Max–Planck–Institut
für biologische Kybernetik
Spemannstraße 38 • 72076 Tübingen • Germany

Technical Report No. 66

Spatial Cognition: Behavioral Competences, Neural Mechanisms and Evolutionary Scaling

Hanspeter A. Mallot[1]

November 1998

The work described in this paper was done at the Max-Planck-Institut für biologische Kybernetik. Additional support was obtained from the Deutsche Forschungsgemeinschaft (DFG, grant number Th 425/8-1) and the Human Frontier Science Program (HFSP, grant number RG–84/97). The introduction of this paper is based on the "Introductory remarks" for the workshop "Mechanisms of Cognition" held at the 26th Göttingen Neurobiology Conference (Mallot and Logothetis 1998). I am grateful to H. H. Bülthoff, M. O. Franz, S. Gillner, B. Schölkopf, S. D. Steck and K. Yashuhara who contributed to the work reviewed here.

[1] AG Bülthoff, E–mail: hanspeter.mallot@tuebingen.mpg.de

This document is available as `/pub/mpi-memos/TR-066.ps` via anonymous ftp from `ftp.kyb.tuebingen.mpg.de` or from the World Wide Web, `http://www.kyb.tuebingen.mpg.de/bu.html`.

Spatial Cognition: Behavioral Competences, Neural Mechanisms and Evolutionary Scaling

Hanspeter A. Mallot

Abstract. Spatial cognition is a cognitive ability that arose relatively early in animal evolution. It is therefore very well suited for studying the evolution from stereotyped to cognitive behavior and the general mechanisms underlying cognitive abilities. In this paper, I will present a definition of cognition in terms of the complexity of behavior it subserves. This approach allows to ask for the mechanisms of cognition, just as the mechanisms of simpler behavior have been addressed in neuroethology. As an example for this mechanistic view of cognitive abilities, I will discuss the view–graph theory of cognitive maps. I will argue that spatial cognitive abilities can be explained by scaling up simple, stereotyped mechanisms of spatial behavior. This evolutionary view of cognition is supported by two types of empirical evidence: Robot experiments show that the simple mechanisms are in fact sufficient to produce cognitive behavior while behavioral experiments with subjects exploring a computer graphics environment indicate that stereotyped and cognitive mechanisms co–exist in human spatial behavior.

1 Introduction: Cognition and Neurobiology

In the theory of brain function and behavior, two major traditions can be distinguished. The first one, which may be called the computational approach, attempts to describe mental processes as judgements, symbols or logical inference. The second one focusses on issues such as control, signal flow in networks or feedback; I will call it the systems approach here. In the field of perception, this distinction is rather old, dating back at least to Hermann von Helmholtz' "unconscious inferences" (Helmholtz 1896) on the one side, and to the Gestaltists on the other. Both approaches have different merits. Computational approaches lend themselves easily for modelling behavioural competences including psychophysical data (Marr 1982, Mallot 1998), without considering the biophysical processes in the brain that underly these competences. Systems explanations, on the other hand, are closer to the neurophysiological correlates of mental processes and are therefore useful in modelling neural activities. Bridging the gap between signal flow in the brain and behavioural competences is not easy, however.

Both approaches can be applied to all aspects of brain function. This is quite clear for perception, and the respective approaches have been mentioned above. It is much less clear for higher competences such as cognition; indeed, cognition is often seen as being accessible only with the computational approach. The assumed relation of cognition and computation is two–fold: first, cognition is often defined introspectively as inference and problem solving, i.e. by notions taken from the computational approach. Second, the explanations offered by the computational approach even for simple processes such as early vision are often formulated in cognitive terms. In fact, Helmholtz' notion of "unconscious inferences" is a clear example of this. For these reasons, researchers who are not interested in the computational approach tend to ignore cognition, whereas others, focussing on computation, might think that all central processing is somehow cognitive.

There is good reason to believe that this confusion can be avoided if cognition is defined as an observable behavioral phenomenon, not as a mechanism of mental processing (Mallot 1997). In the following sections of this paper, I shall argue that cognition can be defined by the complexity of the behavioral competences it supports (Section 2), that mechanisms for non–cognitive behavior can be scaled up to bring about cognitive competences (Section 3), and that in human spatial cognition, cognitive and non–cognitive levels of competence coexist simultaneously (Section 4). I hope that the results and ideas reviewed in this paper make a contribution to an evolutionary theory of higher–level behavior.

