

Latent learning, shortcuts and detours: a computational model

Horatiu Voicu*, Nestor Schmajuk

Department of Psychological and Brain Sciences, Duke University, P.O. Box 90086, Durham, NC 27708-0086, USA

Received 1 November 2001; received in revised form 26 April 2002; accepted 28 April 2002

Abstract

Voicu and Schmajuk (Rob. Auto. Syst. 35 (2001a) 23) described a model of spatial navigation and exploration that includes an action system capable of guiding, with the help of a cognitive system, the search for specific goals as determined by a motivation system. Whereas in the original model the cognitive map stores information about the connectivity between places in the environment, in the present version the cognitive map also stores information about the paths traversed by the agent. Computer simulations show that the network correctly describes experimental results including latent learning in a maze, detours in a maze, and shortcuts in an open field. In addition, the model generates novel predictions about detours and shortcuts in an open field. © 2002 Elsevier Science B.V. All rights reserved.

Keywords: Computer simulations; Cognitive map; Spatial navigation; Computer simulations

1. Introduction

An important task for autonomous agents—including humans, animals and robots—is spatial navigation. A considerable number of models address this problem, including those presented by Keirsey et al. (1984), Schmajuk (1990), Schmajuk and Blair (1993), Payton (1990), Payton and Bihari (1991), Mataric (1991), Penna and Wu (1993), Muller et al. (1996), Reid and Staddon (1998), Trullier and Meyer (1998), Rowe and Alexander (2000) and Staddon (2001). Other models of spatial navigation have been reviewed in Trullier et al. (1997), Schmajuk (1998).

Among these models, Schmajuk and Thieme (1992) have offered a biologically plausible theory that includes (a) an action system consisting of a goal-seeking mechanism with goals set by a motivational system and (b) a cognitive system in which a neural network implements a cognitive map. The cognitive map represents the connectivity between places and the connectivity between places and goals. The goal-seeking mechanism changes from random exploratory behavior to approach behavior when either (a) the goal is found or (b) one place in the cognitive map generates a prediction of the goal that is stronger than the predictions generated by all other alternative places.

Recently, Voicu and Schmajuk (2001a,b) described a modified version of the Schmajuk and Thieme (1992) model. The new model differed from the original one in two ways. First, whereas

* Corresponding author. Tel.: +1-919-660-5659; fax: +1-919-660-5726

E-mail address: hv2@duke.edu (H. Voicu).

the early model assumed no a priori knowledge of the space to be explored, the modified model assumed a representation of the environment as a set of potentially connected and unexamined locations. Second, whereas in the original model the decision of what place to move next was based on the comparison of the predictions of the goal when each of the alternative places is briefly entered, in the new model this decision was based on the comparison of the activation of each of the alternative places when the goal is activated.

In this paper, we offer a model that differs from that presented by Voicu and Schmajuk (2001a,b) in that it differentiates between links that connect places that have not been traversed by the agent and those that have. These last types of links are part of what we refer to as trodden path.

We present computer simulations that illustrate how the modified model describes experimental results deemed to demonstrate animals' ability to use a cognitive map. First, latent learning in a maze, a protocol in which animals are exposed to a maze without being rewarded at the goalbox. When a reward is later presented, animals demonstrate knowledge of the spatial arrangement of the maze, which presumably remains latently stored in the cognitive map until reward is introduced. Second, detour problems in a maze, a protocol in which animals can choose one of two alternative detours only when, supposedly using the cognitive map, they integrate pieces of information learned separately. Third, shortcut problems in an open field, a protocol in which animals select the shortest path to a goal traversing previously unexplored regions of the environment, again apparently using a cognitive map.

In addition, we show novel predictions about a detour problem in an open field, a protocol in which animals avoid an obstacle placed on their path to the goal and then either retake their usual route or advance directly to the goal.

2. A model of spatial exploration

Fig. 1 shows the Schmajuk and Thieme (1992) model of spatial exploration and navigation with the modifications introduced by Voicu and

Schmajuk (2001a,b). The system entails (a) an action system including a goal-seeking mechanism with goals defined by (b) a motivation system, (c) a cognitive system including a neural network which implements a cognitive map, and (d) a short-term memory where the reading of the cognitive map is temporarily stored.

Once the motivation system defines a goal for the simulated animal (e.g. food for hunger, unexamined places for exploration), the action system starts the search for those goals. If the goal is perceived, the simulated animal moves towards it. If it is not perceived, but it can be predicted by the cognitive map, the simulated animal enters the place that best leads to the location of the goal. If the goal is neither perceived nor predicted, then the simulated animal engages in exploratory behavior using the cognitive map.

2.1. A canvas for the cognitive map

The map of the environment is drawn on an empty canvas that represents each location in the space to be mapped. Fig. 2 shows that the canvas is a lattice representing the potential continuity of space. In the figure, dashed-side squares represent places, the circles indicate their centers, and the solid lines linking the circles represent possible movements from one place to another. Although the places used here are square, places with arbitrary forms can also be used as long as they preserve the continuity of space. Places are of approximately the size of the footprint of the agent.

2.2. How locations are defined

We assume that the agent has perfect knowledge of its spatial location. This can be achieved through (a) dead-reckoning, in which the simulated animal adds its movements to its previous position to determine its current position (Golob and Taube, 1999) with limited accuracy as errors in direction and translation can easily accumulate, or (b) a positioning system that determines the simulated animal's location in space by making use of the distance to cues (landmark navigation) as

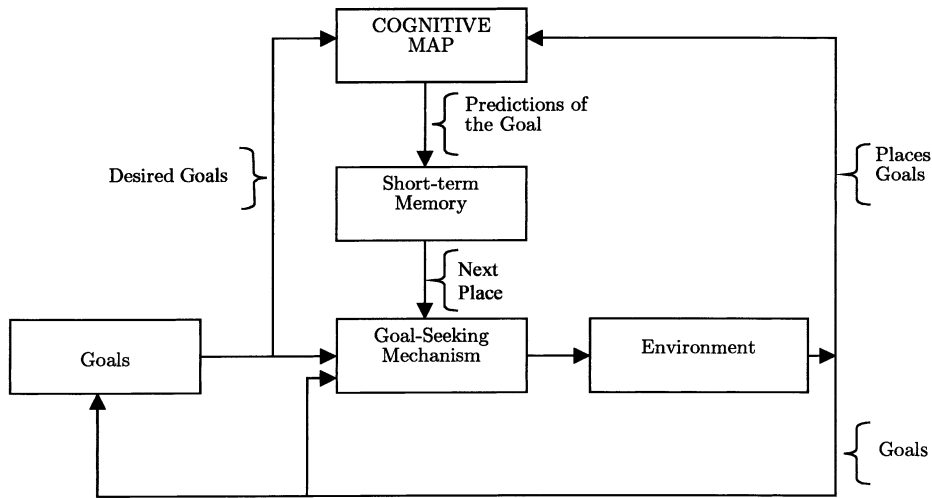


Fig. 1. A system for spatial exploration and navigation. Block diagram of the model showing the interaction between the action system (goal-seeking mechanism), motivation system (goal), cognitive system (cognitive map), short-term memory, and environment.

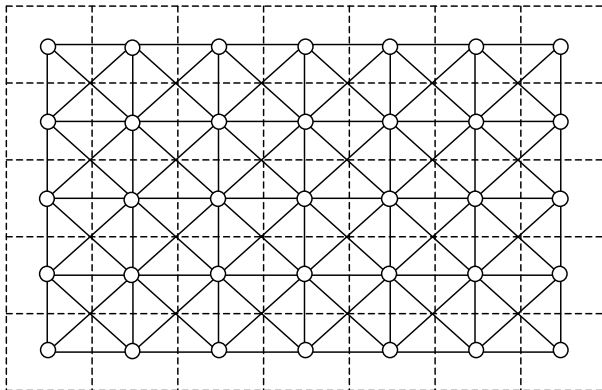


Fig. 2. The canvas. Squares in broken lines represent the places to be explored. Solid lines represent connections between places. The empty canvas is a lattice representing the potential continuity of the space to be explored. Adjacent places are assumed to be linked and each place is designated as unexamined (represented by an open circle.)

coded by their visual angles (Schmajuk, 1990; Zipser, 1985).

2.3. Exploration

During exploration, the motivation system designates all or some unexamined places as desired goals for the action system. Therefore, the agent moves to (a) an unexamined place when is next to

it, or (b) an adjacent place that best predicts, through the cognitive map, a still distant unexamined place. If (a) more than one neighboring place is unexamined, or (b) more than one place generates the same strongest prediction, places are entered according the following priorities: North, West, East, North-West, North-East, South, South-West, and South-East. These priorities define the pattern of exploratory movements carried out by the agent when most places are still unexamined.

