

Purposive behavior and cognitive mapping: a neural network model

N. A. Schmajuk and A. D. Thieme

Department of Psychology, Northwestern University, Evanston, IL 60201, USA

Received July 7, 1991/Accepted in revised form January 18, 1992

Abstract. This study presents a real-time, biologically plausible neural network approach to purposive behavior and cognitive mapping. The system is composed of (a) an action system, consisting of a goal-seeking neural mechanism controlled by a motivational system; and (b) a cognitive system, involving a neural cognitive map. The goal-seeking mechanism displays exploratory behavior until either (a) the goal is found or (b) an adequate prediction of the goal is generated. The cognitive map built by the network is a *topological map*, i.e., it represents only the adjacency, but not distances or directions, between places. The network has recurrent and non-recurrent properties that allow the reading of the cognitive map without modifying it. Two types of predictions are introduced: fast-time and real-time predictions. Fast-time predictions are produced in advance of what occurs in real time, when the information stored in the cognitive map is used to predict the remote future. Real-time predictions are generated simultaneously with the occurrence of environmental events, when the information stored in the cognitive map is being updated. Computer simulations show that the network successfully describes latent learning and detour behavior in rats. In addition, simulations demonstrate that the network can be applied to problem-solving paradigms such as the Tower of Hanoi puzzle.

1 Introduction

Cognitive views of animal and human behavior suggest that behavior can be described in terms of how different goals are pursued (Tolman 1932). According to Tolman, during instrumental learning animals learn *expectancies* that the performance of response $R1$ in a situation $S1$ will be followed by a change to situation

$S2$ ($S1$ - $R1$ - $S2$ expectancies). In contrast with reinforcement theories popular at the time, Tolman suggested that no reward is needed for animals to learn this sequence of events, i.e., associations are learned by simple temporal contiguity. During classical conditioning, animals learn to expect (predict) that a conditioned stimulus (CS) is followed by another CS or by the unconditioned stimulus (US).

Tolman (1932) proposed that multiple expectancies can be integrated into larger units, through a reasoning process called *inference*. One simple example of inference formation is *sensory preconditioning* (Bower and Hilgard 1981). Sensory preconditioning consists of a first phase in which two conditioned stimuli, $CS(A)$ and $CS(B)$, are paired together in the absence of the unconditioned stimulus, US. In a second phase, $CS(A)$ is paired with the US. Finally, when $CS(B)$ is presented alone it generates a weak but distinctive conditioned response (CR): the animal has inferred that $CS(B)$ predicts the US. Tolman hypothesized that a large number of expectancies can be combined in a *cognitive map*. Cognitive maps allow organisms to predict when and where temporally or spatially remote environmental events might be expected.

Tolman (1932) suggested that latent learning and detour learning illustrate the animals' capacity for reasoning. In latent learning, animals are exposed to a maze without being rewarded at the goal box. When a reward is later presented, animals demonstrate knowledge of the spatial arrangement of the maze which remains "latent" until reward is introduced. Detour problems are maze problems that can be solved by integrating separately learned pieces of information into a comprehensive depiction of the environment.

When seeking a goal in a maze, organisms compare the expectancies evoked by different alternative paths. For Tolman, *vicarious trial-and-error behavior* (VTE), i.e., the active scanning of alternative pathways at choice points, reflects the animal's generation and comparison of different expectancies. At choice points, animals sample different stimuli before making a decision. For example, a rat often looks back and forth between

alternative stimuli before approaching one or the other. According to Tolman's "stimulus-approach" view, organisms learn that a particular stimulus situation is appetitive, and therefore it is approached. Supporting this assumption, Mackintosh (1974) suggested that, in the presence of numerous intra- and extra-maze cues, animals typically learn to approach a set of stimuli associated with reward and to avoid a set of stimuli associated with punishment. Only in a totally uniform environment will animals learn to make the correct responses that lead to the goal as suggested by Hull (1943).

2 The neural network

2.1 A biological machine capable of purposive behavior

For Gallistel (1980), Tolman's view of purposive course of action in the environment can be derived from a "biological machine" with two distinct subsystems: an *action system* that controls navigation through the external world, and a *cognitive system* that forms representations of the environment. In agreement with Gallistel, Fig. 1 shows a block diagram of a system composed of (a) an action system, consisting of a goal-seeking neural mechanism controlled by a motivational system, and (b) a cognitive system, involving a neural cognitive map.

Because Tolman never specified how action was related to cognition, Guthrie (1935) criticized Tolman's approach by suggesting that Tolman's rat was left

"buried in thought". In order to establish the connection between cognition and behavior, the action system shown in Fig. 1 implements a "stimulus-approach" view of behavior control by which the system approaches appetitive stimuli and avoids aversive stimuli. Appetitive and aversive stimuli are defined by the system's motivational state. A given motivation elicits a search behavior that receives negative feedback from the appropriate goal. When the goal is perceived, it activates approach responses and inhibits search responses. Motivation decreases when reward is procured at the goal location.

As the biological machine explores the environment under the control of the action system, the cognitive system builds a cognitive map, i.e., an internal representation of the world. Once the cognitive map is available, environmental stimuli that predict the goal activate approach responses and inhibit search responses. The cognitive map is updated whenever a comparator detects a mismatch between predicted and actual inputs. Whereas changes in the cognitive map are regulated by real-time predictions, the cognitive map interacts with the action system by generating fast-time predictions of the goal (Sect. 2.4). The following sections explain how the cognitive map is built and updated, and how maze navigation is accomplished with the assistance of the cognitive map.

2.2 Cognitive mapping and maze learning

To the extent that the layout of a maze can be accurately portrayed in terms of the connections between its contiguous places, it seems reasonable to describe maze connectivity in terms of the associations between the representations of adjacent places. However, because associations between spatially adjacent places are insufficient to predict remote places, navigation to remote goals requires the building and reading of a cognitive map. The cognitive map allows the combination of information about spatially adjacent places, and thereby the inference of the connections to remote places.