2 Complexity of Behavior

2.1 Four levels

Behavior of animals (or robots) may be divided into a number of levels of complexity four of which are illustrated in Fig. 1. In the simplest stimulus–reaction situations found already in individual cells, sensor and effector cannot be distinguished or are so close together that no neural transmission is necessary. This level is not illustrated in Figure 1. By the separation of sensor and effector, reflex–like behaviors arise (level 1 in Fig. 1). Clear illustrations of the surprisingly rich behavioral abilities of systems endowed with such simple machinery are given in the thought experiments of Braitenberg (1984). The most commonly known example is probably his "vehicle 3b", a platform with two sensors in front and two drives in the rear receiving inhibitory input from the sensor located on the opposite ("contralateral") side of the vehicle. If the sensors respond to stimuli originating from certain sources (e.g., light bulbs), this system will avoid the sources since the sensor closer to the source will receive stronger input; in turn, the motor on the side of the source will turn faster, resulting in a turn away from the source. In a corridor, the same mechanism will result in centering behavior. Other behaviors that can be realized by this simple stimulus–response wiring are "attacking" of sources, or approach and "docking".

The second level is reached when sensory information from various sensors is integrated by interneurons or interneuron networks. New sensory input interacts with the activity pattern in the interneurons which forms a kind of working memory. An instructive example of spatio–temporal integration without longterm memory is navigation by path integration. This type of navigation behavior can be implemented already on level 2, by continuously updating a representation of the starting point of a path with the instantaneous motion of the agent (see also below).

Learning can be defined as the change of behavior due to prior experience. On level three, this is achieved by plastic modifications of the spatio–temporal processing. Examples include the fine–tuning of motor programs in skill learning, the association of landmarks (snapshots) and movements in route navigation, or the learning of trigger stimuli. Memory is long–term, but the resulting changes of behavior are still stereotyped in the sense that one stimulus–response pair is simply replaced by another.

Cognitive behavior (level 4 in Figure 1) is characterized by goal–dependent flexibility. The behavior of the agent does no longer depend exclusively on the sensory stimulus and whatever prior experience it might have, but also on the goal which is currently pursued. In the case of navigation, the crucial situation is the behavior at a bifurcation where one of two motion decisions can be chosen. If the agent is able to do this correctly with respect to a distant, not currently visible goal, we will call its behavior cognitive. The difference between route memory and cognitive maps has been lucidly elaborated by O'Keefe & Nadel (1987).

There are also higher levels of complexity in behavior which are not included in Fig. 1. As an example, consider the behavioral definition of consciousness used in the work of Povinelli & Preuss (1995): in their view, consciousness is involved if behavioral decisions are based on assumptions of what some other individual might know or plan to do.

2.2 Application to spatial behavior

Spatial behaviour includes a wide variety of competences that can be classified based on the type and extend of memory they require; for reviews see O'Keefe & Nadel (1978), Collett (1992), Trullier et al. (1997) and Franz & Mallot (1998). With respect to the four levels of complexity given in Fig. 1, the following classification can be given:

Without memory (no remembered goal). Simple tasks like course stabilization, efficient grazing and foraging, or obstacle avoidance can be performed without memory. Traditionally, memory–free orientation movements are called "taxes" (Kühn 1919; see also Tinbergen 1951, Merkel 1980). An example is illustrated in Fig. 2a: an observer with two laterally displaced sensors can travel a straight line between two sources by balancing the sensory input in both detectors. A possible mechanism for this behavior is of course Braitenberg's (1984) "vehicle 3b" discussed already in Section 2.1. While detailed classifications of various types of taxis (see Merkel 1980 for review) have not proved very useful in experimental research, the general concept is central to the understanding of the mechanisms of behavior.

Working memory of a home position is required for path integration (Fig. 2b). Current egomotion estimates are vectorially added to an ego-