As shown by Voicu and Schmajuk (2001a), the exhaustive coverage of an environment performed with the present model produces fewer number of moves than that achieved by random movements as described in Schmajuk and Thieme, (1992).

2.4. The cognitive map

Fig. 3 shows a hetero-associative network (Kohonen, 1977) capable, through recurrent connections, of cognitive mapping. The network includes three types of inputs: places, neighboring places, and goals. Places and neighboring places are defined by their position as represented in the canvas. Goals represent a feature (e.g. food, unexamined places) that, under the appropriate

2.7. Mapping the structure of the environment

Fig. 4 shows how an agent explores and maps the environment. As in Fig. 2, squares in broken lines represent the places to be explored. Unexamined places are indicated with an open circle. An obstacle occupies three adjacent places (1, 4, and 6). As indicated by the rays centered in the gray ellipse, the simulated animal (located in Place 0) can examine surrounding places 1, 2, 3, 4, 5, 6, 7, and 8. As indicated by the arrow, the simulated animal moves North from Place 0 to 2.

Since places 2, 3, 5, 7, and 8 become examined, they are represented by solid circles. Due to the obstacle, places 1, 4, and 6 remain unexamined, and are still represented by open circles. Connections between Place 0 and its surrounding places change to reflect the structure of the environment. Due to the obstacle, connections to inaccessible places 1, 4, and 6 will assume asymptotic values $V_{0,1} = 0$, $V_{0,4} = 0$, and $V_{0,6} = 0$. Since places 3, 5, 7, and 8 are accessible, connections $V_{0,3} = 1$, $V_{0,5} = 1$, $V_{0,7} = 1$, and $V_{0,8} = 1$ remain unchanged. Since the simulated animal moves from Place 0 to Place 2, the connection between these places will become part of the trodden path and assume asymptotic value $V_{0,2} = K$ (see Eq. (3)).

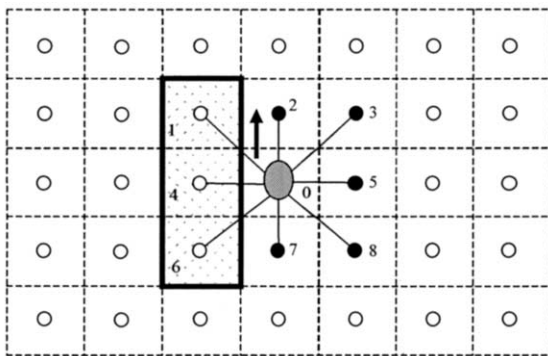


Fig. 4. Mapping the structure of the environment. Squares in broken lines represent the places to be explored. Open circles indicate unexamined places. Solid circles represent examined places. An obstacle occupies three adjacent places (1, 4, and 6). The simulated animal (represented by a gray ellipse in Place 0) examines surrounding places 1, 2, 3, 4, 5, 6, 7, and 8, and moves from Place 0 to 2.

2.8. Mapping the location of the goal

In addition to the associations reflecting the structure of the maze, associations between places j and goals h , $W_{j,h}$, and associations between goals h and places j , $W_{h,j}$, are also represented in the cognitive map. These associations are also stored in modifiable synapses, indicated by open triangles in Fig. 3. Whereas a positive $W_{j,h}$ association means that goal h can be found at place j , a positive $W_{h,j}$ association means that place j is to be entered to find goal h . When the agent is in place j , place j activates neurons goal h , and $g_h = p_j W_{j,h}$ indicates whether goal h is found at place j . Similarly, under a given motivation g_h are activated, and $a_j = g_h W_{h,j}$ indicates whether place j contains goal h .

2.9. Map reading

The cognitive map combines multiple associations $V_{i,j}$ to infer spatially remote goal locations. This is achieved by recurrently reinjecting the signal representing neighboring place j (as predicted by place i according to $V_{i,j}$) into the representation of current place j . Current place j now predicts neighboring place k according to $V_{j,k}$, and the signal representing neighboring place k is reinjected into the representation of current place k . At each reinjection the signal representing a neighboring place is attenuated, for instance, in half. The activation spreads using a breadth first technique. Places that are closer to the activated place have a higher activation than places that are further from that place. Once a place is activated its activation remains constant until the process is halted. Since the connections between obstacles and adjacent places are equal to zero the activation does not spread through obstacles. The process is halted when the representation of the present location of the agent is eventually activated.

If several goals equally salient are active at one time then the activity spreads at the same time from all goal locations determining the model to navigate to the closest goal. If several goals equally distant from the current position of the agent are active at one time then, activation spreads at the

same time from all goal locations determining the model to navigate to the most salient goal.

2.10. Decision making and navigation

Once a goal is determined by the motivation system, activation of the goal determines the activation of the place where the goal is located. This activation is spread through the network as described above, until the place where the simulated animal is located becomes active. The activation of all the places are stored in a short-term memory (see Fig. 1), and the simulated animal now can move towards the goal guided by a gradient ascent rule, that always chooses the neighboring place with the strongest activation.

2.11. Motor learning

Although the description of the path in terms of successive places to be approached is the most flexible way to store the information, eventually, when each place is associated with a specific movement, the sequence of places can be transformed in a sequence of movements. Presumably, with extended practice, the sequence of movements can be generated without sensory input. This latter aspect of the model has not been implemented.

2.12. Reward learning procedure

We assume that when the agent is under a given motivation (h = hunger) and looking for a specific reward (food), all place–goal and goal–place associations are 1 ($W_{j,h} = W_{h,j} = 1$) and therefore, the agent will attempt to visit all unexamined places in search for that reward. This assumption is justified by experimental data showing that hunger might facilitate exploratory behavior (Bolles and De Lorge, 1962; Fehrer, 1956; Zimbardo and Miller, 1958).

If no reward is found in a newly examined place, then $W_{j,h}$ and $W_{h,j}$ decrease gradually from 1 to 0 as follows

$$\Delta W_{j,h} = \Delta W_{h,j} = \alpha p_j (G - z_h), \quad (1)$$

where α is a learning constant, p_j is the representation of place j , z_h is the prediction of goal h ($z_h =$

$W_{j,h} p_j$), and $G = 1$ when the goal is present and 0 otherwise. If reward is found in place j , $W_{j,h}$ and $W_{h,j}$ may stay unchanged.

Notice that under these assumptions and with the reward values shown in Appendix A ($G = 1$), an unexamined place j with $W_{j,h} = 1$ might be as attractive as a rewarded place k with $W_{k,h} = 1$. By modifying the reward values (e.g. $G = 2$), an unexamined place can be less attractive than an examined and rewarded place.

2.13. Connections between places

In the neural network shown in Fig. 3, the connections $V_{i,j}$ and $V_{j,i}$ between the neuron representing place i (where the simulated animal is located) and each neuron representing neighboring places j are updated as follows

$$\begin{aligned} V_{i,j} = V_{j,i} &= 1 \quad \text{if } j \text{ can be accessed from } i, \\ V_{i,j} = V_{j,i} &= 0 \quad \text{otherwise} \end{aligned} \quad (2)$$

2.14. Trodden path

The connections $V_{i,j}$ and $V_{j,i}$ between the neuron representing place i (where the simulated animal is located) and the neuron representing place j that the simulated animal decides to enter are updated using Eq. (3). If place j is accessed from place i then

$$\Delta V_{i,j} = \Delta V_{j,i} = \beta (K - V_{j,i}), \quad (3)$$

where β is a learning constant, $K > 1$ when the simulated animal moves from place i to place j and $K = 1$ when j is accessible from place i but the simulated animal does not move to that location. K represents the upper bound for $V_{i,j}$ and $K \in (1, 1/\eta)$, see Eq. (5). Since $K > 1$ the activation is less attenuated on the trodden path than regular connections (for example see Fig. 13B).

2.15. Map reading

At each place where the simulated animal is located, the motivation system activates the desired goal (e.g. unexamined places or food), and

the goal activates place i where it is found,

$$I_i = W_{h,i} g_h \quad (4)$$

This activation spreads through the network by using

$$a_j(t+1) = \eta \sum_i V_{i,j} a_i(t), \quad (5)$$

where $\eta \in (0, 1)$ is the attenuation factor and t represents time, a_i represent the activation of neuron i and $V_{i,j}$ represents the association between place i and place j .