2.3 Building the cognitive map

Figure 2 shows a network capable of cognitive mapping. The neural network is a real-time mechanism that describes behavior as a moment-to-moment phenomenon. Appendix A presents a formal description of the network as a set of differential equations that govern changes in the values of neural activities and connectivities as a function of time. This formal description of the model employs biologically plausible principles.

The network shown in Fig. 2 includes two types of inputs: Places and Views. Places refer to discernable regions in the maze. Views refer to the sights of places as seen from another place. For simplicity, the current implementation assumes that the system only perceives views of places *adjacent* to the place where it is located. Although in the present version of the model we assume independence between places and views, a given place can be defined by the set of views available from that place. Therefore, views may be regarded as egocentric

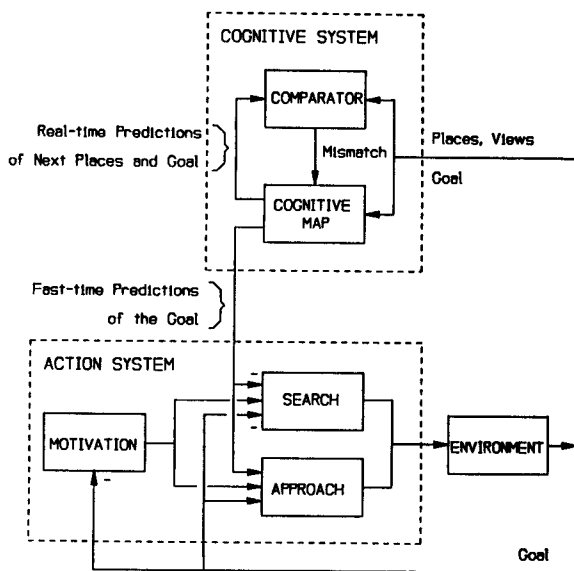


Fig. 1. A biological machine capable of maze navigation. Block diagram of the model showing the interaction between the action system, cognitive system, and environment. *Cognitive system*: A comparator contrasts actual and real-time predictions of views and the goal and the cognitive map is modified when a mismatch is detected. Values of fast-time predictions of the goal are provided to the action system. *Action system*: When present, either the goal or the environmental stimulus best predicting the goal are approached. In the absence of the goal or predictions of the goal, environmental stimuli are randomly approached

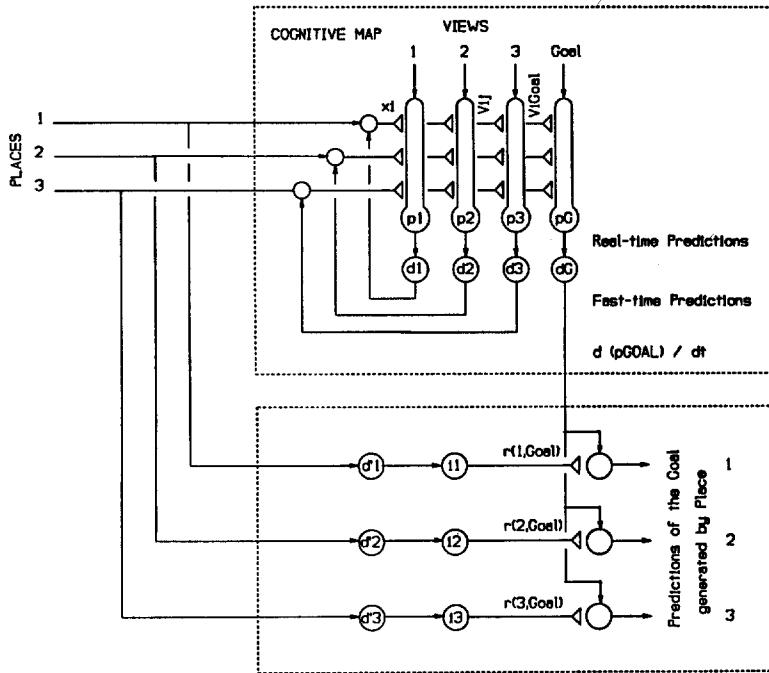


Fig. 2. Cognitive system. The cognitive map is implemented by a neural network that associates Place representations with View representations. The first derivatives of the predictions of View j , p_j , are fed back into the representation of Place i , $x_i = j$. The first derivative of the prediction goal, p_{goal} , is stored in working memories $r_{h,\text{goal}}$. $V_{i,j}$: Long-term memory of place-view associations. d_j and d'_j : Neurons computing first derivatives. i_j : Neurons integrating the pulses generated by neurons d'_j . G : Goal. Arrows: Fixed excitatory connections. Open triangles: Variable excitatory connections

representations (depictions of spatial locations as perceived from the animal's vantage point), and places may be considered allocentric representations (portrayals of spatial locations relative to the positions of other discriminable locations) of the environment.

We assume that both place and view inputs are constant as long as the animal stays in the same Place. Place i in a maze is assumed to give rise to a short-term memory (STM) trace, x_i . Place trace x_i may become associated with the views of other places to form long-term associations $V_{i,j}$ between Place i and View j . When the animal is at Place i and perceives View j , $V_{i,j}$ increases. When the animal is at Place i and cannot perceive View j , $V_{i,j}$ decreases. Each time the animal enters Place i , $V_{i,j}$ associations generate *real-time* predictions of Views j to be seen from Place i . That is, as long as the animal stays in Place i , Place i activates neurons p_j proportionally to $V_{i,j}$, and this activity represents the prediction that View j is available from Place i (3 in Appendix A). If the comparator detects a mismatch between the actual and the predicted view, $V_{i,j}$ is readjusted to reflect the maze configuration (4 in Appendix A).

The cognitive map built by the network is a *topological map*, i.e., it represents only the adjacency, but not distance or direction, of places. $V_{i,j}$ associations are the elementary internal learned representations of the topological connections in the external world. $V_{i,j}$ associations are stored as weights of the modifiable synapses represented by open triangles in Fig. 2.

