

A Sensorimotor Map: Modulating Lateral Interactions for Anticipation and Planning

Marc Toussaint

mtoussai@inf.ed.ac.uk

School of Informatics, University of Edinburgh, Edinburgh, Scotland, U.K.

Experimental studies of reasoning and planned behavior have provided evidence that nervous systems use internal models to perform predictive motor control, imagery, inference, and planning. Classical (model-free) reinforcement learning approaches omit such a model; standard sensorimotor models account for forward and backward functions of sensorimotor dependencies but do not provide a proper neural representation on which to realize planning. We propose a sensorimotor map to represent such an internal model. The map learns a state representation similar to self-organizing maps but is inherently coupled to sensor and motor signals. Motor activations modulate the lateral connection strengths and thereby induce anticipatory shifts of the activity peak on the sensorimotor map. This mechanism encodes a model of the change of stimuli depending on the current motor activities. The activation dynamics on the map are derived from neural field models. An additional dynamic process on the sensorimotor map (derived from dynamic programming) realizes planning and emits corresponding goal-directed motor sequences, for instance, to navigate through a maze.

1 Introduction ---

Köhler's (1917) studies with monkeys were one of the first systematic investigations in the capability of planned behavior in animals. In one of his classic experiments, monkeys had to reach for a banana mounted below the ceiling. After many attempts in vain, one of the monkeys eventually exhibited the behavior that Köhler found so fascinating: the monkey retreated and sat quietly in a corner for minutes, staring at the banana and at some time also staring at a nearby table. It started to saccade several times between the banana and the table while still sitting quietly. Then it suddenly rushed up, grabbed the table, pulled it below the banana, mounted it, and jumped to grab the banana.

Reading these experiment scripts today, one realizes how little we know about the neural processes in the monkey's brain when Köhler read in its face the effort to reason about sequential behaviors to reach a goal. Classical (model-free) reinforcement learning approaches explicitly omit

internal models (Sutton, & Barto, 1998; see also Majors & Richards, 1997). More recent studies in the cognitive sciences converge to the postulate that nervous systems use internal models to perform predictive motor control, imagery, and planning in a way that involves a simulation of actions and their perceptual implications (Grush, 2004). Based on experiments with humans, who were asked to imagine the way from a starting position in a maze to a goal position, Hesslow (2002) formulates three assumptions that may explain a simulation theory of cognitive functions: (1) behavior can be simulated by activating motor structures, as during an overt action, but suppressing its execution; (2) perception can be simulated by internal activation of sensory cortex, as during normal perception of external stimuli; and (3) both overt (executed) and covert (suppressed) actions can elicit perceptual simulation of their normal consequences.

The evidence in favor of internal models and the hypotheses developed in cognitive science raise the challenge to propose concrete models of how neural systems are capable of these processes. Such systems must be able to anticipate the sensorial implications of motor activities, but they also must account for planned, goal-oriented behavior.

The sensorimotor map we propose in this letter provides mechanisms to self-organize a representation of sensorimotor data that encodes the dependencies between motor activity and predictable changes of stimuli (see also Toussaint, 2004). The self-organization process largely adopts the classical approaches to self-organizing neural stimulus representations (von der Malsburg, 1973; Willshaw & von der Malsburg, 1976; Kohonen, 1995) and their extensions with respect to growing representations (Carpenter, Grossberg, Markuzon, Reynolds, & Rosen, 1992; Fritzke, 1995; Bednar, Kelkar, & Miiikkulainen, 2002) and temporal dependencies (Bishop, Hinton, & Strachan, 1997; Euliano & Principe, 1999; Somervuo, 1999; Wiemer, 2003; Varsta, 2002; Klemm & Alstrom, 2002). However, unlike previous self-organizing maps, our model couples sensor and motor signals in a joint representational layer.

The activation dynamics on the sensorimotor map are adopted from dynamic field models of a homogeneous, laterally connected neural layer (Amari, 1977). In the language of neural fields, the anticipation of a new stimulus corresponds to a shift of the activity peak, which is induced by a modulation of the lateral connection strengths. A key ingredient of our model is that the modulation depends on the current motor activities. A motor representation is coupled to the neural field by modulating the lateral connectivity instead of connecting directly to the neural units. By this mechanism, different motor activities lead to different shifts of the peak. The coupling encodes all the information necessary for anticipating a stimulus change depending on the motor activations and also for planning goal-directed motor sequences. On the sensorimotor map, an additional dynamic process similar to spreading activation dynamics (Bagchi, Biswas, & Kawamura, 2000) accounts for planning. The same coupling to the motor

representation allows the system to emit motor excitations that execute the plan.

The next section briefly recalls the relevant aspects of standard neural field dynamics. Section 3 gives an overview of the considered architecture. The sensorimotor map and how it couples to sensor and motor representations is introduced in section 4. Section 5 shows how the topology and parameters of the sensorimotor map can be learned online from data gathered during sensorimotor exploration. A demonstration of anticipation with the sensorimotor map is given in section 6, while section 7 introduces and demonstrates planning. In section 8, we briefly address possible extensions of the basic model before discussing related work in more detail in Section 9. A discussion concludes.

2 Neural Field dynamics

Amari (1977) investigated a spatially homogeneous neural field as an approximation of a dense layer of interconnected neurons. His main interest was in a theory of the dynamics of activity pattern formation on such substrates. The lateral connectivity is assumed to induce local excitation and widespread inhibition, as typically described with a Mexican hat–type interaction kernel. The most elementary interesting stable solution to such a dynamic system is the single peak solution (also called activity bump or packet), where the activity is localized and stabilized around a center while the widespread inhibition emitted from the peak inhibits any spontaneous activation in the neighborhood. This simple solution has some important functional properties: if the peak is induced by a stimulus, it stabilizes its representation against noise; it may even stabilize the representation when the stimulus vanishes or is temporally occluded; it fuses two nearby stimuli while implementing a competition between distal stimuli; and it exhibits some delay to shift the peak to a new position when the stimulus switches (hysteresis). These properties make the model appealing for sensory processing and decision making, as well as motor control, where the dynamics effectively allow the system to filter noisy signals, decide among conflicting signals, and stabilize such decisions (Erlhagen & Schöner, 2002). Consequently, neural fields also find application in motor control and robotics problems (Schöner & Dose, 1992; Schöner, Dose, & Engels, 1995; Iossifidis & Steinhage, 2001; Dahm, Bruckhoff, & Joubin, 1998; Bergener et al., 1999).

