the insect uses discrete panoramic views of the landscape as markers for its location [24–26]. Our analysis shows how the mushroom body circuitry can tie together these discrete points into a cognitive map that supports navigation towards arbitrary goal locations.

In this picture a Kenyon cell that fires only under a specific pattern of receptor activation becomes selective for a specific location in the environment, and thus would play the role of a map cell in the endotaxis circuit (Fig 1). [3] After sufficient exploration of the reward landscape the mushroom body output neurons come to encode the animal's proximity to a desirable goal, and that signal can guide a trial-and-error mechanism for steering. In fact, mushroom body output neurons are known to guide the turning decisions of the insect [27], perhaps through their projections to the central complex [28], an area critical to the animal's turning behavior. Conceivably this is where the insect's basic chemotaxis module is implemented, namely the policy for ascending on a goal signal.

Beyond the cerebellum-like circuits, the general ingredients of the endotaxis model – recurrent synapses, Hebbian learning, many-to-one convergence – are found commonly in other brain areas including the mammalian neocortex and hippocampus. In the rodent hippocampus, an interesting candidate for map cells are the pyramidal cells in area CA3. Many of these neurons exhibit place fields and they are recurrently connected by synapses with Hebbian plasticity. It was suggested early on that random exploration by the agent produces correlations between nearby place cells, and thus the synaptic weights among those neurons might be inversely related to the distance between their place fields [29, 30]. However, simulations showed that the synapses are substantially strengthened only among immediately adjacent place fields [30, 31] (see also our Eqn 21), thus limiting the utility for global navigation across the environment. Here we show that a useful global distance function emerges from the *output* of the recurrent network (Eqns 24, 27, 28) rather than its synaptic structure.

[3] Point cells and Map cells are the same in this picture

Further, we offer a biologically realistic circuit (Fig 1B) that can read out this distance function for subsequent navigation.

5.4 Neural signals

The endotaxis circuit proposes three types of neurons – point cells, map cells, and goal cells – and it is instructive to compare their expected signals to existing recordings from animal brains during navigation behavior. Much of that prior work has focused on the rodent hippocampal formation [32], but we do not presume that endotaxis is localized to that structure. The three cell types in the model all have place fields, in that they fire preferentially in certain regions within the graph of the environment. However, they differ in important respects:

Size and location The place field is smallest for a point cell; somewhat larger for a map cell, owing to recurrent connections in the map network; and larger still for goal cells, owing to additional pooling in the goal network. Such a wide range of place field sizes has indeed been observed in surveys of the rodent hippocampus, spanning at least a factor of 10 in diameter [33, 34]. Some place cells show a graded firing profile that fills the available environment. Furthermore one finds more place fields near the goal location of a navigation task, even when that location has no overt markers [35]. Both of those characteristics are expected of the goal cells in the endotaxis model.

Dynamics The endotaxis model assumes that point cells exist from the very outset in any environment. Indeed, many place cells in the rodent hippocampus appear within minutes of the animal's entry into an arena [33, 36]. Furthermore, any given environment activates only a small fraction of these neurons. Most of the "potential place cells" remain silent, presumably because their sensory trigger feature doesn't match any of the locations in the current environment [37, 38]. In the endotaxis model, each of these sets of point cells is tied into a different map network, which would allow the circuit to maintain multiple cognitive maps in memory [29]. Finally a small change in the environment, such as appearance of a local barrier (Fig 4), can indeed lead to disappearance and appearance of nearby place cells [39].

Goal cells, on the other hand, are expected to appear suddenly when the animal first arrives at a memorable location. At that moment the goal cell's input synapses from the map network are activated and the neuron immediately develops a place field. This prediction is reminiscent of a startling experimental observation in recordings from hippocampal area CA1: A neuron can suddenly start firing with a fully formed place field that may be located anywhere in the environment [40]. This event appears to be triggered by a calcium plateau potential in the dendrites of the place cell, which potentiates the excitatory synaptic inputs the cell receives. A surprising aspect of this discovery was the large extent of the resulting place field, which would require the animal several seconds to cover. This was interpreted as a signature of a new plasticity mechanism that extends over several seconds [41]. Our endotaxis model has a different explanation for this phenomenon: The goal cell's place field extends far in space because it taps into the map network, which has already prepared a large place field prior to the agent finding the goal location. In this picture all the synaptic changes are local in time and space, and there is no need to invoke an extended time scale for plasticity.

5.5 Learning theories

Endotaxis has similarities with *reinforcement learning* (RL) [42]. In both cases the agent explores a number of locations in the environment. In RL these are called *states* and every state has an associated *value* representing how close the agent is to rewards. In endotaxis, this is the role of the virtual odor, represented by the activity of a goal neuron. The value function gets modified through the experience of reward when the agent reaches a valuable resource; in endotaxis this happens via update of the synapses in the goal network ($\mathbf{G}$ in Fig 1B). In both RL and endotaxis, when the animal wishes to exploit a given resource it navigates so as to maximize the value function. Over time that value function converges to a form that allows the agent to find the goal directly from every starting state. The exponential decay of the virtual odor with increasing distance from the target (Eqn 28) is reminiscent of the exponential decay of the value function in RL, controlled by the discount factor, γ [42].

In endotaxis much of the learning happens independent of any reinforcement. During exploration, the circuit learns the topology of the environment, specifically by updating the synapses in the map

network ($\mathbf{M}$ in Fig 1B). The presence of rewards is not necessary for map learning: Until a resource is found for the first time, the value function remains zero because the $\mathbf{G}$ synapses have not yet been established (Eqn 18). Eventually, when the goal is encountered, $\mathbf{G}$ is updated in one shot and the value function becomes nonzero throughout the known portion of the environment. Thus the agent learns how to navigate to the goal location from a single reinforcement (Fig 2). This is possible because the ground has been prepared, as it were, by learning a map. In animal behavior this phenomenon is called *latent learning*. Early debates in animal psychology pitched latent learning and reinforcement learning as alternative explanations [43]. Instead, in the endotaxis algorithm, neither can function without the other (see Eqn 18). In *model-based* reinforcement learning, the agent could learn a forward model of the environment and uses it to update a value function. A key difference is that endotaxis learns the distances between all pairs of states, and can then establish a value function after a single reinforcement, whereas RL typically requires an iterative method to establish the value function [44–46].

The neural signals in endotaxis bear some similarity to the so-called *successor representation* [47, 48]. This is a proposal for how the brain might encode the current state of the agent, intended to simplify the mathematics of time-difference reinforcement learning. Each neuron stands for a possible state of the agent. The activity of neuron j is proportional to the time-discounted probability that the agent will find itself at state j in the future. Thus, the output of the endotaxis map network (Eqns 6, 24) qualitatively resembles a successor representation. However there are some important differences: First, the successor representation depends not only on the structure of the environment, but on the optimal policy of the agent, which in turn depends on the distribution of rewards. Thus the successor representation must itself be learned through a reinforcement algorithm. There is agreement in the literature that the successor representation would be more useful if the model of the environment were independent of reward structure [49]; however, it is believed that "it is more difficult to learn" [47]. By contrast, the map matrix in the endotaxis mechanism is built from a policy of random exploration independent of the reward landscape. Second, no plausible biomorphic mechanism for learning the successor representation has been proposed yet, whereas the endotaxis circuit is made entirely from biologically realistic components.