2.4 Reading the cognitive map

Cognitive maps combine multiple associations $V_{i,j}$ to infer spatially remote goal locations. This can be achieved by recurrently reinjecting the signal represent-

ing View j as predicted by Place i according to $V_{i,j}$, into the representation of Place j . Place j now predicts View k according to $V_{j,k}$, and the signal representing View k is reinjected into the representation of Place k . The process continues until the representation of the goal is eventually activated.

Two major problems may hamper the recurrent signal reinjection process: (a) the extinction of previously stored $V_{i,j}$ associations, and (b) the formation of spurious $V_{i,j}$ associations. This section explains how the network shown in Fig. 2 avoids both problems and achieves cognitive mapping.

Suppose that a maze consists of Place A connected to Place B and Place B connected to Place C . Therefore, a correct internal representation of the maze would consist of $V_{A,B} = 1$, $V_{B,C} = 1$, $V_{B,A} = 1$, $V_{C,B} = 1$, $V_{A,C} = 0$, and $V_{C,A} = 0$. The potential extinction of correct, previously stored Place-View associations can be explained as follows. Assume that the animal is located in Place A . Then, Place A steadily activates a representation of Place B through the $V_{A,B}$ association, and the representation of Place B steadily activates the representation of Place C through the $V_{B,C}$ association. Because the View of Place C is not available at Place A , and hence is not activated, association $V_{B,C}$ might extinguish according to (4) in Appendix A. This undesired result might be described as an extinction produced by the mismatch between the prediction that Place C will be encountered in the future and the present absence of Place C .

The potential acquisition of spurious Place-View associations is described as follows. Assume now that the animal is located in Place B . Place B steadily activates through association $V_{B,A}$ the representation of Place A . Because the View of Place C (visible from

Place B) is steadily active together with the representation of Place A , association $V_{A,C}$ might increase by (4) in Appendix A. This undesired result might be described as the acquisition of a spurious association between (the activated representation of) Place A with the current view of Place C .

The network shown in Fig. 2 allows the combination of multiple $V_{i,j}$ associations while preserving a correct representation of the maze stored in Place-View associations. In Fig. 2, the output of cells p_j excite cells d_j , which generate an output proportional to the first derivative of p_j with respect to time, $d(p_j)/dt$. The output of cells d_j excite the corresponding Place traces, x_j . Since slow-changing signals weakly activate cells d_j , only fast-changing signals are reinjected into x_j . In other words, cells d_j allow fast-changing signals, but not slow-changing signals, to spread their activation over the network. Therefore, the network operates as a heteroassociative *nonrecurrent* network for slow-changing signals and as a *recurrent* network for fast-changing signals¹. These dual recurrent and nonrecurrent properties of the network are fundamental to cognitive mapping.

The extinction of previously stored Place-View associations is now avoided for the following reasons. Suppose that the animal is located in Place A . Then, Place A activates the $V_{A,B}$ synapse but the representation of Place B , x_B , is not steadily activated because p_B does not activate x_B . Therefore, although View C is not available from Place A , association $V_{B,C}$ does not decrease. The acquisition of spurious Place-View associations is also averted. Assume that the animal is located in Place B . Because the representation of Place A is not steadily active along with the view of Place C , the association $V_{A,C}$ does not increase. In sum, undesirable changes in Place-View associations are avoided because the network does not reinject slow-changing signals. In other words, the inclusion of d_j cells allows the network to function as a *nonrecurrent* heteroassociative network for slow-changing signals, thereby preserving the correct internal representation of the maze configuration.

As Tolman, we assume that before making a decision, the animal briefly examines all the *alternative next places* h linked to Place i . These brief inspections (reflected in VTE behavior) result in fast, short, and relatively weak activation pulses of the place inputs and their corresponding traces, x_h . Place traces x_h activate cells p_j in proportion to their $V_{h,j}$ connections. Therefore, view cells, p_j , are activated by fast-changing signals proportional to $x_h V_{h,j}$, that are recurrently reinjected into x_j through cells d_j . Subsequently, x_j activates p_k , p_k activates x_k , and so forth, spreading the activation over the network. Because fast-changing sig-

nals cannot fully activate cells x_j and x_k (see (1) in Appendix A), activation spreads decrementally, and consequently the outputs of cells p_m and d_m are inversely proportional to the distance (measured in number of intermediate places) of the alternative next Place h to Place m .

The output of cells p_j and d_j displays a sequence of pulses that reflects the organization of the maze starting at the alternative place under examination (Fig. 8). This sequence of pulses might include the activation of the representation of the goal. When present, this activation is inversely proportional to the number of reinjections needed to elicit it. In sum, cognitive mapping is achieved by combining multiple associations $V_{i,j}$ as the network recurrently reinjects fast-changing view representations into their corresponding place representations. In other words, the inclusion of d_j cells allows the network to function as a *recurrent* heteroassociative network for fast-changing signals, thereby allowing the inference of remote places.

Notice that in contrast with the *real-time predictions* of View j generated by the slow-changing Place i signal, internal navigation generates *fast-time predictions* of Views: imaginary navigation proceeds at a faster pace than real movement through the maze.

2.5 Maze navigation

As explained previously, before deciding which place it will enter, the action system sequentially evaluates all alternative next Places h linked to its present Place i . When the action system examines all the alternative next places, the output of cell d_{goal} is stored into a working memory $r_{h,\text{goal}}$ (Fig. 2). Working memories $r_{h,\text{goal}}$ are proportional to the prediction of the goal by alternative next Place h (6 in Appendix A). By simultaneously comparing all $r_{h,\text{goal}}$ values provided by the cognitive system, the action system decides which of the sequentially scanned places is the best predictor of the goal. Working memories $r_{h,\text{goal}}$ decay back to zero by the time the animal arrives at the next choice point.