When the location of the agent is activated or the number of iterations is greater than a maximum number of iterations, R , the spreading of activation stops.

2.16. Decision making

The activation of all neurons represents a gradient that is used for planning a path to the closest goal. For each place m , this is accomplished by choosing

$$\text{Next place} = \arg \max_n a_n, \quad (6)$$

where n denotes a neighbor of place m .

Eq. (6) ensures that the agent will move in the direction of the nearest goal if all goals are of the same magnitude. When several neighboring places have an identical goal value, priorities are used to decide the next place. Priorities are given in the following order: North, West, East, North-West, North-East, South, South-West, and South-East.

At each move, the cognitive system updates the representation of the environmental configuration and selects the next place j to be entered. During the decision process, each place j contiguous to the place occupied by the simulated animal, place i , is examined. The place that is most activated by the activation of the desired goal is entered and a new move begins.

In sum, an agent looking for a specific reward will explore the environment by visiting all places and decreasing the attractiveness of those places where reward is not found. When reward is found in a given place, the attractiveness of that place remains unchanged or increases.

3. Computer simulations

This section shows computer simulations illustrating how the model presented in the previous section describes experimental results regarding latent learning, detours, and shortcuts. Parameter values used in the simulations are presented in [Appendix A](#).

3.1. Latent learning

Experimental data. [Blodgett \(1929\)](#) studied how non-rewarded trials affect performance when reward is later introduced in a multiple T-maze. A diagram of the places in the maze is shown in [Fig. 5](#).

Food-deprived rats received one trial a day, in which they were placed in the start box (Place 1) and retrieved after reaching the goal box (Place 19). Several groups of rats were trained. One group received food in the goalbox starting on the first trial. Two other groups were rewarded after three or seven non-rewarded trials. Blodgett found that after only one rewarded trial the performance of the initially non-rewarded groups improved to nearly the same level of the group rewarded on all trials.

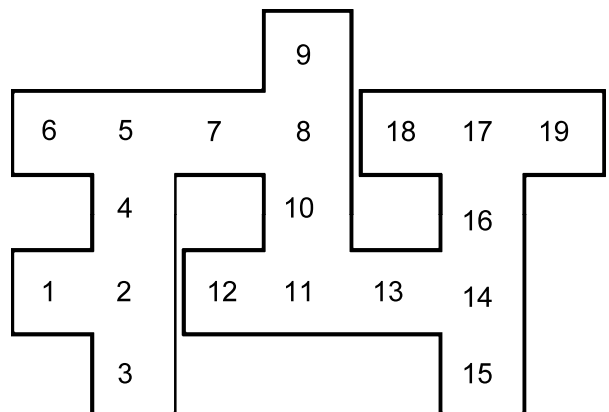


Fig. 5. Latent learning. Places in the multiple T-maze used in latent learning.

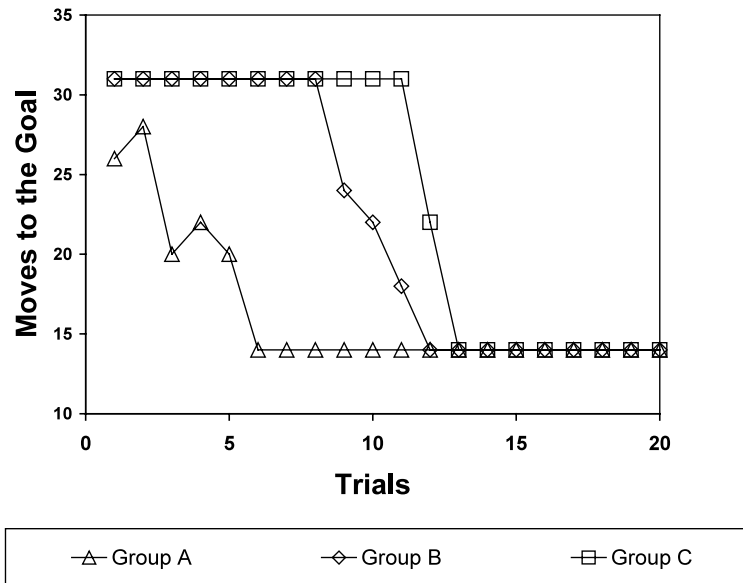


Fig. 6. Latent learning. Number of simulated moves to reach the goal in a multiple-T maze for three different groups. Group A is rewarded at the goalbox on the first trial, Group B on the eighth trial, and Group C on the twelfth trial.

3.2. Simulated results

Fig. 6 shows the number of moves to reach the goal as a function of trials for three different groups. Group A is rewarded at the goalbox on the first trial, Group B on the eighth trial, and Group C on the twelfth trial. Whereas Group A shows a gradual decrease in the number of steps, Group C shows a steep reduction once rewarded. According to Fig. 6, the model shows latent learning because when reward is presented after a period of latency, simulated animals with a pre-constructed cognitive map (Groups B and C) display rapid improvement in performance to the same level of Group A. These results are in accordance with Blodgett (1929) latent learning data. Similar results were obtained by Schmajuk and Thieme (1992), Voicu and Schmajuk (2001b).

Fig. 7A shows the contours of the maze overlapping the canvas used for cognitive mapping. Dashed-side squares represent places, the circles indicate their centers, the thin solid lines linking the circles represent possible movements from one place to another, and the thick solid lines the walls of the maze. The grid shown in Fig. 7A corre-

sponds to all place–place, place–goal, and goal–place associations set to 1 in the cognitive map shown in Fig. 3. By setting all goal–place and place–goal associations equal to 1, we assume that hungry simulated animals initially expect to find food at any place in the maze. These places are indicated by open circles in Fig. 7A.

3.3. Reward presented on the first trial

As the simulated animal moves through the environment, place–place associations are modified to reflect the maze layout. That is, if the simulated animal is in place i and place j is located behind a maze wall, then $V_{i,j}$ and $V_{j,i}$ associations are set to 0. If instead, place j can be accessed by the simulated animal, $V_{i,j}$ and $V_{j,i}$ associations are unchanged. In Fig. 7B, the solid lines linking the circles represent real connections from one place to another.

In addition, because the simulated animal does not find reward at any place with exception of the goalbox, $W_{j,h}$ and $W_{h,j}$ associations gradually decrease for all places with exception of the goalbox, whose associations with the goal remain

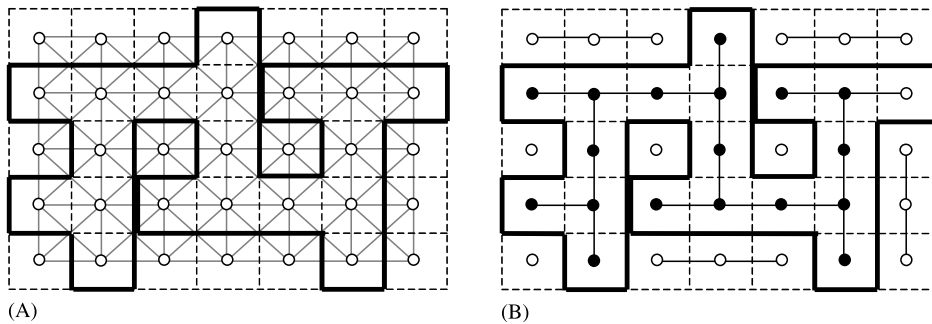


Fig. 7. Latent learning. (A) Walls of the maze superimposed on the empty canvas. All adjacent places are assumed to be linked and each place is designated unexamined and represented by an open circle. (B) The finished map is a lattice representing the actual continuity of the explored space.

unchanged. Solid circles in Fig. 7B indicate examined places, no longer attractive because food has not been found there. Notice that attractive places located outside the maze and that cannot be reached are still indicated by open circles. The open circle inside the maze in Fig. 7B indicates that the goalbox is still attractive because it contains food.

Since at the beginning of experiment most of the places are attractive, the simulated animal moves into the unexamined places and this is reflected in a larger number of steps (and a longer time) to the goal. This number gradually decreases as the attractiveness of the unrewarded places decreases by Eq. (1). When all non-rewarded places become relatively unattractive, the simulated animal moves directly to goalbox in the minimum number of steps (14).

The finished map illustrated in Fig. 7B is a lattice representing the actual continuity of the explored space. Only connected places inside the maze are linked and each place is designated examined and represented by a solid circle. The resulting links and places also constitute a graph representation of the maze, where places represent nodes and links represent edges in the graph.