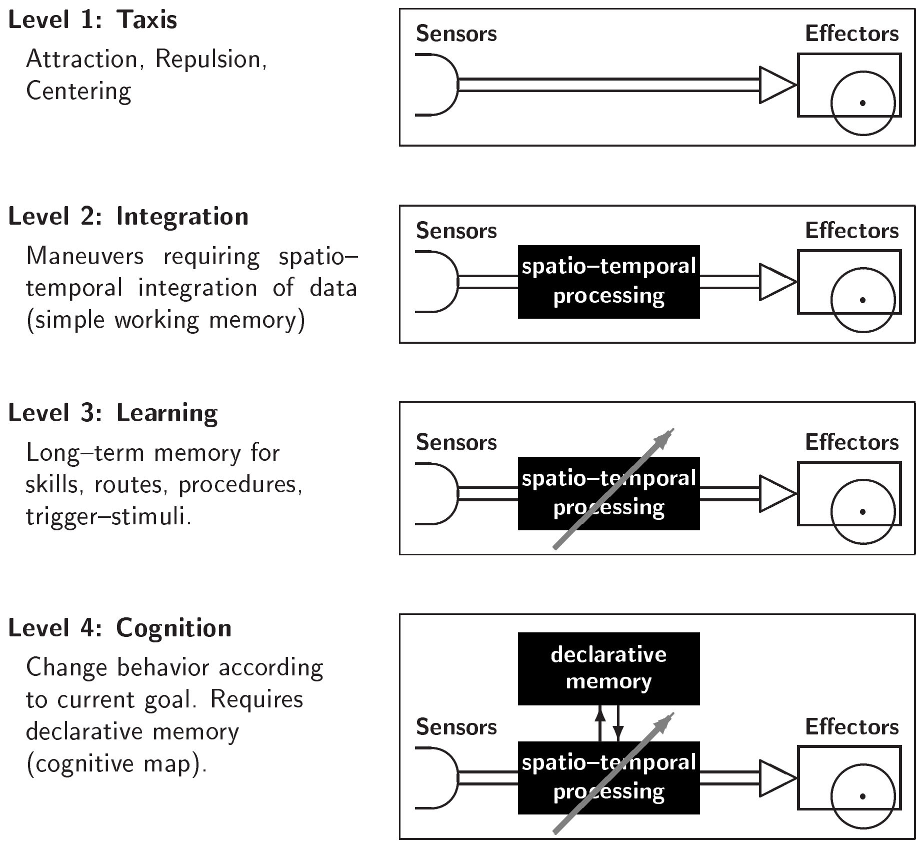


Figure 1: Four levels of complexity of behavior. Level 1 allows reflex–like behavior based on the wiring and the current sensory input. Level 2 includes spatio–temporal processing of inputs arriving at different sensors and at different times. Learning is introduced at Level 3, by allowing for plasticity of the spatio–temporal processing. Except for this plasticity, behavior is still determined completely by the sensory input. At level 4, one sensory input may elicit different behaviors depending on the current goal of the agent. For further explanations see text.

centric representation of the start position thus that the current distance and direction of the start point are always available. This memory is of the working memory type since the places visited or the path travelled are not stored (see Maurer & Séguinot, 1995, for review).

Long–term memory is involved in landmark–based mechanisms, which use a memory of sensory information characteristic of a given place ("local position information"). In *guidance*, motions are performed such as to achieve or maintain some relation to the landmarks. In the example depicted in Fig. 2c, a so–called snapshot taken at the right position is stored in memory. By comparing the current view (visible from position A in Fig. 2c) to the stored reference view (position B), a movement direction can be calculated that leads to an increased similarity of current and stored snapshot (Cartwright & Collett, 1982; Franz et al. 1998b).

A second type of landmark based navigation uses a slightly richer memory. In addition to the snapshot characterizing a place, an action is remembered that the observer performes when recognizing the respective snapshot. In the simplest case, these actions are movements into specific directions (Fig. 2c), but more complicated behaviours such as wall following could also be attached to snapshot recognition (e.g., Kuipers & Byun, 1991). We will refer to this mechanism as "recognition–triggered response" (Trullier et al. 1997). Chains of recognition–triggered responses allow the agent to repeat routes through a cluttered environment. Note that recognition–triggered responses act much like the socalled sign or trigger stimuli (in German: Schlüsselreize) for innate behavior studied in classical ethology (Tinbergen 1951, Lorenz 1978).

Declarative memory is required to plan and travel different routes composed of pieces and steps stored in the memory. At each step, the movement decision will depend not only on the current landmark information, but also on the goal

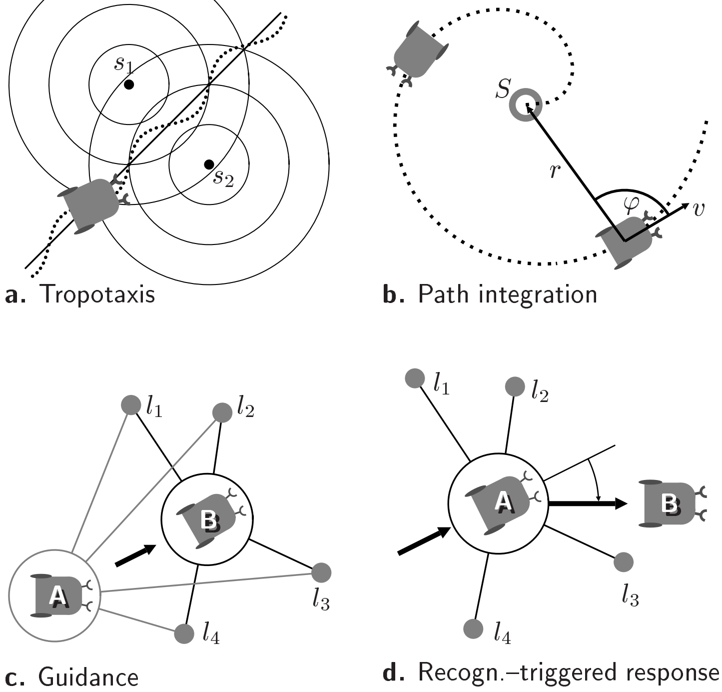


Figure 2: Basic mechanisms of spatial behavior. **a.** Tropotaxis. s_1 and s_2 denote sources that can be sensed by the agent. **b.** Path integration. S start position; (r, φ) current position of start point in egocentric coordinates. **c.** Guidance. The circles surrounding the vehicles symbolize the visual array of the respective position; $l_1, ..., l_4$ are landmarks. The "snapshot" visible at position **B** has been stored. At a location **A**, movement is such that the currently visible snapshot will become more similar to the stored one. **d.** In *recognition–triggered response* memory contains both a snapshot and an action associated with it. When the snapshot is recognized in **A**, an action such as a turn by some remembered angle is executed.