5.6 Outlook

In summary, we have proposed a simple model for spatial learning and navigation in an unknown environment. It includes an algorithm, as well as a fully-specified neural circuit implementation. The model makes quantitative and testable predictions that match a diverse set of observations in behavior, anatomy, and physiology, from insects to rodents (Secs 5.2-5.4). Of course the same observables may be consistent with other models, and in fact multiple navigation mechanisms may be at work in parallel or during successive stages of learning. Perhaps the most distinguishing features of the endotaxis algorithm are its reliance on trial-and-error sampling, and the close relationship to chemotaxis. To explore these specific ingredients, future research could work backwards: First find the neural circuit that controls the random trial-and-error sampling of odors. Then test if that module receives a convergence of goal signals from other circuits that process non-olfactory information. If so, that could lead to the mode switch which routes one or another goal signal to the decision-making module. Finally, upstream of that mode switch lies the soul [50] of the animal that tells the navigation machinery what goal to pursue. Given recent technical developments we believe that such a program of module-tracing is within reach, at least for the insect brain.

A Supplement

The core function of the endotaxis network is to learn the distance between any two points in the environment starting from purely local connectivity. As the agent explores the graph of the environment, the point cells for two adjacent locations briefly fire together. This is the local event that drives synaptic learning in the map population. Eventually the map network learns the global structure of the graph. In particular, for any chosen goal node on the graph, the network computes a virtual odor signal that varies with the agent's location and declines monotonically with the distance from the goal. Using that distance function the agent can navigate to the goal node by the shortest path. In this section we explain how this global distance measure comes about. We start with an analytical result about computing distances on a graph, continue with a formal analysis of how the endotaxis network functions, and proceed to numerical experiments that supplement results in the text.

A.1 A neuromorphic function to compute the shortest distance on a graph

Finding the shortest path between all pairs of nodes on a graph is a central problem of graph theory, known as "all pairs shortest path" (APSP) [51]. Generally an APSP algorithm delivers a matrix containing the distances D_{ij} for all pairs of nodes. That matrix can then be used to construct the actual sequence corresponding to the shortest path iteratively. The Floyd-Warshall algorithm [52] is simple and works even for the more general case of weighted edges between nodes. Unfortunately we know of no plausible way to implement Floyd-Warshall's three nested loops of comparison statements with neurons.

There is, however, a simple function for APSP that operates directly on the adjacency matrix and can be solved by a recurrent neural network. Specifically: If a connected, directed graph has adjacency matrix A_{ij},

$$A_{ij} = \begin{cases} 1, \text{ if node } i \text{ can be reached from node } j \text{ in one step} \\ 0, \text{otherwise, including the } i = j \text{ case} \end{cases} \tag{1}$$

then with a suitably small positive value of γ the shortest path distances are given by

$$D_{ij} = \left\lceil \frac{\log\left[(\mathbf{1} - \gamma\,\mathbf{A})^{-1}\right]_{ij}}{\log\gamma} \right\rceil \tag{2}$$

where $\mathbf{1}$ is the identity matrix, and the half-square brackets mean "round up to the nearest integer".

Proof: The powers of the adjacency matrix represent the effects of taking multiple steps on the graph, namely

$$\left[\mathbf{A}^k\right]_{ij} = N_{ij}^{(k)} = \text{number of distinct paths to get from node } j \text{ to node } i \text{ in } k \text{ steps}$$

where a path is an ordered sequence of edges on the graph. This can be seen by induction as follows. By definition

$$N_{ij}^{(1)} = A_{ij}$$

Suppose we know $N_{ij}^{(k)}$ and want to compute $N_{ij}^{(k+1)}$. Every path from j to i of length $k+1$ steps has to reach a neighbor of node i in k steps. Therefore

$$N_{ij}^{(k+1)} = \sum_l A_{il} N_{lj}^{(k)} \tag{3}$$

The RHS corresponds to multiplication by $\mathbf{A}$, so the solution is

$$N_{ij}^{(k)} = \left[\mathbf{A}^k\right]_{ij}$$

We are particularly interested in the shortest path from node j to node i. If the shortest distance D_{ij} from j to i is k steps then there must exist a path of length k but not of any length $< k$. Therefore

$$D_{ij} = \min_k N_{ij}^{(k)} > 0 \tag{4}$$

Now consider the Taylor series

$$\begin{aligned} \mathbf{Y} &= (\mathbf{1} - \gamma\mathbf{A})^{-1} \\ &= \mathbf{1} + \gamma\mathbf{A} + \gamma^2\mathbf{A}^2 + \ldots \end{aligned} \tag{5}$$

Then

$$Y_{ij} = \sum_{k=0}^{\infty} N_{ij}^{(k)}\gamma^k = N_{ij}^{(D_{ij})}\gamma^{D_{ij}} + N_{ij}^{(D_{ij}+1)}\gamma^{D_{ij}+1} + \ldots \tag{6}$$

We will show that if γ is chosen positive but small enough then the growth of $N_{ij}^{(k)}$ with increasing k gets eclipsed by the decay of γ^k such that

$$\gamma^{D_{ij}} < Y_{ij} < \gamma^{D_{ij}-1} \tag{7}$$

The left inequality is obvious from Eqn 6 because $N_{ij}^{(D_{ij})} \geq 1$ by Eqn 4.

To understand the right inequality, note first that $N_{ij}^{(k)}$ is bounded by a geometric series. From Eqn 3 it follows that

$$N_{ij}^{(k)} < q^k$$

where q is the largest number of neighbors of any node on the graph. So from Eqn 6

$$Y_{ij} < (q\gamma)^{D_{ij}} + (q\gamma)^{D_{ij}+1} + \cdots = \frac{(q\gamma)^{D_{ij}}}{1 - q\gamma} \tag{8}$$

This expression is $< \gamma^{D_{ij}-1}$ (Eqn 7) as long as

$$\gamma < \frac{1}{q + q^{D_{ij}}} \tag{9}$$

In addition, because

$$D_{ij} < n \equiv \text{ number of nodes on the graph}$$

this is satisfied if one chooses γ such that

$$\gamma < \frac{1}{q + q^n} \tag{10}$$

With that condition on γ the inequality 7 holds and taking the logarithm on both sides leads to the desired result:

$$D_{ij} = \left\lceil \frac{\log Y_{ij}}{\log \gamma} \right\rceil$$

A.2 The goal signal in endotaxis

In later sections we show that Y_{ij} can be computed by the endotaxis network, and how the required synaptic weights can be learned from exploration on the graph. For reasons of practical implementation, the network does not operate on Y_{ij} directly but on the scalar products of the column-vectors in $\mathbf{Y}$, namely

$$E_{ij} = \text{“goal signal from node } j \text{ to } i\text{”} = \sum_k Y_{ki} Y_{kj} \tag{11}$$

To understand how that goal signal E_{ij} varies with distance one can follow arguments parallel to those that led to Eqn 6. Using the upper bound by the geometric series (Eqn 8) and inserting in Eqn 11 one finds again that it is possible to choose a γ small enough to satisfy

$$\gamma^{D_{ij}} < E_{ij} < \gamma^{D_{ij}-1} \tag{12}$$

Under those conditions the goal signal E_{ij} decays exponentially with the graph distance D_{ij}.

A.3 Regime of validity of the goal signal

The analytical arguments above all relied on choosing a very small γ. In numerical experiments we found that the exponential dependence of the goal signal E_{ij} on distance (Eqn 12) actually holds over a wide range of γ (Fig 5A).

As γ increases, one enters a regime where the systematic relationship to graph distance (Fig 5B) breaks down and the goal signal becomes non-monotonic: Comparing all node pairs throughout the graph one now finds many instances where the pair with a larger distance produces a stronger goal signal (Fig 5C). This happens because Eqn 12 is no longer satisfied. Nonetheless, it is still possible that an agent ascending on the goal signal gets all the correct local instructions to find the shortest path. To test this we asked whether the goal signal recommends the correct successor node: For every start node j and goal node i one finds the node connected to j with the highest goal signal. If that neighbor is always one step closer to i then navigation will be perfect.