3.4. Reward is not presented on the first trial

As the simulated animal moves through the environment $V_{i,j}$ and $V_{j,i}$ associations are modified to reflect the maze layout. In addition,

because the simulated animal does not find reward at any place, including the goalbox, all $W_{j,h}$ and $W_{h,j}$ associations are gradually decreased. If reward is not introduced before all $W_{j,h}$ and $W_{h,j}$ associations are reduced to 0, the simulated animal will cease to explore and, eventually, will try to reach unexplored places outside the maze. If reward is introduced at the goalbox when the simulated animal is still exploring, goalbox–goal and goal–goalbox associations will increase. If by the time the reward is introduced most of the $W_{j,h}$ and $W_{h,j}$ associations are close to 0, then the goalbox will be the only attractive place in the environment, and therefore, the simulated animal will go directly to the goalbox, as illustrated by Group C in Fig. 6. If, instead, $W_{j,h}$ and $W_{h,j}$ associations are not completely extinguished, then the goalbox will compete with other non-rewarded places, and therefore, the simulated animal will explore the unrewarded places that are farther from the goalbox before moving towards the goalbox, as illustrated by Groups B and C in Fig. 6.

Fig. 8A shows a view of the multiple T-maze shown in Fig. 5 on which lighter shades of gray represent the initially large attractiveness of each place based on being unexamined. Fig. 8B shows that during non-rewarded exploration of the maze, the attractiveness of all places decreases. Finally, Fig. 8C shows that when the goal is finally rewarded, its attractiveness increases.

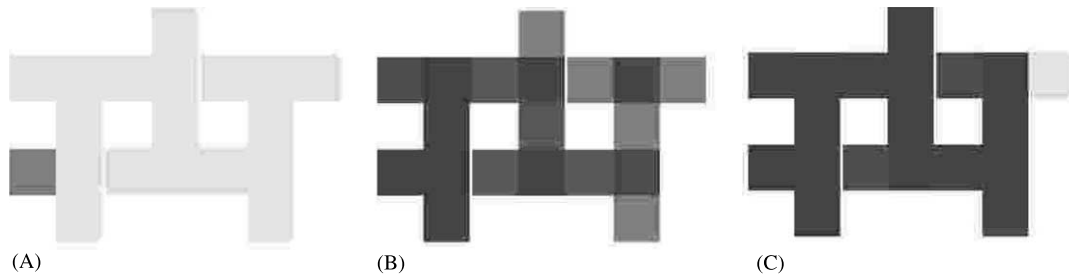


Fig. 8. Latent learning. View of the multiple T-maze shown in Fig. 5. (A) Shades of gray represent the attractiveness of each place based on being unexamined. All places are attractive except the starting place. (B) During non-rewarded exploration of the maze, the attractiveness of all places decreases. (C) When the goal is finally rewarded, its attractiveness increases. Lighter shades of gray represent higher attractiveness.

3.5. Detour problem

3.5.1. Experimental data

Tolman and Honzik (1930) used a maze with three paths, shown in Fig. 9A, in order to analyze how rats solve a detour problem. The shortest path, Path A, includes seven places (1, 2, 3, 4, 5, 6, and 7, the goalbox). The second shortest path, Path B, includes nine places (1, 2, 15, 16, 17, 18, 5, 6, and 7). The longest path, Path C, includes 11 places (1, 2, 8, 9, 10, 11, 12, 13, 14, 6, and 7). Paths A and B share a common place, 5, which Path C does not have. Fig. 1B shows the walls of the maze overlapping the canvas used for cognitive mapping.

During preliminary training, rats are alternately forced to take Paths A, B, and C, by introducing blocks in different segments of the maze. During testing, a block can be placed between places 4 and 5 (see Fig. 10), thereby obstructing only Path A. In

Tolman and Honzik's experiment, most rats, after backing out of Path A, took the second shortest path, Path B. Alternatively, a block can be placed between places 5 and 6 (see Fig. 11), thereby obstructing both paths A and B. In this case, after backing out of Path A, rats took the longest path, Path C, instead of the second shortest path, Path B.

3.5.2. Simulated results

In the simulation, the simulated animal is successively forced to take paths A, B, and C for ten trials each. When learned, Path A is traversed in seven moves, Path B in nine moves, and Path C in 11 moves. After the forced phase, the simulated animal is free to choose among all three paths, and chooses the shortest path, A. Fig. 10 shows that, when a block is introduced in Path A between places 4 and 5, the simulated animal immediately detours to the second shortest path, B.

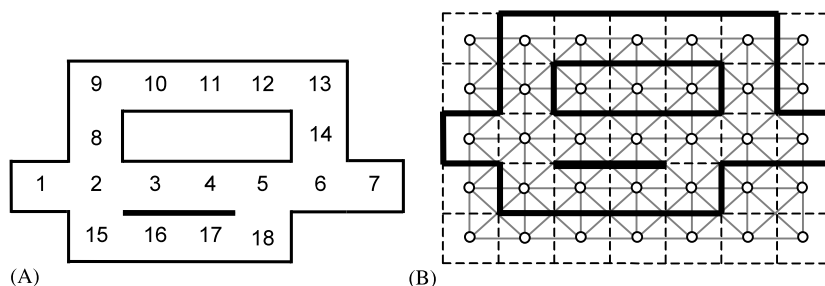


Fig. 9. Detour problems. (A) Walls and numbered places in the maze used to study detour problems. The shortest path, Path A, includes places 1, 2, 3, 4, 5, 6, and 7, the goalbox. The second shortest path, Path B, includes places 1, 2, 15, 16, 17, 18, 5, 6, and 7. The longest path, Path C, includes places 1, 2, 8, 9, 10, 11, 12, 13, 14, 6, and 7. (B) Walls of the maze superimposed on the empty canvas. All adjacent places are assumed to be linked and each place is designated unexamined and represented by an open circle.

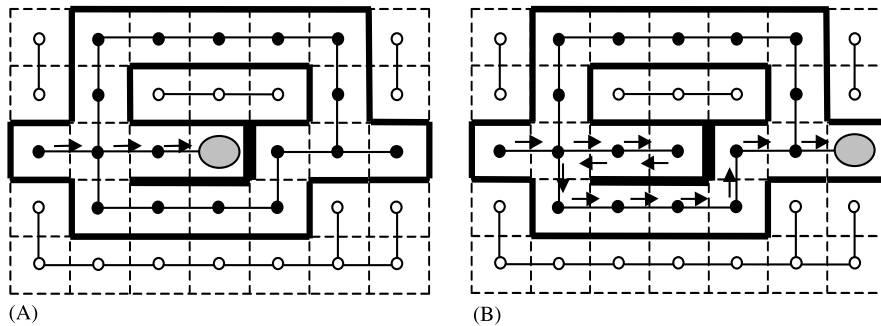


Fig. 10. Short detour problem. (A) The simulated animal is allowed to enter all pathways and chooses Path A. (B) When the block between places 4 and 5 is introduced, the simulated animal detours to Path B.

Fig. 11 shows that, when a block is introduced in Path A between places 5 and 6, the simulated animal immediately detours to the longest path, C. When the path is interrupted by an obstacle and the sequence of movements fails to take the agent to the goal, the system reverts to the goal-activation procedure to establish a new, unobstructed path to the goal. These simulated results are in agreement with Tolman and Honzik's data showing that rats, after backing out of Path A, take the longest path, C, and avoid path B which, though shorter, is also blocked.

3.6. Shortcut experiments

Using a sun-burst maze, Tolman et al. (1946) trained rats to obtain food in a stand connected to the start point by a series of elevated paths and a circular table. A light, which provided the only illumination to the room, was placed behind the goal. When the original path was blocked and a

choice among 18 different paths was presented on the circular table, 36% of the rats chose the path pointing directly toward the goal. The other rats chose the other paths in a chance fashion. One problem with this experiment, as discussed by Tolman et al. (1946), is that animals can be helped by the light to move in the direction of the goal when the original path is closed and the new one is open. If this is the case, the results are readily explained by our model. As shown in Fig. 1, whenever the system perceives the goal, it will move in its direction.