the navigator is pursuing. Following O'Keefe & Nadel (1978), we use the term cognitive map for a declarative memory of space; a cognitive map in this sense does not necessarily contain metric information nor does it have to be two–dimensional or "map–like" in a naive sense.

3 The view–graph approach to cognitive maps

In this section, we present a minimalistic theory of a cognitive map in terms of a graph of recognized views and movements leading the agent from one view to another. For a full account of this theory, see Schölkopf & Mallot (1995). The view–graph generalizes the route memory given as a chain of recognition–triggered responses to a cognitive map. A further generalization to open environments using also a guidance mechanism has been presented by Franz et al. (1998a).

3.1 Places, views, and movements

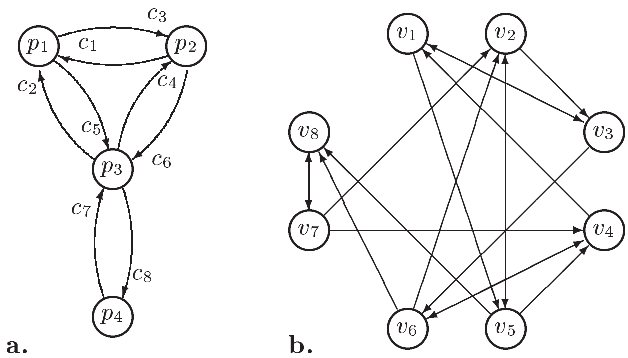
Consider a simple maze composed of *places* $p_1, ..., p_n$ and *corridors* $c_1, ..., c_m$ (Fig. 3a). One way to think of this maze is a graph where the places are the nodes and the corridors are the edges. We consider all corridors to be directional but allow for the existence of two corridors with opposite directions between any two nodes. When exploring this maze, the observer generates a sequence of movement decisions defining a path through the maze. In doing so, he encounters a sequence of *views* from which he wants to construct a spatial memory. In order to study the relation of views, places and movements, we make the following simplifying assumptions. First, we assume that there is a one–to–one correspondence between directed corridors and views. All views are distinguishable and there are no "inner views" in a place that do not correspond to a corridor. Second, one movement is selected from a finite (usually small) set at each time step.

With these assumptions, we can construct the *view–graph* that an observer will experience when exploring a maze (Fig. 3b,c). Its elements are:

- The **nodes** of the view–graphs are the views v_p, which, from the above assumption, are simply identical to the corridors in the place–graph.

- The **edges** of the view–graph indicate temporal coherence: two views are connected, if they can be experienced in immediate temporal sequence. The edges are labelled with the movements resulting in the corresponding view sequence.

The resulting adjacency matrix with movement labels is depicted in Fig. 3c. Note that all edges starting from the same node will have different movement labels. The unlabelled version of the



to view	v_1	v_2	v_3	v_4	v_5	v_6	v_7	v_8
v_1	t_o	–	g_b	g_l	–	t_r	–	–
v_2	–	t_o	–	t_l	g_b	g_r	g_l	t_r
v_3	g_b	g_r	t_o	–	t_l	–	–	–
v_4	–	t_r	–	t_o	g_l	g_b	g_r	t_l
v_5	g_l	g_b	t_r	–	t_o	–	–	–
v_6	t_l	–	g_r	g_b	–	t_o	–	–
v_7	–	–	–	–	–	–	t_o	g_b
v_8	–	t_l	–	t_r	g_r	g_l	g_b	t_o

c.

Figure 3: **a.** Simple maze shown as a directed graph with places p_i and corridors c_j. **b.** Associated view–graph where each node v_i corresponds to one view, i.e. one directed connection in the place graph. Only the edges corresponding to locomotions ("go–labels") are shown. Simpler plots of b. are possible but not required for our argument. **c.** Adjacency matrix of the view–graph with labels indicating the movement leading from one view to another. Go–labels (involving a locomotion from one place to another): g_l (go left), g_r (go right), g_b (go backward). Turn–labels (e.g., probing of corridors): t_l (turn left), t_r (turn right), t_o (stay).