Indeed we found an extended range of values for γ where the goal signal worked flawlessly for navigation between all pairs of nodes (Fig 5C). In this range the goal signal gives the correct turning instructions on a local level, even if it is not globally monotonic with distance across the entire graph. This behavior can also be seen in some of the simulations of random exploration (Fig 3B).

At higher γ values navigation begins to fail (Fig 5D-E). For an increasing number of start/goal pairs the agent gets trapped in a local maximum of signal before arriving at the goal.

Finally above a certain critical value γ_c the goal signal fails catastrophically (Fig 5F). There is a simple mathematical reason for this: Recall that the Taylor expansion (5) has a convergence radius of 1. That means all the eigenvalues of $\gamma\mathbf{A}$ must have absolute value < 1, which requires

$$\gamma < \gamma_c \equiv \frac{1}{\text{largest absolute eigenvalue of } \mathbf{A}} \tag{13}$$

Outside of that convergence radius the expression $(\mathbf{1} - \gamma\mathbf{A})^{-1}$ can no longer be interpreted as counting paths on the graph and therefore loses any connection to graph distance.

A.4 Model formulation

We formalized the endotaxis mechanism of Figure 1B as follows:

The environment is parcelled into a set of discrete locations in space that are sparsely connected to each other. The locations and connectors form a graph that is fully specified by the adjacency matrix A_{ij} (Eqn 1).

We treat neural processing using a textbook linear rate model [10]. Each node on the graph has a point cell corresponding to that location. The point cell fires at a rate of 1 when the agent's position j

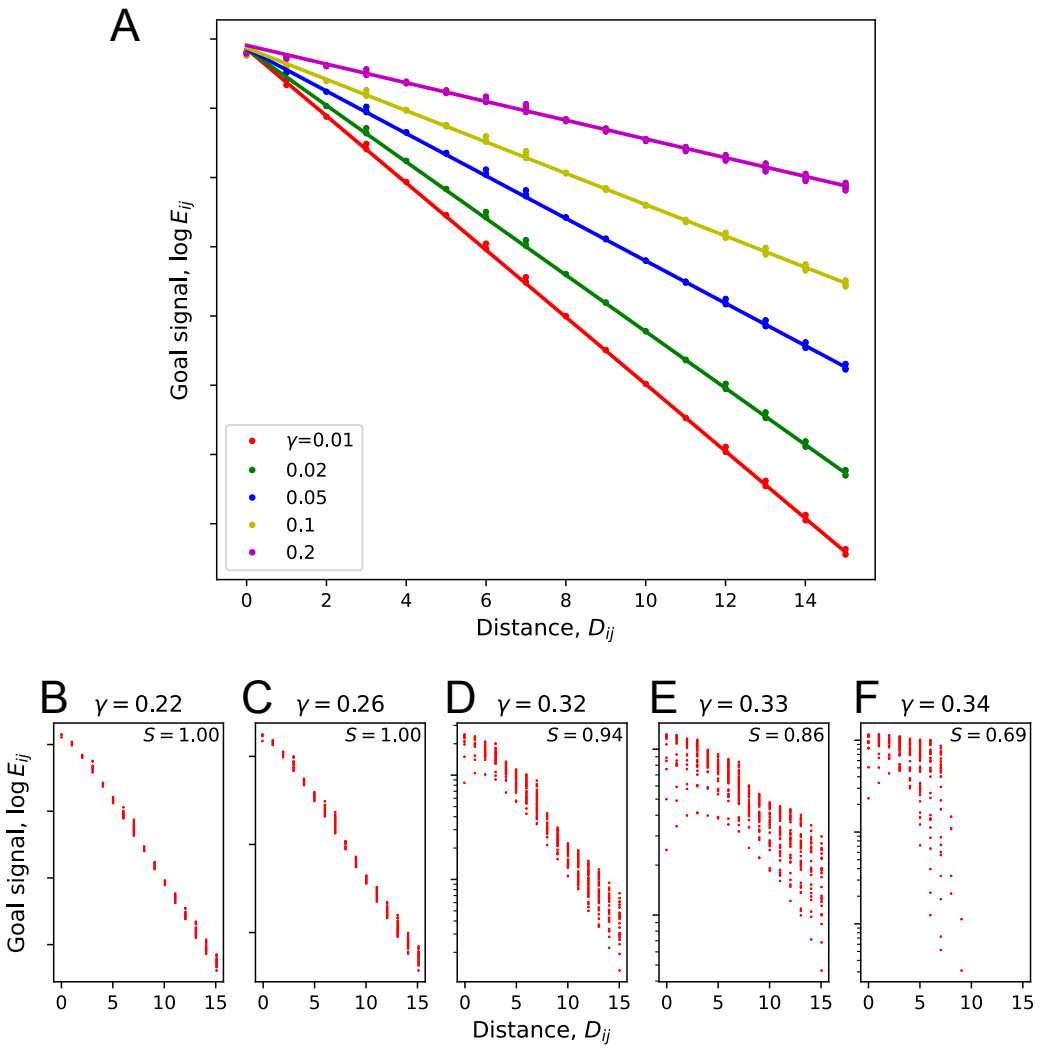

Figure 5: **The goal signal and the choice of** γ. **(A)** The goal signal declines exponentially with graph distance (the tower of Hanoi graph with 4 levels was used for these simulations). Data points indicate the goal signal between all pairs of nodes, computed with different values of γ, and plotted against the distance on the graph between the nodes. Lines are exponential fits to the data. **(B-F)** Detailed plot of goal signal vs distance as γ approaches the critical value γ_c, which for this graph is 0.335 (Eqn 13). The fraction of correct successors S is listed in each panel; as S drops below 1, the goal signal becomes less useful for navigation.

is at that node, and at a lower level w, with $0 < w < 1$, at the neighboring nodes. Thus the firing fields of neighboring point cells overlap somewhat; this produces correlations among point cells along the agent's trajectory which will drive synaptic plasticity.

$$u_i(x) = \text{firing rate of point cell } i \text{ with the agent at node } x \tag{14}$$
$$= \delta_{ix} + w\, A_{ix} \tag{15}$$

where δ_{ix} is the Kronecker delta. The output of the map network (Fig 1B) is

$$\mathbf{v} = \mathbf{u} + \mathbf{M}\mathbf{v} = (\mathbf{1} - \mathbf{M})^{-1}\mathbf{u} \tag{16}$$

where $\mathbf{u}$ is the vector of point cell outputs, $\mathbf{v}$ is the vector of map cell outputs, and $\mathbf{M}$ is the matrix of recurrent synapses among map cells.

A goal cell g receives sensory input s_g from neurons that signal the goal resource available to the agent at the current node:

$$s_g(y) = \text{amount of resource } g \text{ present when the agent is at node } y \tag{17}$$

In addition the goal cell gets input from the map neurons via the network of goal synapses. Thus the vector of goal cell activities with the agent at node x is

$$\mathbf{r}(x) = \mathbf{s}(x) + \mathbf{G}\,\mathbf{v}(x) = \mathbf{s}(x) + \mathbf{G}(\mathbf{1} - \mathbf{M})^{-1}\mathbf{u}(x) \tag{18}$$

The recurrent synapses among map cells undergo Hebbian plasticity. To keep the synaptic strengths bounded some normalization rule is needed. We adopted the standard Oja's Rule [10]:

$$\frac{\mathrm{d}M_{ij}}{\mathrm{d}t} = \beta_{\mathrm{M}}(\alpha_{\mathrm{M}} v_i v_j - M_{ij} {v_i}^2) \tag{19}$$

where β sets the speed of synaptic plasticity and α its strength. The map network has no self-synapses: $M_{ii} = 0$.