Under the "stimulus-approach" view, the alternative next place that best predicts the goal becomes a *subgoal* and is subsequently approached by the action system (see Schmajuk 1990). The goal always has preeminence over its predictions. If the best prediction of the goal is smaller than a certain minimum value, the action system engages in a random exploratory behavior. If the magnitudes of two predictions are identical, the action system decides at random between the two stimuli generating the predictions. We assume that the total time and, consequently, the total number of reinjections allowed to the inspection of the connectivity of each alternative place by the action system is fixed. Maze navigation proceeds until the goal is found or the animal is withdrawn from the maze.

Summarizing, the recurrent and nonrecurrent properties of the cognitive map shown in Fig. 2 allow the exploration of the connectivity of the maze without the alteration of associations between adjacent Places and Views. In other words, the network is capable of navi-

¹ Kohonen (1977) classified associative networks into autoassociative and heteroassociative. Autoassociative networks are those in which the input pattern is associated with itself. Heteroassociative networks are those in which two different input patterns are associated. Autoassociative recurrent networks are autoassociative networks in which the outputs either (a) constitute additional inputs or (b) are reinjected to the inputs of the network.

gating the internal representation of the maze without modifying its contents. The neural architecture of Fig. 2 may be regarded as consisting of two functionally overlapping subsystems. One functional subsystem maintains a topological map of the environment and generates real-time predictions for adjacent places. The other subsystem reads out the cognitive map by generating fast-time predictions of remote places. Fast-time predictions are used to guide the animal's behavior but, given their short duration, they do not modify Place-View associations. Real-time predictions are generated simultaneously with the occurrence of environmental events, and are used to update Place-View associations that reflect the connectivity of the external world. At a given choice point, animals briefly examine all alternative next places connected to their present place, thereby generating fast-time predictions of the goal. These sequentially generated fast-time predictions are stored in working memories that permit the action system to simultaneously compare all the alternative next places. As a result of this comparison, a decision is made and the action system directs the animal to enter the place that best leads to the goal.

3 Computer simulations

This section presents computer simulations obtained with the network for two maze learning paradigms, latent learning and detour learning, that according to Tolman (1932) illustrate animals' capacity for reasoning. Simulations consist of trials divided into "moves". During each move, the cognitive system updates the representation of the maze configuration and predicts the consequences of each of all possible movements, after which the action system selects the next place to be occupied by the simulated animal. A complete description of the simulation protocols is presented in Appendix B. All simulations were carried out with identical parameter values. Parameter values used in the simulations are presented and justified in Appendix C.

3.1 Latent learning

3.1.1 Experimental data. Blodgett (1929) studied how nonrewarded trials affect performance when reward is later introduced in a multiple *T*-maze. A diagram of Blodgett's maze is shown in Fig. 3. Food-deprived rats received one trial a day, in which they were placed in the start box and retrieved after reaching the goal box. Several groups of rats were trained. One group received food reward in the goal box starting on the first trial. Two other groups were rewarded after 3 or 7 nonrewarded trials. Blodgett found that after only one rewarded trial the performance of the initially nonrewarded groups improve to nearly the same level of the group rewarded on all trials.

3.1.2 Computer simulations. During nonrewarded trials, the model constructs a cognitive map of Place-View associations. When no Place-Goal associations are

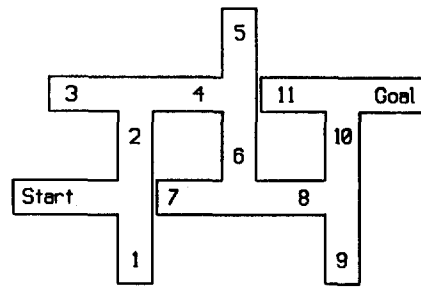


Fig. 3. Latent learning. Diagram of a multiple *T*-maze employed to study latent learning

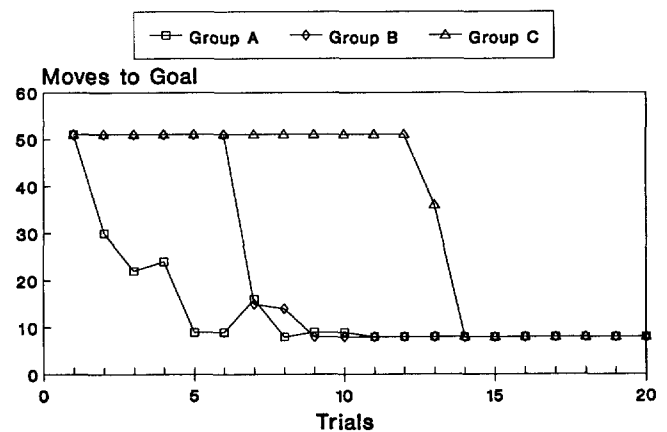


Fig. 4. Latent learning. Number of simulated moves to reach the goal in a multiple-*T* maze for three different groups. Group *A* is rewarded at the goal box on the first trial, Group *B* on the seventh trial, and Group *C* on the thirteenth trial

available, the map is not utilized. When Place-Goal associations are established, the animal is able to determine the correct path, therefore showing faster improvement in performance than unexposed animals.

Figure 4 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 seventh trial, and Group *C* on the thirteenth trial. According to Fig. 4, the model shows latent learning because when reward is presented after a period of latency, animals with a preconstructed 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's (1929) latent learning data.

3.2 Detour problem

3.2.1 Experimental data. In order to study detour performance, Maier (1929) used a complex experimental room, which included tables, ringstands, and elevated paths, to show that rats swiftly take an alternative known path if the usual path to the goal is blocked with a wire mesh.

Tolman and Honzik (1930) used a three-path maze with three pathways (*A*, *B* and *C*) to analyze how rats

solve a detour problem (Fig. 5). Path *A* is the shortest, Path *B* is the next shortest, and Path *C* is the longest. Path *A* is formed by Places, 1, 3, 4, and 8. Path *B* is formed by Places 1, 2, 3, 4, and 8. Path *C* is formed by Places 1, 5, 6, 7, 4, and 8. Paths *A* and *B* share a common place, Place 3, which Path *C* does not have. 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 is placed in the common section of Paths *A* and *B* (Block *b* between Places 3 and 4, in Fig. 5), thereby obstructing both paths. In Tolman and Honzik's experiment, most rats, after backing out of Path *A*, take the longest path, Path *C*, instead of the second shortest path, Path *B*.