We introduce here a discrete implementation of such a neural field, following Erlhagen and Schöner (2002). In this implementation, the activation m_i of a unit i (denoted by m to anticipate the meaning of motor activations) is governed by the dynamics

$$\tau_m \dot{m}_i = -m_i + h_m + A_i + \sum_j w_{ij} \phi(m_j) + [\xi \sim \mathcal{N}(0, \rho_m)]. \quad (2.1)$$

Here, τ_m is the timescale of the dynamics, h_m the resting level, A_i some feedforward input to unit i , ξ a gaussian noise term with variance ρ_m , and $\phi(m)$ a sigmoid. We choose

$$\phi(m) = \hat{m} = \begin{cases} 0 & m < 0 \\ m & 0 \leq m \leq 1 \\ 1 & m > 1 \end{cases} \quad (2.2)$$

as a simple parameterless, piecewise-linear sigmoid.

The crucial term in these dynamics is the interaction strength w_{ij} between units i and j . In spatially homogeneous neural fields, this strength is usually assumed to depend on only the distance between the locations \mathbf{r}_i and \mathbf{r}_j of the two neurons. Namely, for short distances, the interaction is excitatory, while for longer distances, it is inhibitory:

$$w_{ij} = w_E \exp \frac{-(\mathbf{r}_i - \mathbf{r}_j)^2}{2\sigma_E^2} - w_I. \quad (2.3)$$

The parameters here are the strengths of excitation (w_E) and inhibition (w_I), and the width σ_E of the excitatory range.

We generally omit indicating the time dependence of dynamic variables except when we need to refer to the time steps of the Euler integration $m_i^{(t)} = m_i^{(t-1)} + \dot{m}_i^{(t)}$ that we use to simulate the dynamics.

3 Overview of the Sensorimotor Architecture

Figure 1 displays the sensorimotor architecture that we will use in the experiments. The architecture is composed of three layers. The bottom layer is an arbitrary sensor representation. In the experiments, the representation will comprise either 2 units for the x - and y -coordinates of a limb or 40 units encoding range sensor data from a maze.

The top layer is the motor representation, which we choose to be a one-dimensional cyclic neural field. Different units in the field will encode different bearing directions of movements. The dynamics of these units are exactly as given in equation 2.1; the “distance” $|\mathbf{r}_i - \mathbf{r}_j|$ between two units that determines the excitatory kernel in equation 2.3 is taken as the minimal distance on the circle, measured by how many units are between j and i .

The central layer is the sensorimotor map governed by equation 4.1 given below. The key architectural feature is that the motor units project to lateral connections (ij) between two sensorimotor units j and i by multiplicatively modulating the signal transmission of that lateral connection. In contrast, sensor units project directly to sensorimotor units, as it is typical for self-organizing maps.

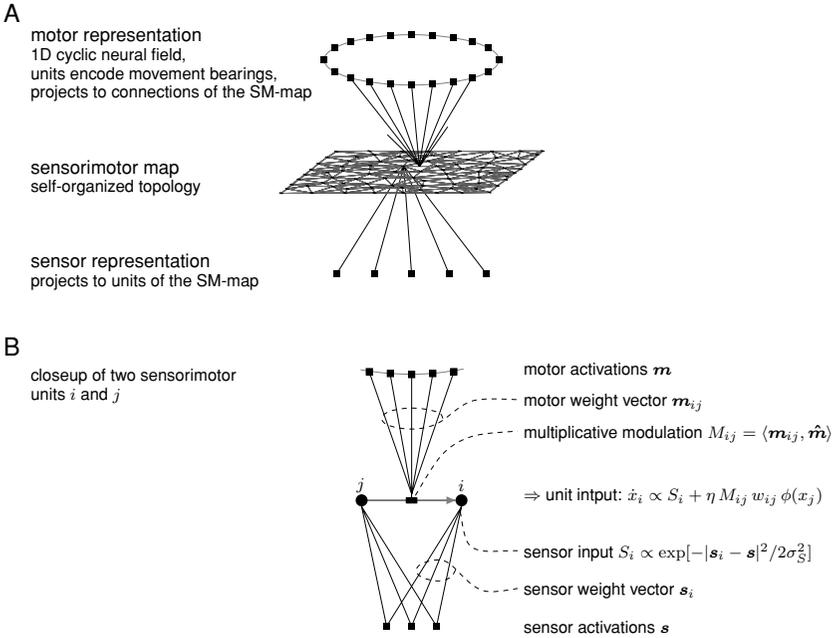


Figure 1: Schema of the considered architecture. (A) The bottom layer is a sensor representation, projecting to units of the sensorimotor map via gaussian kernels. The top layer is a motor representation that projects to lateral connections (ij) between sensorimotor units j and i . (B) This coupling induces a multiplicative modulation of the lateral interactions in the sensorimotor map, which depends on the current motor activations.

4 Modulating the Lateral Interactions

The core of the architecture is the sensorimotor map. Its activation dynamics is very similar to those of neural fields and reads

$$\tau_x \dot{x}_i = -x_i + h_x + S_i + \eta \sum_j [M_{ij} w_{ij} - w_l] \phi(x_j) + [\xi \sim \mathcal{N}(0, \rho_x)]. \quad (4.1)$$

As for the neural field, the first term $-x_i$ induces an exponential relaxation of the dynamics, the second term h_x is the resting level, and the third term S_i is a feedforward input from the sensor representation to unit i . We assume that the sensorial input is given as a (unnormalized) gaussian kernel,

$$S_i = \exp \frac{-(\mathbf{s}_i - \mathbf{s})^2}{2\sigma_S^2}, \quad (4.2)$$

that compares the input weight vector s_i (or codebook vector) of the unit i with the current sensor activations \mathbf{s} .