The synapses from map cells to goal cells also undergo Hebbian plasticity, again via Oja's Rule

$$\frac{\mathrm{d}G_{gi}}{\mathrm{d}t} = \beta_{\mathrm{G}}(\alpha_{\mathrm{G}} r_g v_i - G_{gi} {r_g}^2) \tag{20}$$

Because learning about targets is conceptually different from learning the map of the environment, we allowed $\alpha_{\mathrm{G}}, \beta_{\mathrm{G}}$ to differ from $\alpha_{\mathrm{M}}, \beta_{\mathrm{M}}$. Including the spatial overlap w, the model has 5 parameters.

A.5 How the endotaxis network learns the goal signal

Consider the linear rate model of the map network in Fig 1B and Eqns 16-19. It is well known that a Hebbian recurrent network of this type will learn the correlation structure of its inputs [10, 11]. Evaluating Eqn 19 after synapses have equilibrated leads to

$$M_{ij} = \alpha \frac{\langle v_i v_j \rangle}{\langle {v_j}^2 \rangle} \tag{21}$$

In the limit of small M_{ij}, i.e. if the inputs from point cells dominate, then $v_i \approx u_i$ and one gets to lowest order

$$M_{ij} \approx \alpha \frac{\langle u_i u_j \rangle}{\langle {u_i}^2 \rangle} = \alpha\, w\, A_{ij} \equiv \gamma\, A_{ij} \tag{22}$$

where

$$\gamma = \alpha\, w \tag{23}$$

In this approximation, the recurrent synapses M_{ij} directly reflect the connections among point cells and thus the adjacency matrix of the graph.

The output of the map network (Eqn 16) is

$$\mathbf{v} = (\mathbf{1} - \mathbf{M})^{-\mathbf{1}}\mathbf{u} = (\mathbf{1} - \gamma\mathbf{A})^{-\mathbf{1}}\mathbf{u} \tag{24}$$

So the recurrent network of map cells effectively computes the all-pairs distance function derived above (Eqn 5). If the agent is at node x then the map output $\mathbf{v}(x)$ equals the x-th column vector of the matrix $\mathbf{Y}$ (in the limit of small w and γ):

$$v_i(x) \approx Y_{ix} \tag{25}$$

which declines exponentially with the graph distance D_{ix} (Eqn 7). These distance-dependent humps of activity are schematized in Fig 1C.

The remaining problem is how to use the map output to encode the distance to a specific remembered goal location. Suppose goal g has a rewarding resource only at node y, specifically $s_g(x) = \delta_{xy}$ (Eqn 17). When the agent first arrives at location y, the synaptic plasticity rule (Eqn 20) updates the goal synapses G_{gi} from zero to a profile proportional to the current map output:

$$G_{gi} \sim v_i(y) \tag{26}$$

Subsequent visits will strengthen that profile. From then on, when the agent is at a location $x \neq y$ the virtual odor varies according to Eqn 18:

$$\begin{aligned} r_g(x) &= \mathbf{s}(x) + \mathbf{G}\,\mathbf{v}(x) \\ &\sim 0 + \mathbf{v}(y) \cdot \mathbf{v}(x) \equiv E_{xy} \end{aligned} \tag{27}$$

This corresponds to the goal signal E analyzed above (Eqns 11, 12, Fig 5). Thus the virtual odor computed by the endotaxis network decays exponentially with the agent's distance from the goal

$$E_{xy} \sim \gamma^{D_{xy}} \tag{28}$$

where $\gamma = \alpha\, w$.

The explanation here relied on multiple small-signal approximations. However, our simulations show that navigation based on the virtual odor signal is robust in realistic scenarios that include fully non-linear synaptic update rules and stochastic exploration by a random walk (Figs 2,3,4).

In this framework, the factor γ has an interesting interpretation. Its neural meaning is the strength of recurrent synapses in the map network compared to the feed-forward synapses from point cells (Eqn 22). Ultimately it determines the distance-dependence of the goal signal: For every step along the graph the goal signal declines by a factor of γ (Eqn 28). By analogy to the value function in reinforcement learning [42], one can identify γ as a discount factor or cost that the agent assigns for every step it has to take. This becomes relevant when the agent trades off two goal locations that offer rewards of different magnitude (Fig 4C): an additional step to one of the goals gets compensated if the reward is larger by a factor of $1/\gamma$. If the agent can manipulate γ, for example by varying α in Oja's plasticity rule (Eqns 19,22), that allows it to assign different costs on distance traveled (Fig 4C).

A.6 Limits and extensions of the endotaxis model

To help illuminate the remarkable phenomenon of rapid learning in a complex environment we sought an explanation in terms of biologically realistic processes. This informed the choice of modeling language, using concrete circuits of neurons and synapses, rather than abstract cognitive functions.

Furthermore we kept the model as simple as possible: the cells are single-compartment neurons without elaborate biophysics. The synapses are of a simple Hebbian type. All the input-output functions are linear. Free parameters are kept to the minimum: two each for the synaptic learning rules in the two networks. This simplicity allowed us to understand how and why the model works in analytical detail (Sec A.1-A.5).

Surprisingly this simplest possible model also learns very robustly in simulations over a range of environments. The parameters do not require careful tuning; in fact a single set of 4 numbers works fine for the conditions we studied. In some ways the simulations perform better than real animals. For example in the binary maze the agent can navigate to a reward location flawlessly after discovering it the first time (Fig 3B), whereas real mice solve that problem after $\sim$10 experiences [12]. This inspires confidence that as one adds realistic "bells and whistles" to the model the additional degrees of freedom will not break its operation. A number of extensions seem interesting for future work.

The distance function computed by the network fundamentally relies on the decay of neural activation over multiple synaptic links. In a large environment, and operating with a small γ, the virtual odor signal will span many orders of magnitude (Eqn 28). Real neurons cannot function reliably over such a large dynamic range, but some plausible additions could counteract the decay: A more realistic activation function with a compressive nonlinearity can amplify the signal locally in each neuron. Second, a short-term adaptive gain control might adjust the strength of synapses. In this way map cells far from the animal's current location could become more sensitive and continue to respond to the local trial-and-error movements of the agent.

Another desirable feature would be long-term memory. Animals can learn a cognitive map within minutes, and then retain it for days. Clearly there are multiple time scales for learning and forgetting. In complex brains one supposes that long-term consolidation is handled by transfer of the information between brain areas, for example hippocampus and cortex. Small insect brains don't offer that luxury, but perhaps the goal can be achieved within the endotaxis circuit itself, by endowing synapses with more complex dynamics [53].

A hierarchical extension of the model could be formulated such that an additional set of feedforward weights could read out from the goal signals in the current model formulation, which would allow for weighted preferences of desired goal features. Such a system could be useful for returning to locations with multiple properties that are desirable to the animal, or remembering a unique set of properties that characterize certain goal locations.

A.7 Simulations

Figures 2, 3, and 4 report the results of endotaxis learning while an agent explores the environment. We gave the agent a trajectory, either chosen by design (Fig 2) or as an unbiased random walk through the graph (Figs 3, 4). After every step of the random walk we computed the cell activities in a forward pass from point cells to goal cells. Then we updated the synaptic weights in the two networks $\mathbf{M}$ and $\mathbf{G}$ via a Hebbian learning rule. See Algorithm 1 for details. Matrix operations were implemented in JAX [54], but for the task complexity explored in this paper there was no need for GPU acceleration.