3.2.2 Computer simulations. Figure 6 shows (a) the number of moves to the goal and (b) Place-View associations $V(1,2)$, $V(1,3)$, $V(1,5)$, and $V(3,4)$ as a function of trials, in which different maze configurations are adopted. In the simulation, the animal is successively forced to take paths *C*, *B*, and *A* for 24, 20, and 20 trials respectively. When perfectly learned, Path *A* is traversed in 5 moves, Path *B* in 6 moves, and Path *C* in 7 moves. After the forced phase, the simulated animal is free to choose among all three paths, and chooses the shortest path, *A*. On trial 90, Block *a* is introduced in Path *A* between Places 1 and 3, and the animal immediately detours to the second shortest path, *B*.

Figure 6 shows that, as the animal navigates through Path *C*, association $V(1,5)$ increases. When the animal navigates through Path *B*, associations $V(1,2)$ and $V(3,4)$ increase and association $V(1,5)$ decays to zero. When the animal navigates through Path *A*, associations $V(1,3)$ and $V(3,4)$ increase and association $V(1,2)$ extinguishes. Finally, when all pathways are open, associations $V(1,2)$, $V(1,3)$, $V(1,5)$ and $V(3,4)$ increase. Figure 6 shows that the introduction of Block *a* causes association $V(1,3)$ to decay to zero without altering the other associations, allowing the animal to detour to the second shortest path, *B*.

Figure 7 shows (a) the number of moves to the goal and (b) Place-View associations $V(1,2)$, $V(1,3)$, $V(1,5)$, and $V(3,4)$ as a function of trials, in which different maze configurations are adopted. As in the previous simulation, the animal is successively forced to take paths *C*, *B*, and *A* for 24, 20, and 20 trials respectively. After this forced phase, the simulated animal is free to choose among all three paths, and chooses the shortest

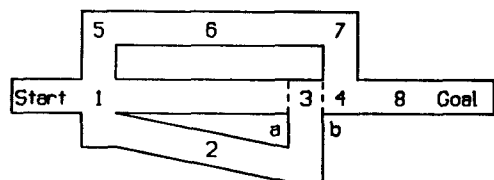


Fig. 5. Detour problems. Diagram of the maze employed to study detour problems. Path *A* is formed by places 1, 3, 4, and 8. Path *B* is formed by places, 1, 2, 3, 4, and 8. Path *C* is formed by places 1, 5, 6, 7, 4, and 8. Dashed lines *a* and *b* indicate blocks

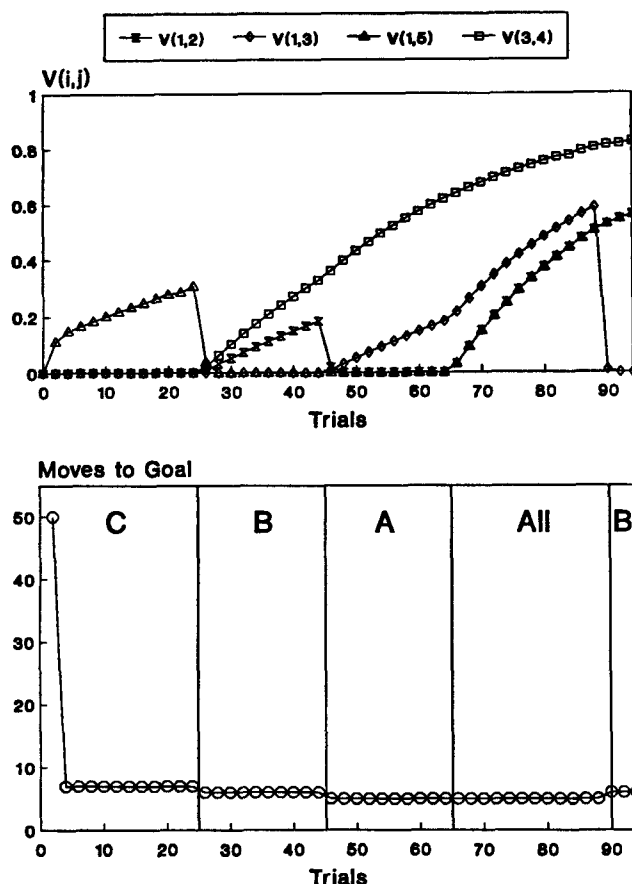


Fig. 6. Short detour problem. Number of simulated moves to reach the goal and Place-View associations as a function of trials, for different maze configurations. Trials 1–24: The animal is forced through the longest path, Path *C*. Trials 25–44: The animal is forced through Path *B*. Trials 45–64: The animal is forced through Path *A*. Trials 65–89: The animal is allowed to enter all pathways and chooses Path *A*. Trials 90–95: Block *a* is introduced on Trial 90 and the animal detours to Path *B*.

path, Path *A*. Before the introduction of the block on Trial 90, variations in Place-View associations are identical to those presented in Fig. 6. On trial 90, Block *b* is introduced between Places 3 and 4, blocking Paths *A* and *B*. Figure 7 shows that the introduction of Block *b* causes association $V(3,4)$ to decay to zero, and, although association $V(1,2)$ remains unchanged, the animal immediately detours to the longest path, *C*. This simulation result is 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.