The fourth term describes the lateral interactions between units in the sensorimotor map. The lateral topology is not necessarily homogeneous but should reflect the topology of the state space and possible state transitions and is given by the lateral weights w_{ij} . In this article, we assume that $w_{ij} = 0$ if there exists no connection and $w_{ij} = 1$ if there exists one (see Toussaint, 2004, and section 8 for a version where w_{ij} is continuous and learned with a temporal Hebb rule). The parameter w_I specifies the global inhibition.

The crucial difference to a standard neural field is the modulation M_{ij} of the lateral interactions. This modulation is how motor signals couple into the sensorimotor map. More precisely, we assume that

$$M_{ij} = \langle \mathbf{m}_{ij}, \hat{\mathbf{m}} \rangle, \quad (4.3)$$

which is the scalar product of the weight vector \mathbf{m}_{ij} and the current motor activations $\hat{\mathbf{m}}$. Thus, lateral interactions are modulated multiplicatively depending on the current motor activation.

The weight vector \mathbf{m}_{ij} , which is associated with every lateral connection (ij), could be thought of the codebook vector of that connection. In a sense, lateral connections “respond” to certain motor activations. Due to the multiplicative coupling, a lateral connection contributes to lateral interaction only when the current motor activity “matches” the weight vector of this connection.

Biologically plausible implementations of such modulation are, for example, pre- or postsynaptic inhibition of the signal transmission. In the case of presynaptic inhibition (Rose & Scott, 2003), synapses attach directly to the presynaptic terminal of other synapses, thereby modulating their transmission. In the case of postsynaptic inhibition (shunting inhibition), inhibitory synapses attach to branches of the dendritic tree near the soma, thereby modulating the transmission of the dendritic input accumulated at this dendritic branch (Abbott, 1991). Generally, modulation is a fundamental principle in biological neural systems (Phillips & Singer, 1997). The modulation may also be regarded as a special variant of sigma-pi neural networks (Mel, 1990; Mel & Koch, 1990).

5 Learning the Sensorimotor Map

The self-organization and learning of the sensorimotor map combines standard techniques from self-organizing maps (von der Malsburg, 1973; Willshaw & von der Malsburg, 1976; Kohonen, 1995) and their extensions with respect to growing representations (Carpenter et al., 1992; Fritzsche, 1995) and the learning of temporal dependencies in lateral connections (Bishop et al., 1997; Wiemer, 2003). The free variables that need to be adapted

are (1) the number of units in the map and their lateral connectivity and (2) the weight vectors s_i and m_{ij} coupling to the sensor and motor layers, respectively. Except for the adaptation of the motor coupling m_{ij} , all the adaptation mechanisms are standard, and we keep their description brief.

The topology. There already exist numerous techniques for the self-organization of representational maps, mostly based on the early work on self-organizing maps (von der Malsburg, 1973; Willshaw & von der Malsburg, 1976; Kohonen, 1995) or vector quantization techniques (Gersho & Gray, 1991). We prefer not to predetermine the state space topology but learn it, and hence adopt the technique of growing neural gas (Fritzke, 1995) to self-organize the lateral connectivity and that of fuzzy ARTMAPs (Carpenter et al., 1992) to account for the insertion of new units when the representation needs to be expanded. We detect novelty when the difference between the current stimulus s and the best matching weight vector s_i becomes too large. We make this criterion more robust against noise by using a low-pass filter (leaky integrator) of this representation error. More precisely, if i^* is the unit with the best match, $i^* = \operatorname{argmax}_i S_i$, we integrate the error measure e_{i^*} via $\tau_e \dot{e}_{i^*} = -e_{i^*} + (1 - S_{i^*})$. Note that $S_{i^*} = 1 \iff s_{i^*} = s$. Whenever this error measure exceeds a threshold $v \in [0, 1]$ termed *vigilance*, $e_{i^*} > v$, we generate a new unit j and reset the error measures, $e_{i^*} \leftarrow 0$, $e_j \leftarrow 0$. Exactly as for growing neural gas, we add new lateral connections between i^* and $j^* = \operatorname{argmax}_{i \neq i^*} S_i$ if they were not already connected. To organize the deletion of lateral connections, we associate an “age” a_{ij} with every connection, which is increased at every time step by an amount of $M_{ij} \phi(x_j)$ and is reset to zero when i and j are the best and second-best matching units. If a connection’s age exceeds a threshold a_{\max} , the connection is deleted.

The sensor and motor coupling. Standard self-organizing maps adapt the input weight vectors s_i of a unit i in a Hebbian way such that s_i converges to the average stimulus for which i is the best matching unit. To avoid introducing additional learning parameters and to make the convergence more robust, we realize this with a weighted averaging,

$$s_i^{(T)} = \frac{1}{\sum_{t=1}^T \alpha_i^{(t)}} \sum_{t=1}^T \alpha_i^{(t)} s^{(t)}, \quad (5.1)$$

where $\alpha_i^{(t)} \in \{0, 1\}$ determines whether i is the best matching unit at time t . The averaging can efficiently be realized incrementally without additional parameters.