Learning and subsequent navigation worked robustly over a range of the α_{M} and β_{M} parameters in Oja's Rule (Fig 6). α_{M} has an absolute upper bound of γ_{c}/w (Eqns 13, 24) which depends on the eigenspectrum of the graph. In practice the Tower of Hanoi graph posed the strongest challenge, presumably because of its size and the large number of loops. For simplicity, we selected model parameters that allow for perfect navigation on that graph and applied the same model without modifications across all the tasks reported here. Note that this is not an exclusive set: smaller values for α_{M} and β_{M} would work as well.

A.7.1 Change in connectivity

To analyze changes in connectivity (Fig 4A.i-ii) we simulated an agent performing a random walk on a ring. At each time step we asked if the agent could navigate to the goal by the shortest path. We assumed that the appearance of a block or a shortcut between two adjacent nodes will alter the sensory cues around both locations (2 and 3 in Fig 4A.i-ii). Therefore the point cells that used to encode those locations drop silent, and the respective map cells lose their afferent input, while still remaining in the recurrent network. At the same time two new point cells appear at those locations, because the new cues match their selectivity. Their map cells now receive afferent input from the

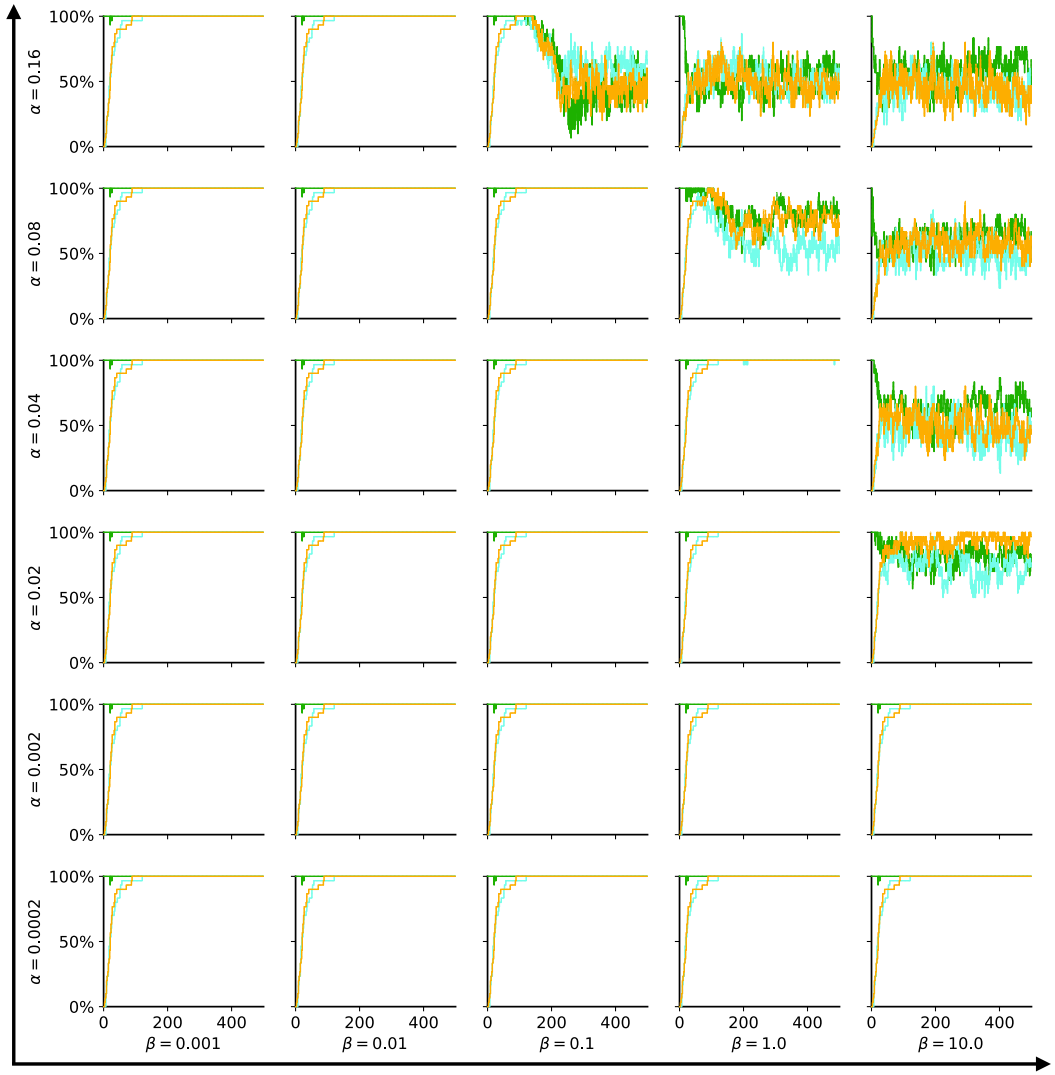

Figure 6: **Dependence of map learning on the parameters α_{M} and β_{M} in Oja's rule.** Each panel is for one combination of α_{M} and β_{M} and shows performance on the Gridworld task (Figs 2, 3 i). The fraction of successful navigations is plotted vs the number of steps in the exploratory random walk, averaged over 30 different walks. The 3 curves show navigation to the 3 goals, color coded as in Fig 3 i.

Algorithm 1 Online Learning via Oja's Rule

j : pre-synaptic neuron
i : post-synaptic neuron
$w = 0.3$ (fractional activity at neighbor nodes)
$s_g = 1$ (except dual-target tasks)
$\alpha_{\mathrm{M}} = 0.05$
$\beta_{\mathrm{M}} = 0.02$
$\alpha_{\mathrm{G}} = 0.5 \cdot \alpha_{\mathrm{M}}$
$\beta_{\mathrm{G}} = 0.03$

$\mathbf{M} \leftarrow 0$
$\mathbf{G} \leftarrow 0$

for *step t in node visit sequence* **do**
  Compute Neural Activity
  $\boldsymbol{u}_{node(t)} \leftarrow 1$
  for *each neighboring node i* **do**
    $\boldsymbol{u}_{node(i)} \leftarrow w$
  end for
  $\boldsymbol{u}_{node(others)} \leftarrow 0$
  $\boldsymbol{v} = \boldsymbol{u} + \mathbf{M}\boldsymbol{v} = (\mathbf{1} - \mathbf{M})^{-1}\boldsymbol{u}$
  $\boldsymbol{g} = \mathbf{G}\boldsymbol{v} + \boldsymbol{s}_{node(t)}$

  Synaptic Learning
  $M_{ij} \leftarrow M_{ij} + \beta_{\mathrm{M}}(\alpha_{\mathrm{M}} v_i v_j - M_{ij} v_i^2)$
  $G_{ij} \leftarrow G_{ij} + \beta_{\mathrm{G}}(\alpha_{\mathrm{G}} g_i v_j - G_{ij} g_i^2)$
end for

respective locations, but their recurrent synapses start at zero weight. The agent then continues a random walk around the ring, subject to the new constraints, and the learning algorithm proceeds as usual.

A.7.2 Dynamics of learning

Figure 7 illustrates the state of the synaptic networks over the course of online learning, as observed during a random walk on the binary maze graph (Fig 3A-ii). The norm of the map matrix $\|\mathbf{M}\|$ increases continuously through steady small updates $\|\mathrm{d}\mathbf{M}\|$. By comparison the goal matrix $\|\mathbf{G}\|$ increases in noticeable steps of $\|\mathrm{d}\mathbf{G}\|$ every time the agent visits a goal location. With sufficiently low α and β, the network learns stably and gradually approaches a steady state. However, as demonstrated in the text, even the first visit to a goal location already produces a goal signal that allows a reliable return to that location.