Internal navigation of the maze occurs in fast-time, i.e., it proceeds at a faster pace than the real movement through the maze. Figure 8 illustrates how fast-time predictions are generated by the network. The upper panel in Fig. 8 shows a real-time simulation of the activity of neurons representing different maze views when the animal briefly glimpses from Place 1 into Places 2, 3, and 5, before Block *b* is introduced. The upper panel in Fig. 8 shows that Place 3 evokes the

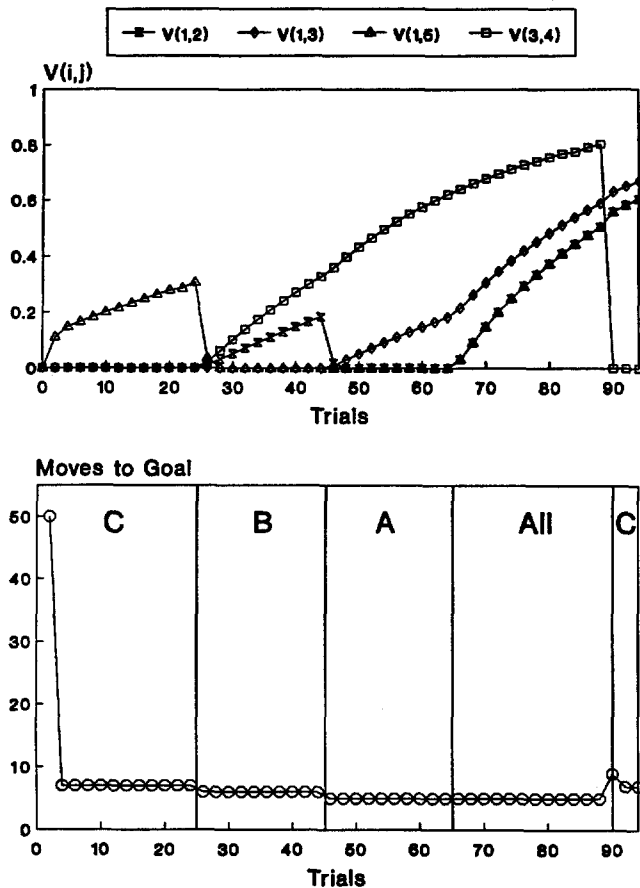


Fig. 7. Long detour problem. Number of simulated moves to reach the goal and Place-View associations as a function of trials, for different maze configurations. Trials 1–24: The animal is forced through the longest path, Path C. Trials 25–44: The animal is forced through Path B. Trials 45–64: The animal is forced through Path A. Trials 65–89: The animal is allowed to enter all pathways and chooses Path A. Trials 90–95: Block *b* is introduced on Trial 90 and the animal detours to Path C

strongest representation of the goal, and therefore, will be entered by the animal. The lower panel in Fig. 8 shows a real-time simulation of the activity of the nodes representing different maze views when the animal briefly glimpses from Place 1 into Places 2, 3, and 5, after finding Block *b*. The lower panel in Fig. 8 shows that after Block *b* is introduced, Place 5 evokes the strongest representation of the goal, and therefore, will be entered by the animal. Notice that the network generates the sequence in which different places in the maze are interconnected independently of the sequence in which these places were originally visited.

4 Discussion

This study presents a real-time neural network that describes purposive behavior and cognitive mapping. The system is composed of (a) an action system, consisting of a goal-seeking neural mechanism controlled

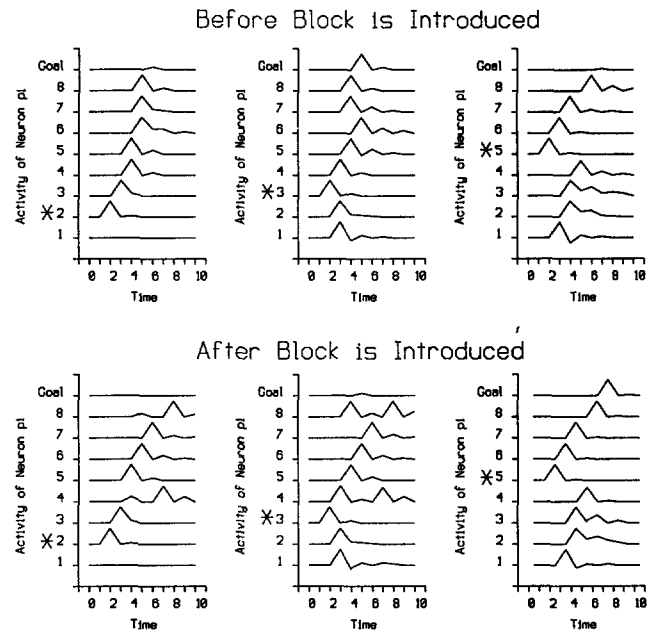


Fig. 8. Long detour problem. Real-time simulations of the predictions of different places in the maze, including the goal, generated by glimpses of Alternative Next Places 2, 3 and 5, from Place 1 before and after introducing Block *b* on Trial 90. Left Panels: Alternative Next Place 2. Central Panels: Alternative Next Place 3. Right Panels: Alternative Next Place 5. Upper Panels: Activity before introducing Block *b*. Lower Panels: Activity after introducing Block *b*. Notice that alternative 3 generates the strongest prediction of the goal (Place 9) before Block *b* is introduced, and alternative 5 generates the strongest prediction of the goal (Place 9) after Block *b* is introduced. In order to maximize display clarity, scales for each activity trace are different in every case except for the goal. Goal activity is scaled to its maximum value for alternative 3 in upper panels and for alternative 5 in the lower panels

by a motivational system; and (b) a cognitive system, involving a neural cognitive map. This “biological machine” is capable of guiding the animal’s search for a specific goal in the environment with the assistance of the cognitive map.

In accordance with Blodgett’s (1929) data, the model shows latent learning because when reward is presented after a period of latency, rats with a preconstructed cognitive map display rapid improvement in performance to the same level of the group rewarded from the first trial. Also in agreement with Tolman Honzik’s (1930) results, the network correctly describes detour behavior in rats by integrating separately learned pieces of information into a comprehensive depiction of the maze.