Using an open field, Chapuis (1988) rewarded dogs after moving them from a start point to two different locations. He showed that, when replaced at the start point, the animals would go first to the closest rewarded location and then would shortcut through unexplored territory to the second rewarded location. Similarly, in order to minimize the total distance traveled to retrieve hidden pieces of food, chimpanzees and vervet monkeys use a

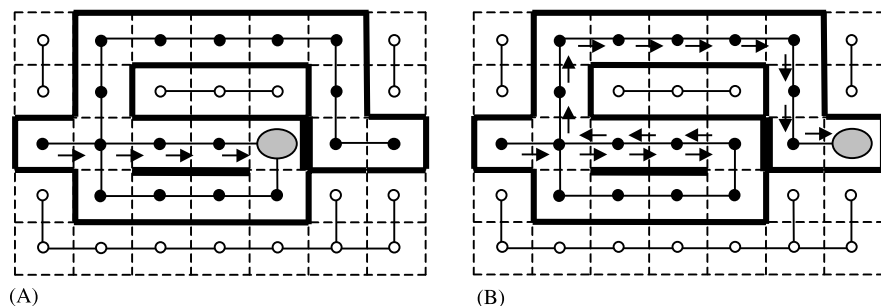


Fig. 11. Long detour problem. (A) The simulated animal is allowed to enter all pathways and chooses Path A. (B) When the block between places 5 and 6 is introduced, the simulated animal detours to Path C.

least-distance strategy, as described, respectively by Menzel (1973), Cramer and Gallistel (1997).

3.6.1. Simulated results

For the Chapuis (1988) experiment, the simulated animal is rewarded at two different locations, goal 1 and goal 2 (Fig. 12A and B) in an open field.

On the test trial (Fig. 12C), when replaced at the start point, the simulated animal goes first to the closest rewarded location (goal 1) and then takes a shortcut through unexplored territory to the second rewarded location (goal 2).

The results are explained in terms of the initial assumption that all locations in space are connected in the cognitive map, and therefore, the simulated animal can easily find the shortest path to goal 1. Once at the location of goal 1, the cognitive map shows that the shortest route to goal 2 is the shortcut goal 1–goal 2, and the simulated animal takes that path, instead of returning to the start point (start).

4. Discussion

The present paper describes latent learning, detour, and shortcuts with a neural network model for spatial navigation. The model assumes the following. First, the cognitive map uses an empty canvas that represents each place in the environment and the potential continuity of space. Second, before exploring an environment all adjacent places are connected. The advantage of this assumption is that it ensures an efficient and thorough mapping of the environment (see Voicu

and Schmajuk, 2001a). Similar approaches have been applied in building autonomous agents (see Schultz, et al., 1999).

Second, we used a goal activation procedure by which the motivation system activates the desired goal, which spreads the activation over all the places in the cognitive map in a progressively attenuated way (places farther away are less activated than place closer to the desired goal). When all the place activations are stored in a short-term memory, the agent can navigate towards the desired goal by choosing the alternative place with the most active memory (gradient ascend). This approach is similar to that described by Deutsch (1960) and more recently used by Mataric (1991), Reid and Staddon (1998), Voicu and Schmajuk (2001b).

4.1. Description of existing data

4.1.1. Latent learning

We offer a description of latent learning that substantially differs from previous descriptions. For instance, Schmajuk and Thieme (1992) assumed that in the absence of a prediction of the reward, animals explore the maze by generating random (undirected) movements until (a) the cognitive map is built and (b) the reward is found. In contrast, the present paper assumed that maze exploration is controlled by movements directed to unexamined places according to established priorities or the empty cognitive map until (a) the cognitive map reflects the design of the maze, (b) the reward is found, and (c) the attractiveness of the unexamined places becomes smaller than the attractiveness of the rewarded place.

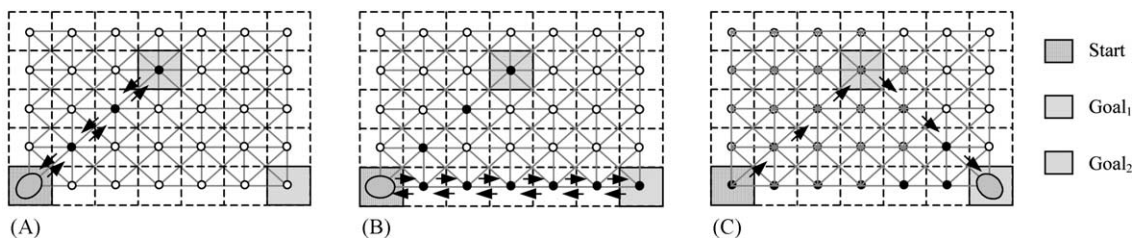


Fig. 12. Shortcut. (A) The simulated animal is trained to obtain reward at goal 1 and is returned to the start place. (B) The simulated animal is trained to obtain reward at goal 2 and is returned to the start place. (C) The simulated animal is placed at the start place and moves first to goal 1 and then shortcuts to goal 2, without returning to the start place.

Interestingly, because both unexamined places and rewarded places are specified as desired goals, the model describes the conflict between the tendency to explore the environment and the tendency to go to the food locations. In the context of reinforcement learning (Sutton and Barto, 1998), the problem of the tradeoff between exploration and exploitation has been described as that in which an agent must prefer actions that have been rewarded in the past but, in order to discover better actions, it has to try actions that have not been selected before. According to Sutton and Barto, the dilemma is that neither exploitation nor exploration can be pursued exclusively but, rather, the agent must try a variety of actions and favor the most rewarded ones. The relationship between exploration versus exploitation has been studied in animals. For instance, Kramer and Weary (1991) reported that when seed density was high, chipmunks returned promptly to the patch they were exploiting, but when seed density was reduced, chipmunks spent more time exploring.

An additional advantage of defining unexamined places as desired goals is that exploration is achieved in a more efficient and thorough manner than in the random case (Voicu and Schmajuk, 2001a).

The non-random exploration of space can be extended to the case of operant conditioning in which animals are rewarded for making one specific type of response (e.g. pressing a bar). Like in reinforcement learning, models of operant behavior (e.g. Schmajuk and Zanutto, 1998) assume that animals try different types of response at random until one is rewarded, which increases its probability of being emitted. According to our approach, all responses have initially a high probability of being generated and those unsuccessfully tried decrease their probability. The advantage of this approach is that methodical exploration of the alternative responses in the response space is more efficient and complete than random attempts and failures.

Finally, defining unexamined places as desired goals for the agent, captures some aspects of Berlyne (1950) theory of exploratory behavior. Berlyne proposed that exploration is controlled by

a curiosity drive, which is activated by novel stimuli. Similarly, in our model, unexamined places are by definition novel places that can be regarded as the goals of the curiosity drive. However, in contrast to Berlyne's view, the novelty of these unexamined places is independent of the organism perceiving them and only dependent on the animal's knowledge that those locations have not been occupied before.

4.1.2. *Detours*

According to the model, once the shortest path to the goal is found using the goal-activation method, the sequence of places to be entered is stored in a short-term memory, and the simulated animal can rapidly navigate towards the goal. This sequence of places can be transformed by the goal-seeking system (see Fig. 1) in a sequence of motor responses to perform at different places or even a sequence of motor responses generated without sensory input. When, as in the detour experiment, the path is interrupted by an obstacle and the sequence of movements fails to take the agent to the goal, the system reverts to the goal-activation procedure to establish a new, unobstructed path to the goal. The reestablishment of a new path to the goal with the help of the cognitive map is similar to the description of detour behavior advanced by Schmajuk and Thieme (1992).

4.1.3. *Shortcuts*

We describe shortcuts in terms of the a priori knowledge that the animal supposedly has about space and that is represented by the canvas illustrated in Fig. 2. According to this view, all points in space are potentially connected for the animal, and therefore, the attractiveness of a trodden path can propagate through these pre-established connections.

In contrast to this cognitive view, which assumes that the goal is not perceived by the agent, in some cases finding a shortcut is simply the result of interpreting sensory information. For example, Wehner and Menzel (1990), Dyer (1991) showed that honeybees take shortcuts between two foraging sites only when large landmarks of the goal site are visible from the starting site.

4.2. Novel predictions

In addition to the detours and shortcuts described above, we have analyzed a theoretical case for which experimental evidence is still lacking. The paradigm, loosely based on that described by Trullier et al. (1997), consists of placing an obstacle on the trodden path between a start place and a goal. After this path is blocked, animals follow the contour of the obstacle until reaching the other side of the barrier and then either (a) reach the closest place on the trodden path and continue navigation towards the goal, or (b) move straight to the goal (again, we assume that the goal is not detectable through sensory information).

According to Trullier et al. (1997), no spatial navigation model had successfully implemented both detours and shortcuts because they involve antagonistic processes. An animal attracted by the goal will not detour, an animal attracted by the trodden path will not shortcut to the goal. Here, we show that our model can describe both cases based on the experience of the animal. We tried to explain under which conditions animals would make a detour and retake the old path or find a shortcut to the goal. In our computer simulations, we attempted to account for these different behaviors in terms of the number of initial training trials (without an obstacle on the path). A shows the training trajectories between start and the goal.