A.7.3 Robustness to noise

We tested how robust the map learning is to noise. Figure 8 illustrates the results using the Gridworld task (Fig 3-i). At each step of the simulation we perturbed each neuron's signal with multiplicative noise, by adding a Gaussian noise variable to the logarithm. Performance of learning and navigation was robust for signal-to-noise ratios of 2 or higher.

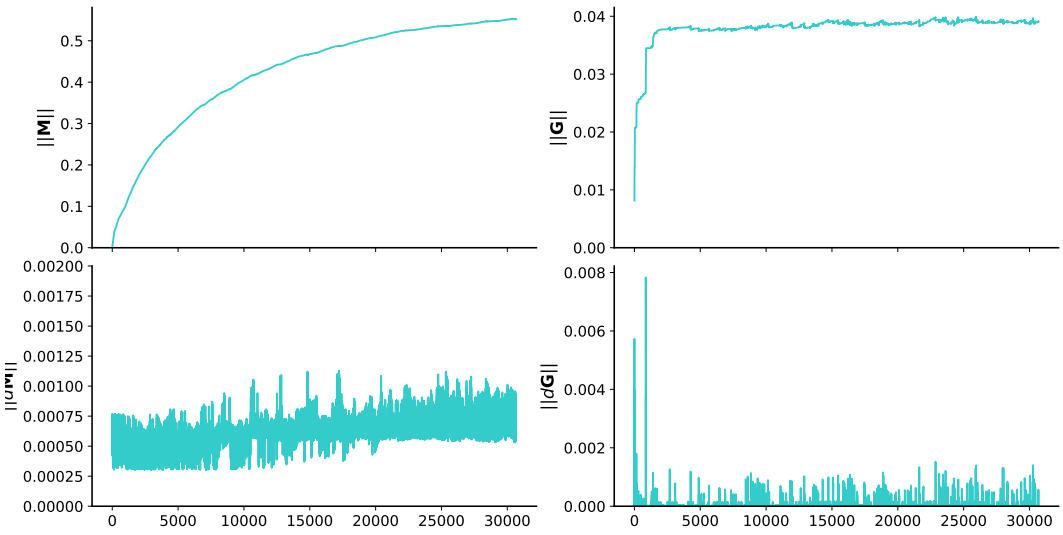

Figure 7: **Dynamics of online learning.** Evolution of the map matrix ($\|\mathbf{M}\|$ and $\|\mathrm{d}\mathbf{M}\|$) and the goal matrix ($\|\mathbf{G}\|$ and $\|\mathrm{d}\mathbf{G}\|$) during exploration of the binary maze graph of Fig 3A ii. See text for details.

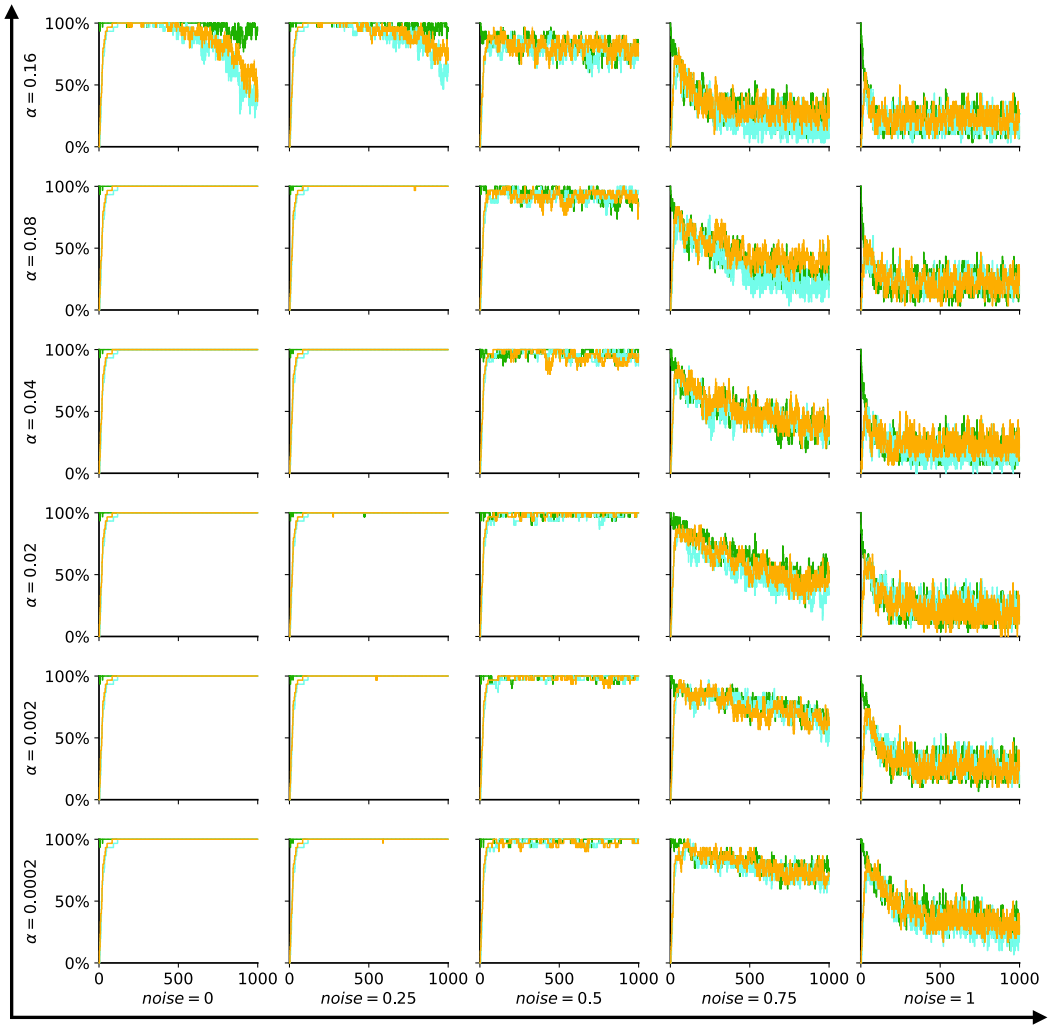

Figure 8: **Learning tolerates perturbation by neural noise.** Each panel shows navigation performance on the Gridworld task (Figs 2, 3 i), plotted as in Fig 6. Each neuron's activity was perturbed by multiplicative noise proportional to the unit's activity. The panels differ by the combination of α_{M} (rows) and noise level (columns). The noise level as a fraction of the unit's firing rate is listed below each column.

References

[1] R. G. M. Morris, P. Garrud, J. N. P. Rawlins, and J. O'Keefe. Place navigation impaired in rats with hippocampal lesions. *Nature*, 297(5868):681–683, June 1982. ISSN 1476-4687. doi: 10.1038/297681a0.

[2] Martin Müller and Rüdiger Wehner. Path integration in desert ants, Cataglyphis fortis. *Proceedings of the National Academy of Sciences*, 85(14):5287–5290, July 1988. ISSN 0027-8424, 1091-6490. doi: 10.1073/pnas.85.14.5287.

[3] Marielena Sosa and Lisa M. Giocomo. Navigating for reward. *Nature Reviews Neuroscience*, pages 1–16, July 2021. ISSN 1471-0048. doi: 10.1038/s41583-021-00479-z.

[4] Thomas S. Collett and Matthew Collett. Memory use in insect visual navigation. *Nature Reviews Neuroscience*, 3(7):542–552, July 2002. ISSN 1471-0048. doi: 10.1038/nrn872.

[5] Keeley L. Baker, Michael Dickinson, Teresa M. Findley, David H. Gire, Matthieu Louis, Marie P. Suver, Justus V. Verhagen, Katherine I. Nagel, and Matthew C. Smear. Algorithms for Olfactory Search across Species. *The Journal of Neuroscience*, 38(44):9383–9389, October 2018. ISSN 0270-6474. doi: 10.1523/JNEUROSCI.1668-18.2018.