4.1 Other formal approaches to cognitive mapping

Alternative formal approaches to cognitive mapping have been proposed by Deutsch (1960); Lieblisch and Arbib (1982); and Hampson (1990). Although these approaches share aspects of the model introduced in the present study, our network differs from these models in several respects. The main difference between our network and Deutsch’s (1960) model is that whereas our

network generates forward predictions of the goal, Deutsch's model generates backward activation of the nodes. Unlike Lieblach and Arbib's (1982) "world graph," in which nodes represent connections between places and the drive-reduction properties of a given place, the cognitive map implemented by our network simply represents connection between places. Finally, in contrast to Hampson's (1990) use of stimulus-response operators, our network adopts a purely stimulus-response approach strategy.

4.2 Problem solving

Tolman (1932) suggested that the relations between initial and goal positions in a maze can be represented by a directed graph², and he called this graph a "means-end field". Many years later, artificial intelligence theories described problem solving as the process of finding a path from an initial to a desired state through a directed graph (Winston 1977). Directed graphs can be transformed into trees by terminating paths leading to previously visited nodes. Two basic methods of tree searching have been proposed: depth-first and breadth-first searches. In the depth-first method, one alternative is selected and pursued at each decision point until forward movements are blocked. Then forward exploration begins again from the nearest decision point with unexplored alternatives. In a breadth-first method, all alternatives at each decision point are equally explored. According to Winston (1977), when search trees are very large both depth- and breadth-first techniques become inefficient. The inefficiency of a sequential approach to tree searching is avoided by our neural network. Our neural network combined depth- and breadth-first procedures: at each decision point, the network first explores *in depth* all alternatives and then selects the best one.

In order to ascertain the power of the network in tree-structured problem solving other than maze learning, we applied the network to the Tower of Hanoi task. We represented the task in terms of the possible transitions between different ring arrangements on the three poles (Hampson 1990). Simulations show that in a few trials, the network learns to solve the problem in the minimal number (seven) of movements.

5. Conclusion

The present paper introduces a real-time, biologically plausible neural network capable of describing cognitive mapping. Computer simulations show that the network successfully describes latent learning and detour behavior in rats. In addition, simulations demonstrate that the network can be applied to problem-solving paradigms, such as the Tower of Hanoi puzzle.

² A directed graph is a set of nodes (representing states) and a set of links (representing possible transitions between states).

Appendix A: formal description of the model

Changes in the representation of Place i , x_i , are given by

$$dx_i/dt = -k_1 x_i + k_2(1 - x_i)T_i, \quad (1)$$

where k_1 is a decay constant and k_2 is a rise constant. x_i is bound between 0 and 1.

T_i is given by

$$T_i = \text{Place } i + k_3 f[d(p_i)/dt], \quad (2)$$

where $\text{Place } i = 1$ if the animal is at Place i , 0.1 if the animal glances into Place i while at another position, and 0 otherwise. k_3 is the reinjection constant. $f[d(p_i)/dt]$ is the *fast-time prediction* of Place i . $f[d(p_i)/dt] = d(p_i)/dt$ if $d(p_i)/dt > 0$, and $f[d(p_i)/dt] = 0$ otherwise. Therefore, the representation of Place i , x_i , is active when the animal is at Place i or when Place i is predicted by the network

The *real-time prediction* of View j , p_j is given by

$$p_j = \sum_i V_{i,j} x_i, \quad (3)$$

where $V_{i,j}$ is the association between Place i and View j .

Changes in association $V_{i,j}$ between Place i and View j are given by

$$dV_{i,j}/dt = k_4 x_i (\text{View } j - p_j). \quad (4)$$

According to (4) $V_{i,j}$ increases when View j and x_i are active together (with $k_4 = k_4'$), and decreases when x_i is active in the absence of View j (with $k_4 = k_4''$). View j is active whenever Place j is visible from Place i contiguous to Place j . x_i is active either when (a) the animal is in Place i , or (b) the animal glimpses into Place k associated with the view of Place i through $V_{k,i}$. When the animal is in Place i , x_i rises rapidly to an asymptotic value of one for 90 or more time units (see Appendix B). In contrast, when the animal glimpses into Place k , x_k is a short and weak pulse (1) and $d(p_i)/dt$ is a short and weak pulse activated by x_k proportionally to $V_{k,i}$ (2). Therefore, x_i is essentially proportional to Place i ((1) and (2)). Consequently, $V_{i,j}$ is barely modified when the animal glimpses into different places associated with Place i and essentially reflects the association between Place i and View j .

Before making a decision, the animal examines all the alternative next places h linked to Place i for one time unit. As mentioned, during these brief inspections Place h assumes the value of 0.1. The output of neurons d'_h is proportional to $f[d(\text{Place}_h)/dt]$, where $f[x] = 0.1$ if $x > 0$ and $f[x] = 0$ otherwise. Therefore, neurons d'_h generate a one time unit pulse whenever a place is entered by the animal or when it is briefly examined as a candidate for the next movement.

Neurons i compute a STM of the places h examined by the animal, according to

$$di_h/dt = -k_5 i_h + k_6(1 - i_h)f[d(\text{Place}_h)/dt]. \quad (5)$$

By (5) neuron i_h acts as a leaky integrator of the pulse generated by neuron d'_h . The output of neuron i_h is proportional to $f[x]$, where $f[x] = 1$ if $x > k_7$ (a

threshold) and $f[x] = 0$ otherwise. With the parameters defined in Appendix C, neurons i_h generate a pulse approximately 30 time units long whenever Place_h is briefly examined as a candidate for the next movement.

The output of neurons i_h can be regarded as a STM of the alternative next place h examined by the animal. STM i_h is associated with the fast-time prediction of the goal, $d(p_{\text{goal}})/dt$, generated when the animal examines next Place h , in working memory $r_{h,\text{goal}}$ according to

$$d(r_{h,\text{goal}})/dt = -k_8 r_{h,\text{goal}} + k_9 i_h (1 - r_{h,\text{goal}}) d(p_{\text{goal}})/dt. \quad (6)$$

According to (6) $r_{h,\text{goal}}$ increases when the STM of next alternative Place h , i_h , and $d(p_{\text{goal}})/dt$ are active together and spontaneously decreases at a slow rate. $r_{h,\text{goal}}$ is bound between 0 and 1.