We follow the same approach to adapt the motor coupling,

$$m_{ij}^{(T)} = \frac{1}{\sum_{t=1}^T \alpha_{ij}^{(t)}} \sum_{t=1}^T \alpha_{ij}^{(t)} \hat{m}^{(t)}. \quad (5.2)$$

Here, the averaging weight $\alpha_{ij}^{(t)} \in \{0, 1\}$ is chosen such that m_{ij} learns the average motor signals that lead to an increasing postsynaptic and a decreasing presynaptic activity. In that way, m_{ij} learns which motor signals contribute, on average, to a transition from the stimulus s_j to a stimulus s_i . The simplest realization of this rule is

$$\alpha_{ij}^{(t)} = \begin{cases} 1 & \text{if } \dot{x}_i > 0 \text{ and } \dot{x}_j < 0 \\ 0 & \text{else} \end{cases}. \quad (5.3)$$

5.1 Experiments. All experiments will consider the problem of controlling a limb with position $\mathbf{y} \in [-1, 1]^2$ in a two-dimensional plane or maze. In this experiment, the sensor representation is directly the 2D coordinate of this limb, that is, $\mathbf{s} = \mathbf{y}$ (see section 8 for an example where the sensor representation is based on range measurements). The motor representation is given by 20 units, $\hat{\mathbf{m}} \in [0, 1]^{20}$, which encode 20 different bearing directions $\varphi_i \in \{0^\circ, 18^\circ, \dots, 342^\circ\}$. Activations of motor units directly lead to a limb movement with velocity $\dot{\mathbf{y}}$ according to the law

$$\begin{pmatrix} \dot{y}_1 \\ \dot{y}_2 \end{pmatrix} = \sum_{i=1}^{20} \hat{m}_i \begin{pmatrix} \cos(\varphi_i) \\ \sin(\varphi_i) \end{pmatrix}. \quad (5.4)$$

At the borders or walls of a maze, this law is violated such that \dot{y}_1 or \dot{y}_2 is set to zero when otherwise the border or wall would be crossed.

In the first experiment, the limb performs random movements that are induced by explicitly coupling a random signal A_i into the motor layer (see equation 2.1). A random signal A_i is generated by randomly picking a motor unit i^* and choosing $A_{i^*} = 1$ while $A_i = 0$ for all $i \neq i^*$. The signal is not randomized at every time step; instead, at each time step with a probability .8, the signal remains unchanged, and with a probability .2, a new i^* is chosen.

These movements generate the data—the sequences of sensor and motor signals $\mathbf{m}^{(t)}$ and $\mathbf{s}^{(t)}$ —from which the sensorimotor map learns the dependencies between motor signals and stimulus changes. Our choice of parameters for the dynamics of the sensorimotor map and motor layer is shown in Table 1. Those for adaptation are $\tau_e = 10$, $\nu = .2$, and $a_{\max} = 300$. During the learning phase, the lateral coupling (which will induce anticipation) is switched off ($\eta = 0$).

Figure 2A displays the topology of the sensorimotor map that has been learned for the 2D plane after various time steps. In all displays, the units are positioned according to their sensor weight vectors s_i . Concerning the topology, we basically reproduce the standard behavior of growing neural gas: in the early phase, the map grows as more and more regions are

Table 1: Parameters.

	τ	h	w_E	σ_E	w_I	ρ	η	σ_S
Sensorimotor map	2	0	-	-	.5	.01	0	.05
Motor layer	5	-1	1	2	.6	.01	-	-

explored. In the late phase, unnecessary connections are deleted, leading to a Voronoi-like graph.

Figures 2B and 2C are two different illustrations of the learned motor weight vectors \mathbf{m}_{ij} . To compute these diagrams, we first associate an angle $\theta_{ij} = \angle(\mathbf{s}_j - \mathbf{s}_i)$ with every connection in the sensorimotor map. These angles θ_{ij} correspond to the true geometrical direction of a transition from j to i . Figure 2B displays tuning diagrams for 10 different motor units. For a given motor unit k , we consider all connections (ij) and draw a line with orientation θ_{ij} and length $(\mathbf{m}_{ij})_k$. The diagrams exhibit that motor units that represent a certain bearing φ_k have larger weights to connections with similar bearing θ_{ij} . The tuning curve 2C displays the same data in another way: for every motor unit k and connection (ij) , the weight $(\mathbf{m}_{ij})_k$ is plotted against the difference $\theta_{ij} - \varphi_k$.

Finally, Figure 2D displays the learning curve with respect to an error measure for the weight vectors \mathbf{m}_{ij} : as every motor unit k corresponds to a bearing φ_k , every activation pattern $\hat{\mathbf{m}}$ over motor units corresponds to an average bearing $\varphi(\hat{\mathbf{m}})$ (cf. equation 5.4). The weight vectors \mathbf{m}_{ij} are such activation patterns and thus correspond to average bearings $\varphi(\mathbf{m}_{ij})$. The error measure is the absolute difference between this bearing $\varphi(\mathbf{m}_{ij})$ and the geometrical direction θ_{ij} , averaged over all connections (ij) . The graph shows that this error measure does not fully converge to zero. Indeed, most of this error is accumulated at the border of the region for an obvious reason: according to the “physics” we defined, a motor command that would diagonally cross a border leads to a movement parallel to the border instead of a full stop. Thus, at the borders, a whole variety of motor commands exists that all lead to the same movement parallel to the border. Connections between two units parallel to a border thus learn an average of motor commands that also includes diagonal motor commands.

6 Anticipation

The sensorimotor map as introduced so far is sufficient for short-term anticipations. When the sensorimotor space is explored as previously with random movements and given the map as learned in the previous example,