Fig. 13A shows that the simulated animal moves straight to the goal after one trial of training. Fig. 12B shows that the simulated animal reaches a place on the trodden path removed from the obstacle. Finally, Fig. 13C shows that the simulated animal avoids the obstacle and retakes the trodden path after ten trials of training. According to the model, animals will take shortcuts to the goal when the attractiveness of points on the trodden path is weaker than the attractiveness of points on a shorter route to the goal (see Fig. 13A). This happens when the strength of the place–place connections on the trodden path is comparable to the strength of the initial connections between non-examined places in the environment. In contrast, animals will perform a detour around the obstacle when the attractiveness of points on the trodden path is stronger than the attractiveness of points on a shorter route to the goal (see Fig. 13C). In this case, the place–place connections on the trodden path are much stronger than the connections assumed to exist between non-examined places in the environment. In other words, our system takes a detour because it perceives a place on the trodden path more attractive than others in the environment. Once this place is found, the agent continues to move on the trodden path.

If the obstacle is introduced after only one training trial, then the simulated animal performs

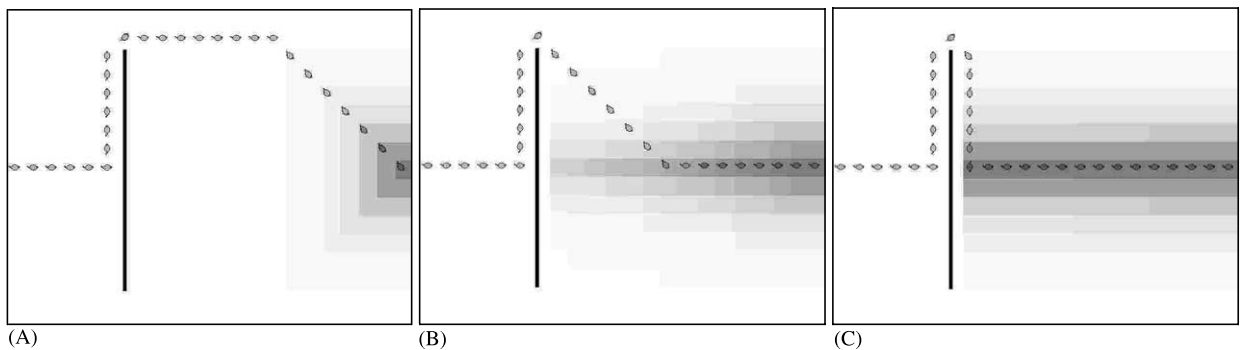


Fig. 13. Detour in open field. (A) Behavior on the first trial with the obstacle, after one trial of training. The simulated animal avoids the obstacle and shortcuts to the goal. (B) Behavior on the first trial with the obstacle, after five trials of training. The simulated animal avoids the obstacle and a while after passing the barrier regains the line of the trodden path. (C) Behavior on the first trial with the obstacle, after ten trials of training. The simulated animal avoids the obstacle and returns back to the trodden path. Shades of gray represent the activation of each place when the model performs gradient ascent to produce the path to the goal. The black thick line represents the obstacle.

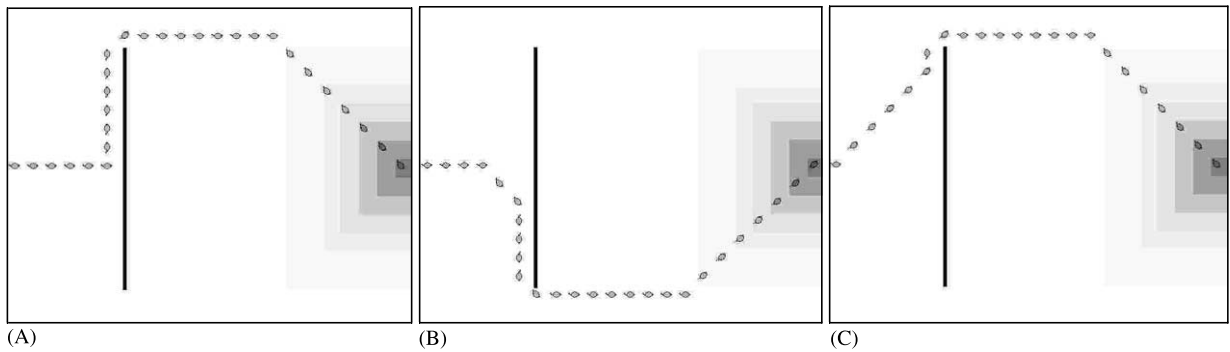


Fig. 14. Detour and shortcut with no initial training. (A) Behavior on the first trial with the obstacle, after one trial of training. The simulated animal avoids the obstacle by turning left and moves directly to the goal. (B) Behavior on the second trial with the obstacle, after one trial of training. The simulated animal avoids the obstacle by turning right and moves directly to the goal. (C) Behavior on the third trial with the obstacle, after one trial of training. The simulated animal avoids the obstacle by heading directly to the left-hand side of the obstacle and then moves directly to the goal. Shades of gray represent the activation of each place when the model performs gradient ascent to produce the path to the goal. The black thick line represents the obstacle.

as shown in Fig. 14A. During the next trial, the simulated animal gives a similar performance as in the first trial except that it deviates to the right instead of left (see Fig. 14B). This happens because the internal representation of the environment does not have a complete description of the obstacle (only the part on the left hand side has a representation in the cognitive map). Since the right hand side of the obstacle is not represented yet in the cognitive map, the simulated animal sees it as a shorter way to the goal as opposed to the path taken in the previous trial. At the end of the second trial the simulated animal has a complete representation of the obstacle. During the third trial the model follows the path shown in Fig. 14C. This shows that if the simulated animal has a complete representation of the obstacle, then it avoids the obstacle without bumping into it. The choice of deviating first to the left and then to the right depends on the assigned priorities (see Section 2).

4.3. Experimental data

There is no experimental data that particularly tests the use of trodden paths in animals. However, several experimental studies show that animals sometimes rely on cognitive maps and trodden paths to reach a given goal and sometimes rely on path integration and shortcuts to navigate to a

known place. As compared with dead-reckoning (see Section 2.2) path integration is defined as the process of updating continuously a vector that points towards the goal location (Schmidt et al., 1992). For example, Zeil and Layne (2002) showed that fiddler crabs base their navigation on path integration and visual input. They perform radial movements from and to the home burrow and when they decide to move to another sector of the environment they navigate to the burrow first. When their initial position is shifted they move in a straight line towards a position where they would find the burrow if their initial position had not been modified. This behavior clearly shows that fiddler crabs rely on path integration.

Regarding detours, fiddler crabs can solve them by relying either on visual information or solely on path integration. If they rely on visual information, they are able to avoid an obstacle before bumping into it. After circumventing the obstacle they rely on path integration and move straight to the goal. While they use visual information the path integration system is updated continuously. If, on the other hand, they rely solely on path integration, they first hit the obstacle and then, after circumventing it, they move straight to the goal (see Fig. 15A).

Schmidt et al. (1992) showed that, during foraging, desert ants use path integration by updating a vector that points towards their nest.

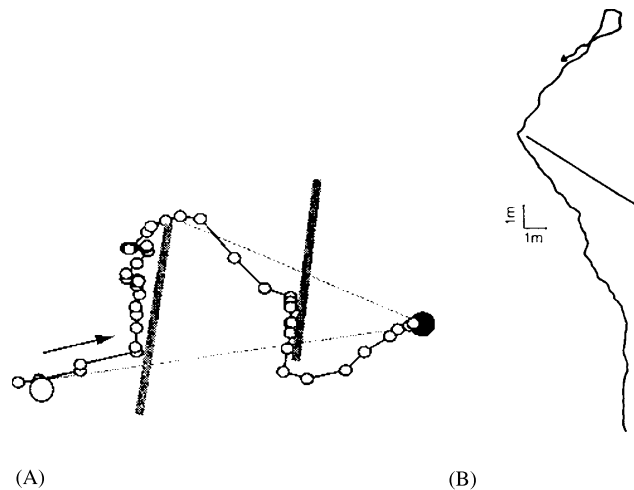


Fig. 15. (A) Detouring in fiddler crabs. Cardboard barriers (thick black lines) are placed between the crab and its burrow. Thin lines show the direction of the home vector. From Zeil and Layne (2002), (B) path followed by ant (*Cataglyphis bicolor*) when a barrier interrupts its path. From Schmidt et al. (1992).