The action system chooses next place h , with the highest value of $r_{h,\text{goal}}$. If the largest $r_{h,\text{goal}}$ is less than a threshold value (k_{10}), the action system randomly chooses the next place. If two $r_{h,\text{goal}}$ have identical values, the action system randomly decides between those two places.

Equations (1) to (5) are biologically plausible. As shown by Grossberg (1975), (1) is based on the properties of neural membranes. Neurons computing first derivatives, as those described in (2) have been reported in different sensory systems (Adrian 1928). Finally, (4) is in agreement with Kelso et al. (1986) and Stanton and Sejnowski's (1989) data showing that the connectivity between two neurons, as reflected by long-term potentiation (LTP), increases when the presynaptic membrane is active in the presence of postsynaptic depolarization, and with Stanton and Sejnowski's (1989) data showing that LTP decreases when the presynaptic membrane is active in the absence of postsynaptic depolarization.

Appendix B: simulation protocol

Simulations consist of a number of trials, this number depending on the paradigm being replicated. Each trial is divided into "moves". During each move, the cognitive system updates the representation of the maze configuration and predicts the consequences of each of all possible movements. A move ends when the action system selects the next place to be occupied by the animal. Each move begins with a stabilization of 30 time units during which x_i , the STM trace for the place that the animal is occupying, Place_i , reaches asymptotic value. During VTE behavior, each Place h connected to Place i is briefly examined for one time unit. The total time allowed for the prediction of the consequences of moving to Place h , $r_{h,\text{goal}}$, is 30 time units. In order to avoid interference, 30 time units separate the examination of each alternative next place. When all next places have been examined, the place to be entered is selected by the action system and a new move begins. At the beginning of each trial the animal is placed into the start box. A maximum of 60 moves is allowed on

any given trial. A trial ends when the animal reaches the goal or when the maximum number of moves per trial is exceeded.

Appendix C: parameter values

Computer simulations generate values of the relevant variables only at discrete time units. In our simulations we assumed that one time step is equivalent to 1 ms. Parameters used in all simulations were $k_1 = 0.99$, $k_2 = 0.25$, $k_3 = 1$, $k_4' = 10^{-3}$, $k_4'' = 2$, $k_5 = 0.03$, $k_6 = 0.8$, $k_7 = 0.0375$, $k_8 = 10^{-3}$, $k_9 = 0.9$, and $k_{10} = 10^{-5}$. Parameters k_1 and k_2 were selected to obtain a moderate rise and a rapid decay in Place short-term memory, x_i . Parameter k_3 , which controls the rate of attenuation over repetitive reinjections in the network, was chosen to obtain an adequate signal at the neuron representing the goal. Greater values of k_3 would allow more reinjections without appreciable attenuation. Parameters k_4' and k_4'' were selected to achieve relatively slow acquisition and fast extinction of Place-View associations, $V_{i,j}$. A relatively slow acquisition rate k_4' is needed to describe latent learning and a fast extinction rate k_4'' is required to generate a fast detour. Parameters k_5 , k_6 and k_7 were selected to generate a STM of the alternative next place that lasts long enough (approximately 30 time units) to be associated with the predicted consequence of entering such places. k_8 and k_9 were selected (a) to store values of working memory $r_{h,\text{goal}}$ for a length of time that allows the comparison and selection of the best predicting path and (b) to ensure that working memory $r_{h,\text{goal}}$ decays to zero before the next "move" begins.

Acknowledgements. The authors are thankful to Tad Blair, John Gabrieli, and Roger Ratcliff for their helpful comments on a previous version of the manuscript. This project was supported in part by BRSG SO7 RR07028-22 awarded by the Biomedical Research Support Grant Program, Division of Research Resources, National Institutes of Health and by Contract N00014-91-J-1764 from the Office of Naval Research.

References

- Adrian ED (1928). The basis of sensation, the action of sense organs. Christophers, London
- Blodgett HC (1929) The effect of the introduction of reward upon maze performance of rats. University of California Publications in Psychology 4:113-134
- Bower GH, Hilgard ER (1981) Theories of learning. Prentice Hall, Englewood Cliffs, New Jersey.
- Deutch JA (1960) The structural basis of behavior. University of Chicago Press, Chicago
- Gallistel CR (1980) The organization of action: a new synthesis Erlbaum, Hillsdale, New Jersey
- Grossberg S (1975) A neural model of attention, reinforcement, and discrimination learning. Intern Rev Neurobiol 18:263-327
- Guthrie ER (1935) The psychology of learning. Harper and Row, New York
- Hampson SE (1990) Connectionistic problem solving. Birkhäuser, Boston
- Hull CL (1943) Principles of behavior. Appleton-Century-Crofts, New York

- Kelso SR, Ganong AH, Brown TH (1986) Hebbian synapses in hippocampus. *Proc Nat Acad Sci USA* 83:5326–5330
- Kohonen T (1977) *Associative memory*. Springer, Berlin Heidelberg New York
- Lieblich I, Arbib MA (1982) Multiple representations of space underlying behavior. *Behav Brain Sci* 5:627–659
- Mackintosh NJ (1974) *The psychology of animal learning*. Academic Press, London
- Maier NRF (1929) Reasoning in white rats. *Comp Psychol Monogr* 6:29
- Schmajuk NA (1990) Role of the hippocampus in temporal and spatial navigation: an adaptive neural network. *Behav Brain Res* 39:205–229
- Stanton PK, Sejnowski TJ (1989) Associative long-term depression in the hippocampus induced by Hebbian covariance. *Nature* 339:215–218
- Tolman EC (1932) *Purposive behavior of animals and men*. Irvington New York
- Tolman EC, Honzik CH (1930) “Insight” in rats. *University of California Publications in Psychology* 4:215–232
- Winston PH (1977) *Artificial intelligence* Addison-Wesley, Reading, Mass