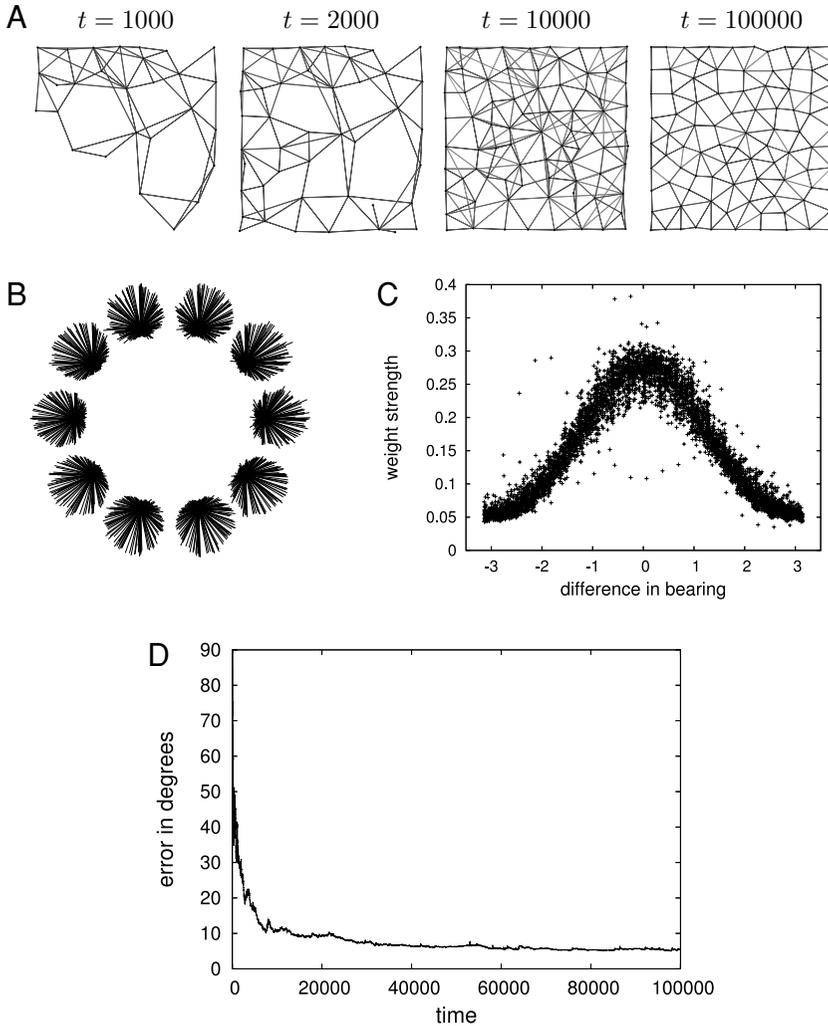


Figure 2: (A) The topology of the sensorimotor representation learned for the 2D region at different times. (B) Tuning diagrams for 10 of the 20 motor units (we display only every second unit to save space): for a motor unit k , lines with length $(m_{ij})_k$ and orientation θ_{ij} are drawn. (C) The tuning curve of motor units for all motor units and lateral connections: the weight $(m_{ij})_k$ is plotted against the difference of orientation of the motor unit (φ_k) and the connection (θ_{ij}). (D) The learning curve of an error measure for the difference in bearing represented by m_{ij} and ϕ_{ij} . Errors occur mostly at the borders. See section 5.1 for more details.

we may compare the actual current stimulus s to what the sensorimotor map currently represents,

$$\bar{s} = \frac{1}{\sum_i x_i} \sum_i x_i s_i. \quad (6.1)$$

We term the quantity \bar{s} the *represented stimulus*, which may in general differ from the true stimulus s ; we term the difference $\Delta\bar{s} = \bar{s} - s$ the *representational shift*. The approximate nature of the representation is one obvious source of representational shift: even when the lateral couplings are turned off ($\eta = 0$), there might be small shifts because the map is course-grained. In our case, most of such representation errors stem, again, from the borders. Since there exist no units to represent positions beyond a border and since the activations x_i typically have a gaussian-like shape over the units i , the represented stimulus \bar{s} for a stimulus s at the border will always have a slight inward shift of the order of σ . The results we give will omit this effect by discarding data from the border of the region.

We collected data for three different strength $\eta \in \{0, .2, .5\}$ of lateral interaction. The two measures we discuss are the norm,

$$RSN = |\Delta\bar{s}|, \quad (6.2)$$

of the representational shift and the directional match RSD of the representational shift with the true change in stimulus $\Delta s^{(t)} = s^{(t+1)} - s^{(t)}$ that occurs due to the motor activations,

$$RSD = \frac{\langle \Delta\bar{s}, \Delta s \rangle}{|\Delta\bar{s}| |\Delta s|} \in [-1, 1]. \quad (6.3)$$

The results are displayed in Figure 3. All numbers are the averages (and standard deviations) over 2205 data points taken when the limb moves, in random directions as described previously, in the central area of the plane.

For $\eta = 0$, we find that the norm of the representational shift ($RSN = .015 \pm .009$) is, as expected, very small when compared to the kernel width $\sigma = .05$. The shift direction is not correlated to the true stimulus change ($RSD = .0054 \pm .7$). Thus, for $\eta = 0$, the internally represented stimulus \bar{s} is fully dominated by the true stimulus s , and small representational shifts stem from the approximate nature of the representation.

For $\eta = .2$ we find significantly larger shifts, $RSN = .075 \pm .042$. More important, though, we find a strong correlation in the direction of the representational shift and the true future change of the stimulus, $RSD = .89 \pm .27$. For $\eta = .5$, both effects are even stronger: $RSN = .20 \pm .17$ and

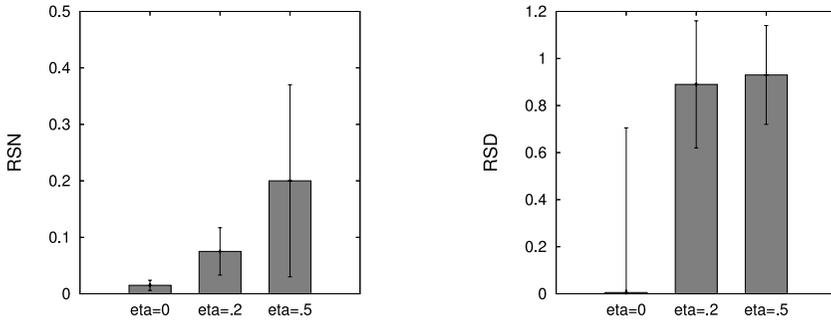


Figure 3: The norm RSN of the representational shift and the correlation measure RSD between representational shift and the current change of stimulus, for different strengths η of the lateral coupling in the sensorimotor map. With nonzero lateral coupling, the represented stimulus is shifted in the same direction as the current true stimulus change.