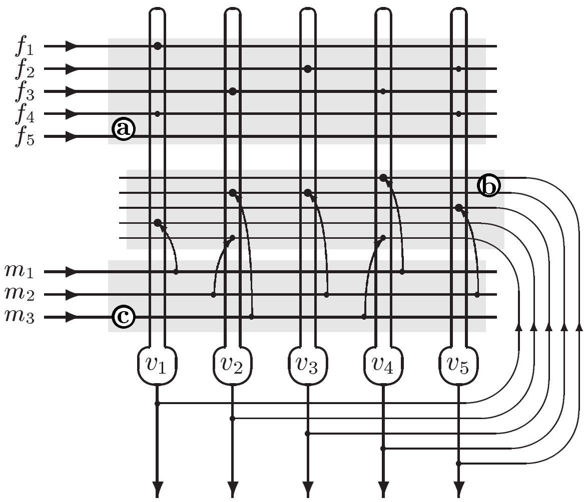


Figure 4: Wiring diagram of the neural network. $(f_1, ..., f_J)$: feature vector corresponding to current view. $m_1, ..., m_K$: motion input. $v_1, ..., v_N$: view cells. The dots in the view cell distal parts ("dendrites") symbolize synapitc wieights. **a.** Input weights $\rho_{nj} : f_j \to v_n$ subserve view recognition. **b.** Map layer weights $\alpha_{ni} : v_i \to v_n$ represent connections between views. They can be modified by facilitating weights $\beta_{k,ni}$ (**c.**), indicating that view v_n can be reached from v_i by performing movement m_k.

view–graph defined here is the *interchange graph* (e.g., Wagner 1970) of the place–graph.

The view–graph contains the same information as the place–graph. In fact, the place–graph can be recovered from the view–graph since each place corresponds to a complete bipartite subgraph of the view–graph (for a proof, see Schölkopf & Mallot 1995). Using the view–graph as a spatial memory, however, is a more parsimonious solution. Computation is simplified in two points: First, in order to construct a view–graph memory, it is not necessary to decide which views belong to the same place. Second, when reading out the labelled view–graph, the labels can be used directly as motion commands. This is due to the fact that in the view–graphs, labels are specified in an ego–centric way (e.g., "left", "right", etc.). In contrast, in a place–graph memory, labels have to be world centered ("north", "south", etc.). In order to derive movement decisions from world–centered labels, an additional reference direction or compass would be required.

3.2 Learning mazes from view sequences

A neural network for the learning of view–graphs from sequences of views and movements is shown in Fig. 4; for details see Schölkopf & Mallot (1995). The network consists of one layer of "view–cells" (v_n in Fig. 4) and a mixed auto– and heteroassociative connectivity. View input enters the network as a feature vector (f_j in Fig. 4). For each view–cell v_n, a set of input weights ρ_{nj} subserves view recognition. The input weights ρ are learned during exploration of the maze by a competitive learning rule: if unit v_i is the most active unit at time t (the "winner" neuron), its input weights will be changed in a way such that next time the same stimulus occurs, the unit will react even stronger. This learning rule is similar to the one introduced by Kohonen (1982) for the self–organization of feature maps. Unlike the self–organizing feature map, learning does not spread to neighboring neurons in our network. Thus, adjacency in our network will not reflect view similarity.

View–recognition is facilitated by neighborhood information or expectations implemented by the feedback connections shown in Fig. 4b. If view cell v_i is active at a time t, it will activate other view cells by means of the "map–weights" α_{ni}. These map weights thus increase the probablity that a unit connected to the previous winner unit will be the most active one in the next time step. Map weights reflect the adjacency in the view–graph. Map weights are learned by the simple rule that the weight connecting the past and the present winner unit will be increased in each time step.

The labels of the view–graph are implemented in the network by a set of "facilitating weights" $\beta_{k,ni}$ shown in Fig. 4c. They receive input from a set of movement units m_k whose activity represents the movement decisions ("left", "right" or some finer sampling) of the agent. If a movement is performed, the according movement cell (m_k, say) will be active. If a positive weight $\beta_{k,ni}$ has been learned, the map weight α_{ni} will be facilitated. By this mechanism, the activity bias distributed to all neighbors of unit v_i by the map weights will be focussed to the one neighbor that can in fact be reached by the present movement. During learning, the faciliating weight $\beta_{k,ni}$ is set to some constant value if movement m_k coincided with the last increase of α_{ni}; otherwise, $\beta_{k,ni}$ is zero.

Some simulation results obtained with this neural network include the following:

Convergence. For the toy–maze of Fig. 3a, a network with 8 input lines and 20 view cells (no movement input) converges in 60 presentation steps. By this time, view specificities together with the approriate map–layer connections have evolved.

View Recognition. The fead–back connections in the network (Fig. 4b) help recognize the views. If noise is added to the views, the map layer weights reduce the signal–to–noise ratio required for recognition by a factor of about 2 (3 dB). This indicates that the topological structure stored in the map layer weights is used to distinguish similar, but distant views.