They measured the angle taken by ants towards a goal location after circumventing a barrier (see Fig. 15B). The results show that the ants adapt their path integration vector depending on how far the barrier is from the goal location. Although these results clearly show that ants use path integration during foraging, they do not explain the variety of behaviors reported by Santschi (1913). As shown in Fig. 16 ants can return to the trodden path or perform a shortcut to the goal after the obstacle is circumvented. The data shown in Fig. 16 do not provide an explanation for path selection. Both behaviors (i.e. shortcut and detour) can be observed within the same species and in the same individual.

Another example of insects capable of performing detours is provided by Thorpe (1950). The digger wasp (*Ammophila pubescens*) while carrying its prey performs immediate detours once obstacles are placed on its way to the nest. As shown in Fig. 17 once the obstacle is circumvented the wasp attends the known path. Since the path selected does not point towards the goal location this clearly shows that the wasp does not rely on path integration. Instead it could rely on a trodden path or an odor trail that leads to the goal position.

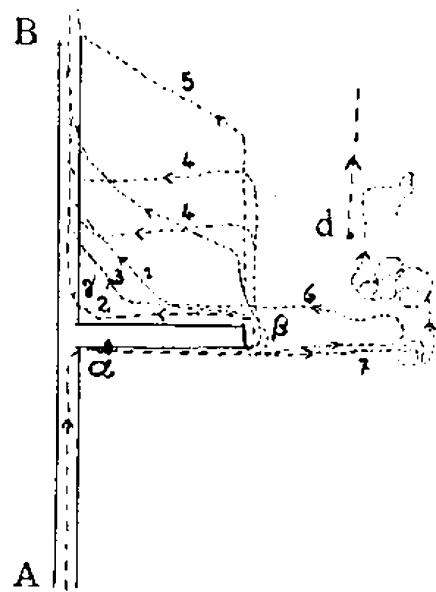


Fig. 16. Trajectories of different species of ants when they encounter a barrier along the way to a goal. From Santschi (1913).

Although our model performs similar behavior to that showed in Figs. 16 and 17, we cannot conclude that ants and wasps have cognitive maps. Instead of having an internal representation of the environment (i.e. a cognitive map) they could rely

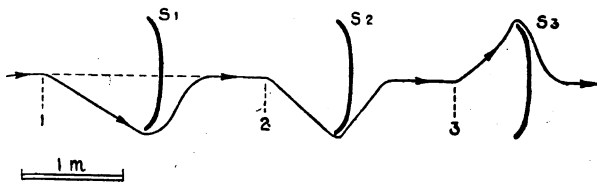


Fig. 17. Trajectory of the digger wasp (*Ammophila pubescens*) that pull its prey to the nest at about 10–15 m away. From Thorpe (1950).

on systems that can act, in some cases, like a cognitive map (e.g. odor trails in the case of the wasp).

4.4. Related work

Different models were tested using spatial tasks similar to those described above. For instance, a computational model that uses schema-based learning has been used by Corbacho and Arbib (1995) to simulate how frogs and toads learn to detour. A schema describes a control system that has inputs, outputs and an internal state. Multiple schemas can be combined to form coordinated control programs (Corbacho and Arbib, 1995) that are used to model behavioral data. An important advantage of this approach is the ability to provide a functional and structural decomposition of the modeled adaptive system. Corbacho and Arbib (1995) showed that frogs have the ability to perform detours when they are confronted with an array of equally spaced chopsticks. The frogs are able to perceive the size of the barrier and the location of the goal through the space between the chopsticks. They manage to circumvent the obstacle and reach the goal. Visual information plays a significant role in this experiment, therefore, one would predict that frogs do not have to rely on an internal representation of the environment to solve this task. The results are similar to those shown by our model in Fig. 13D where the simulated animal has a complete representation of the obstacle.

Muller et al. (1996) developed a theory of cognitive mapping that is based on known assumptions about hippocampal neurophysiology, including recurrent connections in CA3, long term

potentiation and firing properties of the place cells. The internal representation of the environment is built by strengthening or weakening the connections between places that have overlapping fields. Finding the shortest path between two places is done by executing a graph search. Both the starting place and the goal are activated and the search algorithm minimizes the sum of the synaptic connections along the path. According to Muller et al. (1996) the model is able to solve both shortcuts and detours. However, experiencing a certain path back and forth several times does not strengthen in particular the connections between adjacent places along that path. This means that the results of the search algorithm are not influenced by how many times the trodden path has been experienced. As a consequence, the model always takes a shortcut towards the goal after an obstacle has been avoided.

Reid and Staddon (1998) presented a model for spatial navigation that uses a diffusion process for guiding navigation. The environment is simulated by using a finite grid, each cell on the grid having assigned a value that denotes the probability of containing reward. In order to explore the environment the agent chooses the next place that has the highest probability value and after each movement is performed the distribution of reward is updated by using an equation of a diffusion process. The model focuses only on how the information stored in the cognitive map is used for guiding navigation. It does not provide details on how the cognitive map is built.

An important property of the model is that instead of being attracted to the nearest reward and thus minimizing the distance traveled, the agent moves towards the region of the environment with the highest value of reward. The agent performs an efficient exploration given that it knows the reward distribution in the environment. The model is able to simulate Chapuis (1988) shortcut experiment and to produce detour behavior for a U shaped obstacle. Since the model does not include a cognitive map, previous experiences (e.g. trodden paths) cannot shape future behavior. Therefore, when making a detour, the simulated animal will be attracted directly by the goal and therefore, move straight to it regardless of the

number of times the simulated animal experiences the path between the starting place and the goal location.

Trullier and Meyer (1998) presented a cognitive graph theory of the hippocampus that differs from the approaches proposed by Schmajuk and Thieme (1992), Muller et al. (1996) in that their cognitive map stores associations between two places and the direction needed to reach one place from the other. Due to the nature of how spatial information is stored, (e.g. place–direction–place) the model does not include a spreading activation mechanism for path planning. The experience of the animal is stored internally as a sequence of places. The model is able to show latent learning but since it is not built to deal with changing environments neither shortcuts nor detour simulations are presented by the authors.

Reinforcement learning models address the problem of behaving optimally and, in some cases, have used spatial navigation tasks as testbeds. There are many algorithms that have the ability to learn tasks using only reinforcement signals (e.g. Sutton, 1988, 1990; Watkins, 1989; Watkins and Dayan, 1992; Singh and Sutton, 1996). However, these methods are not able to show latent learning. Furthermore, they need a large number of trials to converge. Once the algorithm has converged the learned information is static. Any change in the environment is not readily reflected in the behavior of the agent. This translates in an inability to react promptly to shortcuts and detours. Furthermore, in order to be able to detect changes in the environment like shortcuts and detours the agent has to maintain a level of exploration that results in a deviation from optimal behavior. Once a change is detected another large number of trials are needed before the new optimal behavior is performed. Another limitation of this approach is that finding the optimal way to select action does not necessarily produce behavior similar to that of animals.

Witkowski (1997) proposed a dynamic expectancy model that includes a reinforcement learning system able to acquire knowledge without the need of an external goal. The model is based on a blend of theories concerning early stages of development in infants, cognitive mapping and reinforcement

learning. From the perspective of the model, these simulated animals are implementing a low-level scientific discovery process (Witkowski, 1997). They gather data, build hypotheses and perform experiments. All this processing is integrated in a graph-like structure. When goals need to be reached, they are added to the graph structure and appropriate actions are produced. In contrast to reinforcement learning techniques in which goals are needed for learning, the dynamic expectancy model sets its own internal goals by checking predictions generated by the hypothesis pool. The model shows latent learning in an experimental setup similar to that designed by Blodgett (1929). Since reinforcement is generated internally and is not dependent on external reward, the model is able to gather information about the environment and use it to its own benefit when an external goal becomes available. The model is also able to simulate the detour experiment performed by Tolman and Honzik (1930). Although the model is implemented using list manipulation, it is similar to our approach in that it uses a graph-like structure and spreading activation to produce action upon the environment.

5. Conclusion

The present paper presents a neural network model of spatial navigation characterized by three basic features. First, we assume that the environment is a collection of potentially connected locations that the animal is motivated to visit during exploration. Second, the connectivities between adjacent places are updated when the agent moves between these places. Third, the decision of which place to move next is based on the comparison of the activation of each of the alternative next places when the goal is activated. The model describes experimental data on latent learning, detours, and shortcuts. In addition, the model predicts the conditions in which animals will retake their usual route or advance directly to the goal after circumventing an obstacle placed on their path to the goal.