$RSD = .93 \pm .21$. For any η , the norm of the true stimulus change is $|\Delta s| = 0.036 \pm 0.017$.

The results clearly show that the representational shift $\Delta \bar{s}$ encodes an anticipation of the true change of stimulus, that is, the represented stimulus \bar{s} is an anticipation of a future stimulus that will be perceived depending on the current motor activations. The motor modulation of the lateral interactions is able to direct the representational shift toward the direction that corresponds to the motor signals.

This effect can be seen much better visually, watching the recordings¹ of the activations on the sensorimotor map and the dynamics of the two positions that correspond to \bar{s} and s (see also Figure 4). For $\eta = 0$, both \bar{s} and s move very coherently, almost always overlapping; only at the borders there is a systematic inward shift. For $\eta = .2$, the activity peak of the field x is always slightly ahead of the true stimulus; the represented position \bar{s} always runs ahead of the true limb position s . When the motor activations change, \bar{s} sweeps in front of s toward the new movement bearing.

For $\eta = .5$ the situation becomes more dramatic. The lateral interaction become dominating such that the field activations x actually run away from the true stimulus, traveling self-sustained in the direction of the current movement. This “wave” breaks down at the border of the sensorimotor map, and the activation peak is recreated at the current stimulus. Thus, the represented position \bar{s} travels quickly away from the true limb position s in the movement direction until it hits the border and restarts from s .

¹Access and watch the recordings online at www.marc-toussaint.net/projects.

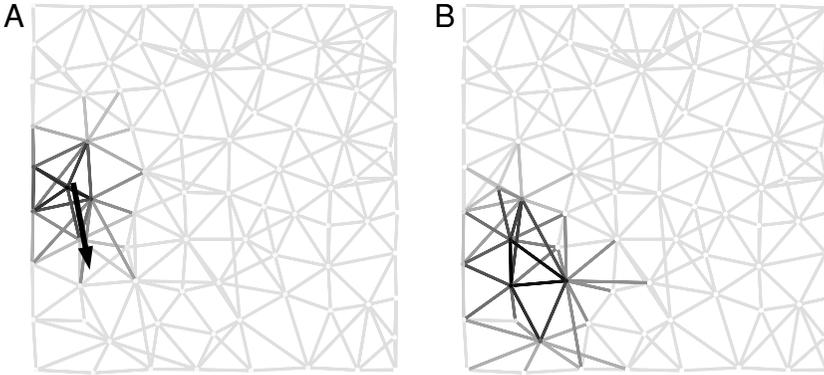


Figure 4: Anticipation of future stimuli. (A) The forward excitation S_i , which encodes the true current stimulus s . The gray shading indicates the value of $S_i \in [0, 1]$; for better visibility, edges (ij) are shaded with the average value $(S_i + S_j)/2$. The black arrow indicates the direction encoded by the current motor activations. (B) The activation field x_i on the sensorimotor map. It exhibits a significant shift in the direction of movement, thus encoding an anticipation of future stimuli depending on the current motor activations. See also note 1.

7 The Dynamics of Planning

To organize goal-oriented behavior, we assume that, in parallel to the activation dynamics of x , there exists a second dynamic process that can be motivated from classical approaches to reinforcement learning (Bertsekas & Tsitsiklis, 1996; Sutton & Barto, 1998). Recall the Bellman equation,

$$V_{\pi}^*(i) = \sum_a \pi(a|i) \sum_j P(j|i, a)[r(j) + \gamma V_{\pi}^*(j)], \quad (7.1)$$

yielded by the expectation $V^*(i)$ of the discounted future return $R(t) = \sum_{\tau=1}^{\infty} \gamma^{\tau-1} r(t+\tau)$ (for which $R(t) = r(t+1) + \gamma R(t+1)$). Here, i is a state index, and γ is the discount factor. We presumed that the received rewards $r(t)$ actually depend on only the state and thus enter equation 7.1 only in terms of the reward function $r(i)$ (we neglect here that rewards may directly depend on the action). Behavior is described by a stochastic policy $\pi(a|i)$, the probability of executing an action a in state i . Given the property 7.1 of V^* , it is straightforward to define a recursion algorithm for an approximation V

of V^* such that V converges to V^* . This recursion algorithm is called *value iteration* (Sutton & Barto, 1998) and reads

$$\tau_v \Delta V_\pi(i) = -V_\pi(i) + \sum_a \pi(a|i) \sum_j P(j|i, a) [r(j) + \gamma V_\pi(j)], \quad (7.2)$$

with a “reciprocal learning rate” or time constant τ_v . Note that equation 7.1 is the fixed point equation of equation 7.2.

Equation 7.2 provides an iterative scheme to compute the state-value function V based on only local information. The practical meaning of the state-value function is that it quantifies how desirable and promising it is to reach a state i , also accounting for future rewards to be expected. If rewards are given only at a single goal state, V has its maximum at this goal and is the higher the easier the goal can be reached from a given state. Thus, if the current state is, i it is a simple and efficient rule of behavior to choose an action a that will lead to the neighbor state j with maximal $V(j)$ (the greedy policy). In that sense, $V(i)$ provides a smooth gradient toward desirable goals. Note, though, that direct value iteration presumes that the state and action spaces are known and finite and that the current state and the world model $P(j|i, a)$ are known.