Maze Reconstruction. From the weight matrix, we derive an estimate of the adjacency matrix A of the view graph by deleting the all–zero rows and columns, thresholding the remaining entries, and suitable reordering of the rows and columns. Using the reconstruction method described in Schölkopf & Mallot (1995), the underlying place–graph could be recovered after about 20 learning steps. This is due to a redundancy of the view–graph: each place corresponds to an complete bipartite subgraph consisting of the entries and exits of the place. Thus, if view a is connected to views b and c and view d is connected to view b, a connection from view d to view c can be predicted. From this property of the view-graph, optimal strategies for further exploration of the maze can be derived.

Robot Navigation. A modified Khepera robot was used to explore a hexagonal maze by Mallot et al. (1995). The robot was equipped with "two–pixel–vision". i.e. two infra–red sensors looking downward to the textured floor. The sequence of black and white signals obtained when approaching a junction was used as view input to the network. The robot was able to explore a maze of 12 places and 24 "views" in about an hour. Afterwards, shortest paths to all views could be planned and travelled.

3.3 View–graphs in open environments

The theory presented so far applies to mazes, i.e. environments with discrete decision points and strong movement restrictions. In oder to apply the view–graph approach to open environments, two problems have to be addressed: First, discrete points or centers have to be defined based on sensory saliency and strategic importance (e.g., gateways). Second, a homing mechanism has to be implemented that allows the agent to approach the centers from a certain neighborhood or catchment area.

View–based solutions to both problems have been presented by Franz et al. (1998a,b). An agent starts the exploration of a maze by recording the view visible from its initial position. During exploration, the agent continuously monitors the difference between the current view of the environment and the views already stored in memory. If the difference exceeds a threshold, a new view is stored; in the view–graph, this new view is connected to the previously visited one. The second problem, approaching a familiar view, is solved by scene–based homing (see Fig. 2c): From a comparison of current and stored view, the agent calculates a movement direction in which to move in order to increase the similarity between stored and current view. During exploration, this second mechanism is also used for "link verification": if the agent encounters a view similar to one stored in its memory, it tries to home to this view. If homing is successful, i.e. if stored and curred view get sufficiently similar, a link is added to the view–graph. The mechanism has been tested with a robot us-

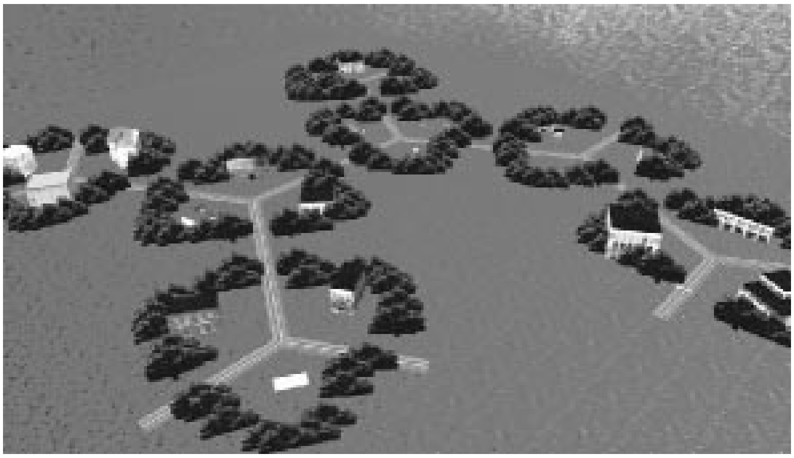


Figure 5: Aerial view of Hexatown. The white rectangle in the left foreground is view 15, used as "home"–position in our experiments. The aerial view was not available to the subjects.

ing a panoramic vision device navigating an arena with model houses.

At first glance, the view–graph approach might not seem natural for open environments. However, in a view–based scheme, the manifold of all pictures obtainable from all individual positions and viewing directions in the arena (the "view manifold") cannot be stored completely. The sketched exploration scheme is an efficient way to sample the view–manifold and represent it by a graph whose mesh size is adapted to the local rate of image change, i.e. to the information content of the view–manifold. The threshold for taking a new snapshot has to be set in a way to make sure that the catchment areas of adjacent nodes have sufficient overlap.

4 Mechanisms of human spatial behavior

We have tested the view–based approach to cognitive maps in a series of behavioral experiments using the technology of virtual reality (Bülthoff et al. 1997, van Veen et al. 1998). The basic structure of the experimental environment, called Hexatown, is depicted in Fig. 5 (Gillner 1997, Gillner & Mallot 1998). It consists of a hexagonal raster of streets where all decision points are three–way junctions. Three buildings providing landmark information are located around each junction. Subjects can move through the environment by selecting "ballistic" movement sequences (60 degree turns or translations of one street segment) by clicking the buttons of a computer mouse (see Gillner & Mallot, 1998, for details). Areal views are not available to the subjects. In the version appearing in Fig. 5, the information given to the subjects is strictly view–based, i.e. at any one time, no more than one of the landmark objects is visible.