[6] E. C. Tolman. Cognitive maps in rats and men. *Psychological Review*, 55(4):189–208, 1948. ISSN 0033-295X. doi: 10.1037/h0061626. WOS:A1948UY69500001.

[7] Lucia F. Jacobs. From chemotaxis to the cognitive map: The function of olfaction. *Proceedings of the National Academy of Sciences*, 109(Supplement 1):10693–10700, June 2012. ISSN 0027-8424, 1091-6490. doi: 10.1073/pnas.1201880109.

[8] Francisco Aboitiz and Juan F. Montiel. Olfaction, navigation, and the origin of isocortex. *Frontiers in Neuroscience*, 9, 2015. ISSN 1662-453X. doi: 10.3389/fnins.2015.00402.

[9] H. C. Berg. A physicist looks at bacterial chemotaxis. *Cold Spring Harb Symp Quant Biol*, 53 Pt 1:1–9, 1988.

[10] Peter Dayan and L. F. Abbott. *Theoretical Neuroscience: Computational and Mathematical Modeling of Neural Systems*. Computational Neuroscience. MIT Press, Cambridge, Mass., 2001.

[11] Mathieu Galtier, Olivier Faugeras, and Paul Bressloff. Hebbian Learning of Recurrent Connections: A Geometrical Perspective. *Neural computation*, 24:2346–83, May 2012. doi: 10.1162/NECO_a_00322.

[12] Matthew Rosenberg, Tony Zhang, Pietro Perona, and Markus Meister. Mice in a labyrinth exhibit rapid learning, sudden insight, and efficient exploration. *eLife*, 10:e66175, July 2021. ISSN 2050-084X. doi: 10.7554/eLife.66175.

[13] A. David Redish. Vicarious trial and error. *Nature reviews. Neuroscience*, 17(3):147–159, March 2016. ISSN 1471-003X. doi: 10.1038/nrn.2015.30.

[14] K. S. Lashley. Visual discrimination of size and form in the albino rat. *Journal of Animal Behavior*, 2(5):310–331, 1912. ISSN 0095-9928(Print). doi: 10.1037/h0071033.

[15] Michael Tarsitano. Route selection by a jumping spider (Portia labiata) during the locomotory phase of a detour. *Animal Behaviour*, 72(6):1437–1442, December 2006. ISSN 0003-3472. doi: 10.1016/j.anbehav.2006.05.007.

[16] Diogo Santos-Pata and Paul F. M. J. Verschure. Human Vicarious Trial and Error Is Predictive of Spatial Navigation Performance. *Frontiers in Behavioral Neuroscience*, 12:237, October 2018. ISSN 1662-5153. doi: 10.3389/fnbeh.2018.00237.

[17] Spike W. S. Lee and Norbert Schwarz. Bidirectionality, mediation, and moderation of metaphorical effects: the embodiment of social suspicion and fishy smells. *Journal of Personality and Social Psychology*, 103(5):737–749, November 2012. ISSN 1939-1315. doi: 10.1037/a0029708.

[18] Curtis C. Bell, Victor Han, and Nathaniel B. Sawtell. Cerebellum-like structures and their implications for cerebellar function. *Annual Review of Neuroscience*, 31:1–24, 2008. ISSN 0147-006X. doi: 10.1146/annurev.neuro.30.051606.094225.

[19] S. M. Farris. Are mushroom bodies cerebellum-like structures? *Arthropod Struct Dev*, 40: 368–79, July 2011. doi: 10.1016/j.asd.2011.02.004.

[20] D. M. Wolpert, R. C. Miall, and M. Kawato. Internal models in the cerebellum. *Trends in Cognitive Sciences*, 2(9):338–347, September 1998. ISSN 1364-6613. doi: 10.1016/S1364-6613(98)01221-2.

[21] Martin Heisenberg. Mushroom body memoir: From maps to models. *Nature Reviews Neuroscience*, 4(4):266–275, April 2003. ISSN 1471-0048. doi: 10.1038/nrn1074.

[22] Markus Knaden and Paul Graham. The Sensory Ecology of Ant Navigation: From Natural Environments to Neural Mechanisms. In M. R. Berenbaum, editor, *Annual Review of Entomology, Vol 61*, volume 61, pages 63–76. 2016. ISBN 978-0-8243-0161-3. doi: 10.1146/annurev-ento-010715-023703.

[23] Kathrin Steck, Bill S. Hansson, and Markus Knaden. Smells like home: Desert ants, Cataglyphis fortis, use olfactory landmarks to pinpoint the nest. *Frontiers in Zoology*, 6(1):5, February 2009. ISSN 1742-9994. doi: 10.1186/1742-9994-6-5.

[24] Barbara Webb and Antoine Wystrach. Neural mechanisms of insect navigation. *Current Opinion in Insect Science*, 15:27–39, June 2016. ISSN 2214-5745. doi: 10.1016/j.cois.2016.02.011.

[25] Cornelia Buehlmann, Beata Wozniak, Roman Goulard, Barbara Webb, Paul Graham, and Jeremy E. Niven. Mushroom Bodies Are Required for Learned Visual Navigation, but Not for Innate Visual Behavior, in Ants. *Current biology: CB*, 30(17):3438–3443.e2, September 2020. ISSN 1879-0445. doi: 10.1016/j.cub.2020.07.013.

[26] Xuelong Sun, Shigang Yue, and Michael Mangan. A decentralised neural model explaining optimal integration of navigational strategies in insects. *eLife*, 9:e54026, June 2020. ISSN 2050-084X. doi: 10.7554/eLife.54026.

[27] Y. Aso, D. Sitaraman, T. Ichinose, K. R. Kaun, K. Vogt, G. Belliart-Guerin, P. Y. Placais, A. A. Robie, N. Yamagata, C. Schnaitmann, W. J. Rowell, R. M. Johnston, T. T. Ngo, N. Chen, W. Korff, M. N. Nitabach, U. Heberlein, T. Preat, K. M. Branson, H. Tanimoto, and G. M. Rubin. Mushroom body output neurons encode valence and guide memory-based action selection in Drosophila. *Elife*, 3:e04580, 2014. doi: 10.7554/eLife.04580.

[28] Feng Li, Jack W Lindsey, Elizabeth C Marin, Nils Otto, Marisa Dreher, Georgia Dempsey, Ildiko Stark, Alexander S Bates, Markus William Pleijzier, Philipp Schlegel, Aljoscha Nern, Shin-ya Takemura, Nils Eckstein, Tansy Yang, Audrey Francis, Amalia Braun, Ruchi Parekh, Marta Costa, Louis K Scheffer, Yoshinori Aso, Gregory SXE Jefferis, Larry F Abbott, Ashok Litwin-Kumar, Scott Waddell, and Gerald M Rubin. The connectome of the adult Drosophila mushroom body provides insights into function. *eLife*, 9:e62576, December 2020. ISSN 2050-084X. doi: 10.7554/eLife.62576.

[29] Robert U. Muller, John L. Kubie, and Russ Saypoff. The hippocampus as a cognitive graph (abridged version). *Hippocampus*, 1(3):243–246, 1991. ISSN 1098-1063. doi: 10.1002/hipo.450010306. URL `https://onlinelibrary.wiley.com/doi/abs/10.1002/hipo.450010306`. _eprint: https://onlinelibrary.wiley.com/doi/pdf/10.1002/hipo.450010306.

[30] A. D. Redish and D. S. Touretzky. The role of the hippocampus in solving the Morris water maze. *Neural Computation*, 10(1):73–111, January 1998. ISSN 0899-7667. doi: 10.1162/089976698300017908.