Appendix A: Parameter values

In Eq. (5), parameter η , which controls the attenuation at each reinjection in the network, was chosen ($\eta = .5$) to obtain an adequate signal at the neurons representing the location of the agent. Parameter R , the maximum number of reinjections, was selected ($R = 625$) to ensure that the representation of the location of the agent is activated from any place in the environment where the goal can be found.

In Eq. (3), the upper bound K , for the case when place j is accessed from place i , the rates of change in the associations between two places on the trodden path, α , and a place and reward, β , were chosen ($K = 1.5$, $\alpha = 0.2$, and $\beta = 0.01$) to adequately describe the experimental data.

References

- Berlyne, D.E., 1950. Novelty and curiosity as determinants of exploratory behavior. *British Journal of Psychology* 41, 68–80.
- Blodgett, H.C., 1929. The effect of the introduction of reward upon maze performance of rats. University of California Publications in Psychology 4, 113–134.
- Bolles, R.C., De Lorge, J., 1962. Effect of hunger on exploration in a familiar locale. *Psychological Reports* 10, 54.
- Chapuis, N., 1988. Les opérations structurantes dans la connaissance de l'espace chez les mammifères: Detour, raccourci et retour, Thèse de Doctorat d'Etat-es-Sciences, Université & d'Aix, Marseille II, France.
- Corbacho, F., Arbib, M.A., 1995. Learning to Detour. *Adaptive Behavior* 3 (4), 419–468.
- Cramer, A.E., Gallistel, C.R., 1997. Vervet monkeys as travelling salesmen. *Nature* 387, 464.
- Deutsch, J.A., 1960. *The Structural Basis of Behavior*. Cambridge University Press, Cambridge.
- Dyer, F.C., 1991. Bees acquire route-base memories but not cognitive maps in a familiar landscape. *Animal Behaviour* 41, 239–246.
- Fehrer, E., 1956. The effects of hunger and familiarity of locale on exploration. *Journal of Comparative and Physiological Psychology* 49, 549–552.
- Golob, E.J., Taube, J.S., 1999. Head direction cells in rats with hippocampal or overlying neocortical lesions: evidence for impaired angular path integration. *Journal of Neuroscience* 19, 7198–7211.
- Keirsey, D.M., Mitchell, J.S.B., Payton, D.W., Preys, E.P., 1984. Multilevel path planning for autonomous vehicles, applications of artificial intelligence. *SPIE Proceedings* 485, 133–137.
- Kohonen, T., 1977. *Associative Memory*. Springer, Berlin.
- Kramer, D.L., Weary, D.M., 1991. A field study of time allocation to environmental tracking by foraging chipmunks. *Animal Behaviour* 41, 443–449.
- Mataric, M.J., 1991. Navigating with a rat brain: a neurobiologically-inspired model of robot spatial representation. In: *From Animals to Animats*, vol. 1. MIT Press, Cambridge, MA.
- Muller, R.U., Stead, M., Pach, J., 1996. The hippocampus as a cognitive graph. *J. Gen. Physiol.* 107, 663–694.
- Menzel, E.W., Jr, 1973. Chimpanzee spatial memory organization. *Science* 182, 943–945 (01).
- Payton, D.W., 1990. Internalized plans: a representation for action. *Robotics and Autonomous Systems* 6, 89–103.
- Payton, D.W., Bihari, T.E., 1991. Intelligent real-time control of robotic vehicles. *Communications of the ACM* 34 (8), 49–62.
- Penna, M.A., Wu, J., 1993. Models for map building and navigation. *IEEE Transactions on Systems, Man and Cybernetics* 23, 1276–1301.
- Reid, A.K., Staddon, J.E.R., 1998. A dynamic route finder for the cognitive map. *Psychological Review* 105, 585–601.
- Rowe, N.C., Alexander, R.S., 2000. Finding optimal-path maps for path planning across weighted regions. *International Journal of Robotics Research* 19, 83–95.
- Santschi, F., 1913. Comment s'orientent les Fourmis. *Revue Suisse de Zoologie* 21 (12), 347–425.
- Schmajuk, N.A., 1990. Role of the hippocampus in temporal and spatial navigation: an adaptive neural network. *Behavioral Brain Research* 39, 205–229.
- Schmajuk, N.A., 1998. Introduction to the special issue on biologically inspired models of navigation. *Adaptive Behavior* 6, 369–370.
- Schmajuk, N.A., Thieme, A.D., 1992. Purposive behavior and cognitive mapping: an adaptive neural network. *Biological Cybernetics* 67, 165–174.
- Schmajuk, N.A., Blair, H.T., 1993. Place learning and the dynamics of spatial navigation: an adaptive neural network. *Adaptive Behavior* 1, 355–387.
- Schmajuk, N.A., Zanutto, B.S., 1998. Escape, avoidance, and imitation: a neural network approach. *Adaptive Behavior* 6, 63–129.
- Schmidt, I., Collett, T.S., Dillier, F.-X., Wehner, R., 1992. How desert ants cope with enforced detours on their way home. *Journal of Comparative Physiology, A* 171, 285–288.
- Schultz, A.C., Adams, W., Yamauchi, B., 1999. Integrating exploration localization and planning with a common representation. *Autonomous Robots* 6, 293–308.
- Singh, S.P., Sutton, R.S., 1996. Reinforcement learning with replacing eligibility traces. *Machine Learning* 22, 123–158.
- Staddon, J.E.R., 2001. *Adaptive Dynamics: The Theoretical Analysis of Behavior*. MIT Press.
- Sutton, R.S., 1988. Learning to predict by the methods of temporal differences. *Machine Learning* 3, 9–44.
- Sutton, R.S., 1990. Integrated architectures for learning, planning, and reacting based on approximating dynamic programming. In: Porter, B.W., Mooney, R.J. (Eds.),

- Machine Learning, Proceedings of the Seventh International Conference on Machine Learning. Morgan Kaufmann Publishers, pp. 216–224.
- Sutton, R.S., Barto, A.G., 1998. Reinforcement Learning: An Introduction. MIT Press, Cambridge, MA.
- Thorpe, W.H., 1950. A note on detour experiments with *Ammophila pubescens* Curt. (Hymenoptera; Sphecidae). Behaviour 2, 163–257.
- Tolman, E.C., Honzik, C.H., 1930. Insight in rats. University of California Publications in Psychology 4, 215–232.
- Tolman, E.C., Ritchie, B.F., Kalish, D., 1946. Studies in spatial learning. I. Orientation and the shortcut. Journal of Experimental Psychology 36, 13–24.
- Trullier, O., Wiener, S., Berthoz, A., Meyer, J.-A., 1997. Biologically-based artificial navigation systems: review and prospects. Progress in Neurobiology 51, 483–544.
- Trullier, O., Meyer, J.A., 1998. Animat Navigation Using a Cognitive Graph. In: R. Pfeifer, B. Blumberg, J.-A. Meyer and S.W. Wilson (Eds.). From Animals to Animats 5. Proceedings of the Fifth International Conference on Simulation of Adaptive Behavior.
- Voicu, H., Schmajuk, N.A., 2001a. Three-dimensional cognitive mapping with a neural network. Robotics and Autonomous Systems 35, 23–36.
- Voicu, H., Schmajuk, N.A., 2001b. Exploration, navigation, and cognitive mapping. Adaptive Behavior, 8, 207–224.
- Watkins, C.J.C.H., 1989. Learning from Delayed Rewards, King's College, Cambridge University (Ph.D. Thesis).
- Watkins, C.J.C.H., Dayan, P., 1992. Technical Note: Q-learning. Machine Learning 8, 279–292.
- Wehner, R., Menzel, R., 1990. Do insects have cognitive maps. Annual Review of Neuroscience 13, 403–414.
- Witkowski, C.M., 1997. Schemes for Learning and Behaviour: A New Expectancy Model. Ph.D. Thesis, Department of Computer Science, Queen Mary Westfield College, University of London.
- Zeil, J., Layne, J., 2002. Path integration in Fiddler Crabs and its relation to habitat and social life. In: Wiese, K. (Ed.), The Crustacean Nervous System. Springer.
- Zimbardo, P.G., Miller, N.E., 1958. Facilitation of exploration by hunger in rats. Journal of Comparative and Physiological Psychology 51, 43–46.
- Zipser, D., 1985. A computational model of hippocampal place fields. Behavioral Neuroscience 99, 1006–1018.