The most important results obtained with the Hexatown environment are the following:

Map knowledge. Subjects can acquire map knowledge in a virtual maze. In a series of search tasks where subjects were released at some position and had to find a landmark shown to them as a print–out on a sheet of paper, subjects were able to infer the shortest ways to the goal in the later search tasks (Gillner & Mallot, 1998). Each individual search corresponded to a route learning task; the advantage for later search tasks indicates that some goal–independent knowledge was transferred from the known routes to the novel tasks, which is an indication of map knowledge in the sense of O'Keefe & Nadel (1978). Other indications of map knowledge were the subjects' ability to estimate distances in the maze and the sketch maps drawn as the last part of the experiment.

Stereotyped behavior. In addition to map knowledge, subjects have also a more stereotyped form of knowledge, i.e. associations between views and movements, or recognition–triggered responses (Gillner & Mallot 1998). By evaluating the sequences of views and movement decisions generated by the subjects when navigating the maze, we found a clear tendency to simply repeat the previous movement decision when returning to an already known view. This implies that subjects use the strategy of recognition–triggered response, which is a stereotyped strategy useful in route navigation.

Place vs. view. In systematic experiments with landmark transpositions, we could show that recognition–triggered response is triggered by the recognition of individual objects, not of the configurations of objects making up a place (Mallot & Gillner 1998). After learning a route, each object together with its retinal position when viewed from the decision point (left peripheral, central, right peripheral) is associated with a movement triggered by the recognition of this object. When objects from different places are recombined in a way that their associated movement votes are consistent, no effect in subjects' performance was found. If however, objects are combined in inconsistent ways (i.e. if their movement votes differ), subjects get confused and the number of erroneous motion decisions increases. It is interesting to note that this result is different from findings in guidance tasks (Poucet 1993, Jacobs et al. 1998), where the configuration of all landmarks at a place seems to be stored in memory.

Interaction of cues. In order to study different types of landmark information, we added distal landmarks to the environment, placed on a mountain ridge surrounding Hexatown (Steck & Mallot 1998). In this situation, various strategies can be used to find a goal: subject could ignore the distant landmarks alltogether, they could rely on the distant ones exclusively, or they could use both types in combination. We tried to identify these strategies by replacing the distant landmarks after learning, so that different patterns of movement decisions can be expected for each of the above strategies. We found that different strategies are used by different subjects and by the same subject at different decision points. When removing one landmark type from the maze after learning, subjects who had relied on this landmark type earlier were still able to use the previously neglected type. This indicates that both types of information were present in memory but one was ignored in the cue–conflict situation.

5 Discussion: Evolutionary scaling of spatial behavior

The theoretical and experimental work gathered in this paper is motivated by the following ideas:

1. Spatial behavior includes a fair number of different competences, ranging from stereotyped orientation behavior to way–finding and path planning, and on to communication about space.

2. These competences and their underlying mechanisms form a hierarchy not only in the sense of increasing complexity but also in the sense of the evolution of behavior. Simple mechanisms can be scaled–up to realize more complex competences (see also Mallot et al., 1992, for a discussion of "preadaptations" in the evolution of intelligent systems). We have shown that recognition–triggered response can be used as a buidling block for a cognitive map and we would like to suggest as a working hypothesis that this relation reflects the course of evolution.

3. In this mechanistic/evolutionary view, the distinction between stereotyped and cognitive behavior, clear–cut as it may seem when looking af Figure 1, looses much of its strength. If there are evolutionary connections between recognition–triggered response and cognitive maps, why shouldn't they coexist in the same navigating system? Our data from the Hexatown experiment show that this is in fact the case.

References

Braitenberg, V., 1984. *Vehicles. Experiments in Synthetic Psychology.* Cambridge, MA: The MIT Press.

Bülthoff, H. H., Foese-Mallot, B. M. & Mallot, H. A., 1997. Virtuelle Realität als Methode der modernen Hirnforschung. In H. Krapp & T. Wägenbauer (eds.), *Künstliche Paradise — Virtuelle Realitäten. Künstliche Räume in Literatur–, Sozial– und Naturwissenschaften* (pp. 241 – 260). München: Wilhelm Fink Verlag.

Cartwright, B. A. & Collett, T. S., 1982. How honey bees use landmarks to guide their return to a food source. *Nature*, **295**, 560 – 564.

Collett, T., 1992. Landmark learning and guidance in insects. *Philosophical Transactions of the Royal Society (London) B*, **337**, 295 – 303.

Franz, M. O. & Mallot, H. A., 1998. Biomimetic robot navigation. Technical Report 65, Max-Planck-Institut für biologische Kybernetik, Tübingen.

Franz, M. O., Schölkopf, B., Mallot, H. A., & Bülthoff, H. H., 1998a. Learning View Graphs for Robot Navigation. *Autonomous Robots*, **5**, 111 – 125.