[31] R. U. Muller, M. Stead, and J. Pach. The hippocampus as a cognitive graph. *The Journal of General Physiology*, 107(6):663–694, June 1996. ISSN 0022-1295. doi: 10.1085/jgp.107.6.663.

[32] May-Britt Moser, David C. Rowland, and Edvard I. Moser. Place Cells, Grid Cells, and Memory. *Cold Spring Harbor Perspectives in Biology*, 7(2):a021808, February 2015. ISSN 1943-0264. doi: 10.1101/cshperspect.a021808.

[33] M. A. Wilson and B. L. McNaughton. Dynamics of the hippocampal ensemble code for space. *Science (New York, N.Y.)*, 261(5124):1055–1058, August 1993. ISSN 0036-8075. doi: 10.1126/science.8351520.

[34] Kirsten Brun Kjelstrup, Trygve Solstad, Vegard Heimly Brun, Torkel Hafting, Stefan Leutgeb, Menno P. Witter, Edvard I. Moser, and May-Britt Moser. Finite scale of spatial representation in the hippocampus. *Science (New York, N.Y.)*, 321(5885):140–143, July 2008. ISSN 1095-9203. doi: 10.1126/science.1157086.

[35] Stig A. Hollup, Sturla Molden, James G. Donnett, May-Britt Moser, and Edvard I. Moser. Accumulation of Hippocampal Place Fields at the Goal Location in an Annular Watermaze Task. *Journal of Neuroscience*, 21(5):1635–1644, March 2001. ISSN 0270-6474, 1529-2401. doi: 10.1523/JNEUROSCI.21-05-01635.2001.

[36] Loren M. Frank, Garrett B. Stanley, and Emery N. Brown. Hippocampal Plasticity across Multiple Days of Exposure to Novel Environments. *Journal of Neuroscience*, 24(35):7681–7689, September 2004. ISSN 0270-6474, 1529-2401. doi: 10.1523/JNEUROSCI.1958-04.2004.

[37] Charlotte B. Alme, Chenglin Miao, Karel Jezek, Alessandro Treves, Edvard I. Moser, and May-Britt Moser. Place cells in the hippocampus: Eleven maps for eleven rooms. *Proceedings of the National Academy of Sciences*, 111(52):18428–18435, December 2014. ISSN 0027-8424, 1091-6490. doi: 10.1073/pnas.1421056111.

[38] J. Epsztein, M. Brecht, and A. K. Lee. Intracellular determinants of hippocampal CA1 place and silent cell activity in a novel environment. *Neuron*, 70:109–20, April 2011. doi: 10.1016/j.neuron.2011.03.006.

[39] Alice Alvernhe, Etienne Save, and Bruno Poucet. Local remapping of place cell firing in the Tolman detour task. *The European Journal of Neuroscience*, 33(9):1696–1705, May 2011. ISSN 1460-9568. doi: 10.1111/j.1460-9568.2011.07653.x.

[40] Katie C. Bittner, Aaron D. Milstein, Christine Grienberger, Sandro Romani, and Jeffrey C. Magee. Behavioral time scale synaptic plasticity underlies CA1 place fields. *Science*, 357 (6355):1033–1036, September 2017. ISSN 0036-8075, 1095-9203. doi: 10.1126/science.aan3846. URL `https://science.sciencemag.org/content/357/6355/1033`. Publisher: American Association for the Advancement of Science Section: Report.

[41] Jeffrey C. Magee and Christine Grienberger. Synaptic Plasticity Forms and Functions. *Annual Review of Neuroscience*, 43(1):95–117, 2020. doi: 10.1146/annurev-neuro-090919-022842. URL `https://doi.org/10.1146/annurev-neuro-090919-022842`. _eprint: https://doi.org/10.1146/annurev-neuro-090919-022842.

[42] Richard S. Sutton and Andrew G. Barto. *Reinforcement Learning: An Introduction*. MIT Press, November 2018. ISBN 978-0-262-03924-6.

[43] D. Thistlethwaite. A critical review of latent learning and related experiments. *Psychological Bulletin*, 48(2):97–129, 1951. ISSN 0033-2909. doi: 10.1037/h0055171.

[44] David Ha and Jürgen Schmidhuber. Recurrent world models facilitate policy evolution. In S. Bengio, H. Wallach, H. Larochelle, K. Grauman, N. Cesa-Bianchi, and R. Garnett, editors, *Advances in Neural Information Processing Systems*, volume 31. Curran Associates, Inc., 2018. URL `https://proceedings.neurips.cc/paper/2018/file/2de5d16682c3c35007e4e92982f1a2ba-Paper.pdf`.

[45] Richard S Sutton. Integrated architectures for learning, planning, and reacting based on approximating dynamic programming. In *Machine learning proceedings 1990*, pages 216–224. Elsevier, 1990.

[46] Thomas M Moerland, Joost Broekens, and Catholijn M Jonker. Model-based reinforcement learning: A survey. *arXiv preprint arXiv:2006.16712*, 2020.

[47] Peter Dayan. Improving generalization for temporal difference learning: The successor representation. *Neural Computation*, 5(4):613–624, July 1993. ISSN 0899-7667. doi: 10.1162/neco.1993.5.4.613.

[48] Kimberly L. Stachenfeld, Matthew M. Botvinick, and Samuel J. Gershman. The hippocampus as a predictive map. *Nature Neuroscience*, 20(11):1643–1653, November 2017. ISSN 1546-1726. doi: 10.1038/nn.4650.

[49] Jesse P. Geerts, Fabian Chersi, Kimberly L. Stachenfeld, and Neil Burgess. A general model of hippocampal and dorsal striatal learning and decision making. *Proceedings of the National Academy of Sciences*, 117(49):31427–31437, December 2020. ISSN 0027-8424, 1091-6490. doi: 10.1073/pnas.2007981117.

[50] Plato. The apology. ca 400 BCE.

[51] Uri Zwick. Exact and approximate distances in graphs — A survey. In Friedhelm Meyer auf der Heide, editor, *Algorithms — ESA 2001*, pages 33–48, Berlin, Heidelberg, 2001. Springer Berlin Heidelberg. ISBN 978-3-540-44676-7.

[52] Robert W. Floyd. Algorithm 97: Shortest path. *Communications of the ACM*, 5(6):345, June 1962. ISSN 0001-0782. doi: 10.1145/367766.368168.

[53] L. F. Abbott and W. G. Regehr. Synaptic computation. *Nature*, 431:796–803, October 2004. doi: 10.1038/nature03010.

[54] James Bradbury, Roy Frostig, Peter Hawkins, Matthew James Johnson, Chris Leary, Dougal Maclaurin, George Necula, Adam Paszke, Jake VanderPlas, Skye Wanderman-Milne, and Qiao Zhang. JAX: composable transformations of Python+NumPy programs, 2018. URL `http://github.com/google/jax`.

Data and code availability

Data and code to reproduce the reported results are available at `https://github.com/tonyzhang25/Zhang-2021-Endotaxis`. Following acceptance of the manuscript they will be archived in a permanent public repository.

Acknowledgments

Funding: This work was supported by the Simons Collaboration on the Global Brain (grant 543015 to MM and 543025 to PP), by NSF award 1564330 to PP, and by a gift from Google to PP.

Author contributions: Conception of the study TZ, MR, PP, MM; Numerical work TZ, PP; Analytical work MM; Drafting the manuscript MM; Revision and approval TZ, MR, PP, MM.

Competing interests: The authors declare no competing interests.

Colleagues: We thank Kyu Hyun Lee and Ruben Portugues for comments.