In transferring these classical ideas to our model, we assume that the system is given a goal stimulus g , that is, it is given the command to reach a state that corresponds to perceiving the stimulus g . Just as ordinary stimuli induce an input S_i to the field activations x_i , we let the goal stimulus induce a reward excitation,

$$R_i = \frac{1}{Z} \exp \frac{-(s_i - g)^2}{2\sigma_R^2}, \quad (7.3)$$

for each unit i , where Z is chosen such that $\sum_i R_i = 1$. Besides the activations x_i , we introduce an additional field over the sensorimotor map, the value field v_i , which is in analogy to the state-value function $V(i)$. The dynamics are

$$\tau_v \dot{v}_i = -v_i + R_i + \gamma \max_j (w_{ji} v_j), \quad (7.4)$$

and well comparable to equation 7.2. One difference is that v_i estimates the “current-plus-future” reward $r(t) + \gamma R(t)$ rather than the future reward only. In the upper notation, this corresponds to the value iteration $\tau_v \Delta V_\pi(i) = -V_\pi(i) + r(i) + \sum_a \pi(a|i) \sum_j P(j|i, a) [\gamma V_\pi(j)]$. As it is commonly done for value iteration, we assumed π to be the greedy policy. More precisely, we considered only that action (i.e., that connection (ji)) that leads to the neighbor state j with maximal value $w_{ji} v_j$. In effect, the summations over a as well as over j can be replaced by a maximization over j .

Finally, we replaced the probability factor $P(j|i, a)$ by w_{ji} . In practice, the value field will relax quickly to its fixed point $v_i^* = R_i + \gamma \max_j (w_{ji} v_j^*)$ and stay there if the goal does not change.

The quasi-stationary value field v_i together with the current (nonstationary) activations x_i allow the system to generate motor excitations that lead toward the goal. More precisely, the gradient $v_j - v_i$ in the value field indicates how desirable motor activations m_{ji} are when the current "state" is i . Goal-directed motor excitations can thus be generated as a weighted average of the motor activations m_{ji} that have been learned for the connections,

$$A = \frac{1}{Z} \sum_{i,j} x_i w_{ji} (v_j - v_i) m_{ji}, \quad (7.5)$$

where Z is chosen to normalize $|A| = 1$. These excitations enter the motor activation dynamics, equation 2.1. Hence, the signals flow between the sensorimotor map, and the motor system is in both directions. In the anticipation process, the signals flow from the motor layer to the sensorimotor map: motor signals activate the corresponding connections and cause lateral, predictive excitations. In the action selection process, the signals are emitted from the sensorimotor map back to the motor layer to induce the motor excitations that should turn predictions into reality.

7.1 Experiments. To demonstrate the planning capabilities of the sensorimotor map, we consider a 2D maze. In the first phase, a sensorimotor map is learned that represents the specific maze environment, using random explorations as described in section 5. Figure 5A displays the topology of the learned sensorimotor map after 100,000 iterations, now with a kernel $\sigma_S = .01$.

In the planning phase, a goal stimulus is applied that corresponds to the position indicated by a triangle in Figure 5B. This goal stimulus induces reward excitations R_i on units that match the goal stimulus closely. The value field dynamics, equation 7.4, quickly relaxes to its fixed point, which is displayed in Figure 5C. The parameters we used are $\tau_u = 5$, $\gamma = .9$, and $\sigma_R = \sigma_S/4$. As expected, the value field activations are high for units representing the proximity of the goal location and decay smoothly along the connectivity of the sensorimotor map. Note that this value field is not a decaying function of the Euclidean distance to the goal, but approximately a decaying function of the topological distance to the goal, that is, the shortest path length with respect to the learned topology.

Figure 5B illustrates a trial where the limb is initially located in the upper-right corner of the maze. The activation field x_i represents this current location. Together with the gradient of the value field at the current location (see equation 7.5), motor excitations are induced that let the limb

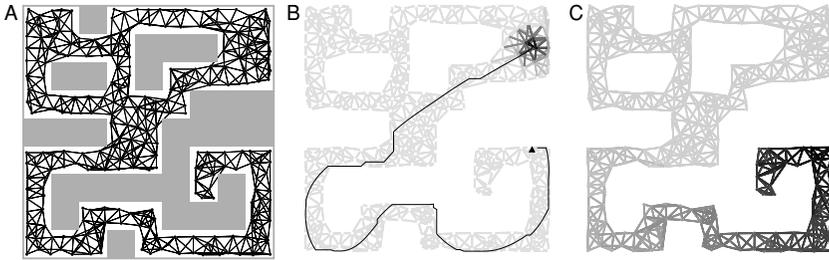


Figure 5: Experiments with a maze. (A) The topology of the sensorimotor map learned. (B) The activation field x_i on the sensorimotor at the start of the trajectory. (C) The attractor state of the value field on the sensorimotor map, spreading from the goal location in the lower right.

move toward the bottom left. As the limb moves, the sensorimotor activities x_i follow its current position, and new motor excitations are induced continuously, which leads to a movement trajectory ascending the gradient of the value field until the goal is reached. In the experiment, once the goal is reached, we switch the goal to a new random location, inducing new reward excitations R_i . The value dynamics, equation 7.4, respond quickly to this change and relax to a new fixed point, providing the gradient to the new goal.

A standard quality measure for planning techniques is the required time to goal. Figure 6 displays the time intervals between switching the goals, which are the times needed to reach the new goal position from the previous goal position. First, we see that the goal is always reached in finite time, indicating that planning is always successful. Further, the graph compares the time to reach the goal with the length of the shortest path. This shortest path length was computed from a coarse (block-wise) discretization of the maze with dynamic programming. The clear correlation between the time to reach the goal and the shortest path length shows that the movement of the limb indeed follows a planned shortest path trajectory from the initial position to the goal.

Another indicator for successful action selection is whether the current movement is in the direction of the value gradient. Figure 7A displays the bearing of the local value gradient, $\angle(\sum_i x_i (v_j - v_i) (s_j - s_i))$, and the bearing of the current movement, $\angle(\dot{\mathbf{j}})$, for the first 300 time steps of the experiment. We observe a clear correlation between both bearings, though with a systematic time delay. This time delay is approximately six time steps, as can be read from Figure 7B, and corresponds to the timescale of the motor dynamics, $\tau_m = 5$. (See note 1 to access more recordings of planning experiments).