Franz, M. O., Schölkopf, B., Mallot, H. A., & Bülthoff, H. H., 1998b. Where did I take that snapshot? Scene–based homing by image matching. *Biological Cybernetics*, **79**, 191 – 202.

Gillner, S., 1997. *Untersuchungen zur bildbasierten Navigationsleistung in virtuellen Welten.* Ph.D. thesis, Fak. Biologie, Universität Tübingen.

Gillner, S. & Mallot, H. A., 1998. Navigation and Acquisition of Spatial Knowledge in a Virtual Maze. *Journal of Cognitive Neuroscience*, **10**, 445 – 463.

Helmholtz, H. v., 1896. *Handbuch der physiologischen Optik* (2. edition). Hamburg, Leipzig: Voss.

Jacobs, W. J., Thomas, K. G. F., Laurance, H. E., & Nadel, L., 1998. Place learning in virtual space II: Topgraphical relations as one dimension of stimulus control. *Learning and Motivation*, **29**, 288 – 308.

Kohonen, T., 1982. Self–Organized Formation of Topologically Correct Feature Maps. *Biological Cybernetics*, **43**, 59 – 69.

Kühn, A., 1919. *Die Orientierung der Tiere im Raum*. Jena: Gustav Fischer Verlag.

Kuipers, B. J. & Byun, Y.-T., 1991. A robot exploration and mapping strategy based on a semantic hierarchy of spatial representations. *Journal of Robotics and Autonomous Systems*, **8**, 47 – 63.

Lorenz, K., 1978. *Vergleichende Verhaltenforschung*. Wien: Springer Verlag.

Mallot, H. A., 1997. Behavior–oriented approaches to cognition: theoretical perspectives. *Theory in Biosciences*, **116**, 116 – 220.

Mallot, H. A., 1998. *Sehen und die Verarbeitung visueller Information. Eine Einführung*. Wiesbaden: Vieweg Verlag.

Mallot, H. A., Bülthoff, H. H., Georg, P., Schölkopf, B., & Yasuhara, K., 1995. View–based cognitive map learning by an autonomous robot. In *Proceedings of the International Conference on Artificial Neural Networks (ICANN) '95 and Neuro–Nîmes '95*, vol. II, pp. 381 – 386 Paris. EC2 & Cie.

Mallot, H. A. & Gillner, S., 1998. View–based vs. place–based navigation: What is recognized in recognition–triggered responses? Technical Report 64, Max-Planck-Institut für biologische Kybernetik, Tübingen.

Mallot, H. A., Kopecz, J., & von Seelen, W., 1992. Neuroinformatik als empirische Wissenschaft. *Kognitionswissenschaft*, **3**, 12 – 23.

Mallot, H. A. & Logothetis, N. K., 1998. Mechanisms of Cognition (Introductory Remarks). In N. Elsner & R. Wehner (eds.), *New Neuroethology on the Move*, vol. 1, pp. 278 – 283 Stuttgart. 26th Göttingen Neurobiology Conference, Thieme Verlag.

Marr, D., 1982. *Vision*. San Francisco: W. H. Freeman.

Maurer, R. & Séguinot, V., 1995. What is Modelling For? A Critical Review of the Models of Path Integration. *Journal of theoretical Biology*, **175**, 457 – 475.

Merkel, F. W., 1980. *Orientierung im Tierreich*. Stuttgart, New York: Gustav Fischer Verlag.

O'Keefe, J. & Nadel, L., 1978. *The hippocampus as a cognitive map*. Oxford, England: Clarendon.

Poucet, B., 1993. Spatial cognitive maps in animals: New hypotheses on their structure and neural mechanisms. *Psychological Review*, **100**, 163 – 182.

Povinelli, D. J. & Preuss, T. M., 1995. Theory of Mind — Evolutionary History of a Cognitive Specialization. *Trends in Neurosciences*, **18**, 418 – 424.

Schölkopf, B. & Mallot, H. A., 1995. View–based cognitive mapping and path planning. *Adaptive Behavior*, **3**, 311 – 348.

Steck, S. D. & Mallot, H. A., 1998. The role of global and local landmarks in virtual environment navigation. Technical Report 63, Max-Planck-Institut für biologische Kybernetik, Tübingen.

Tinbergen, N., 1951. *The Study of Instinct*. Oxford: Clarendon Press.

Trullier, O., Wiener, S. I., Berthoz, A., & Meyer, J.-A., 1997. Biologically based artificial navigation systems: Review and Prospects. *Progress in Neurobiology*, **51**, 483 – 544.

Veen, H. A. H. C. v., Distler, H. K., Braun, S. J., & Bülthoff, H. H., 1998. Navigating through a virtual city: Using virtual reality technology to study human action and perception. *Future Generation Computer Systems*, **14**, 231 – 242.

Wagner, K., 1970. *Graphentheorie*. Mannheim, Wien, Zürich: Bibliographisches Institut.