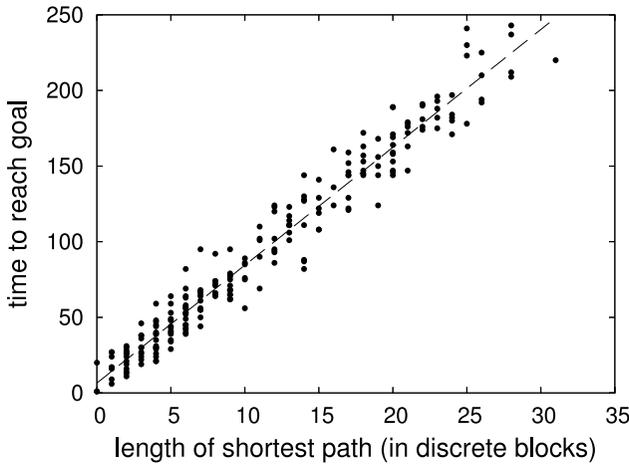


Figure 6: Times needed to move from a random start position to a random goal position in the maze when the sensorimotor map plans and controls the movement. The time is plotted against the length of the shortest path connection, which was computed from a coarse (block-wise) discretization of the maze with dynamic programming.

8 Extensions

We tried to keep the model introduced so far simple and focused on the key mechanisms. This basic model can be extended in many straightforward ways to realize other desired functionalities.

For instance, the path generated by the sensorimotor map in the previous example (see Figure 5B) clearly hits the walls very often. This should be no surprise since there is no mechanism of obstacle avoidance implicit in the model so far. But is easy to apply a standard obstacle avoidance technique: given distance signals $d_i \in [0, 1]$ from 20 range sensors (in the 20 different bearings φ_i) around the limb, an inhibition (e.g., proportional to $(1 - d_i)^3$) can be directly coupled into the motor activations m_i . Figure 8A displays a trajectory generated with this obstacle avoidance switched on.

Perhaps more important is the question of whether the local lateral weights w_{ij} should be learned instead of fixed to 1 if a connection exists and 0 otherwise. In Toussaint (2004) we presented a learning scheme for these weights based on the temporal Hebb rule. One of the main reasons to consider the continuous plasticity of these lateral weights was that this allows the model to adapt to change. We decided to keep the simpler alternative where $w_{ij} \in \{0, 1\}$. The adaptability can also be achieved with the basic mechanism of adapting the topology to a change of the world, that is, keep adding or deleting connections.

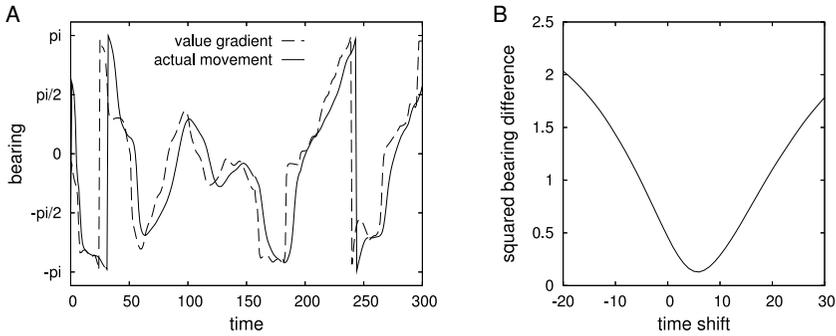


Figure 7: (A) The movement direction $\angle(\mathbf{j})$ and the direction of the value gradient $\angle(\sum_i x_i (v_j - v_i) (s_j - s_i))$ for the first 300 time steps. (B) Similar to a convolution between both curves, we plot the average squared difference $\sum_i \|h(t) - f(t + \tau)\|^2$ between both curves when one of them is shifted by a time delay τ . (We chose a norm $\|\cdot\|^2$ that accounts for the cyclic metric in angle space.) The typical time shift is $\tau^* = 6$.

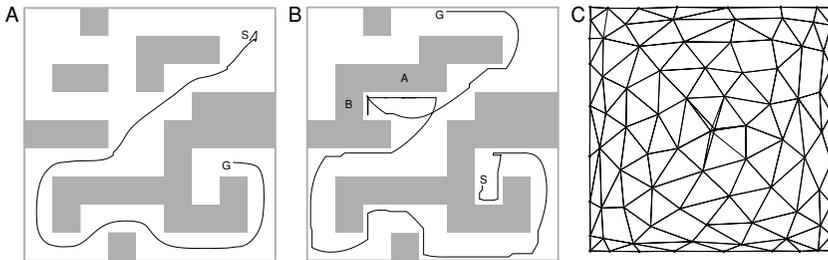


Figure 8: Results of three different extensions of the sensorimotor map. (A) A trajectory with obstacle avoidance. (B) A trajectory from start S to goal G when the paths were blocked at A and B. (C) A sensorimotor map learned from range sensors. See section 8 for details.

Recall our rule to delete connections. As for growing neural gas (Fritzke, 1995), we associate an “age” a_{ij} with every connection and delete the connections when it exceeds a threshold a_{max} . The difference from Fritzke is that we increase all connections’ ages by an amount of $M_{ij} \phi(x_j)$ at every time. As a result, if during execution of a planned trajectory, an anticipated transition to a new stimulus does not occur, then the connections that contribute to this anticipation (for which $M_{ij} \phi(x_j)$ is high) will eventually be deleted. This adaptation of the topology has a crucial influence on the dynamics of the value field. If all connections of a specific pathway are deleted, the attractor of the value field rearranges to guide a way around this blocked pathway.

Figure 8B displays such an interesting trajectory, generated for $a_{\max} = 10$: the limb is initially located at S and the goal location is G. The system has learned a representation of the original maze, as given in Figure 5A. But the maze has now been modified by blocking the pathways at A and B. The system first tries to follow a direct path, which is blocked at A. When the limb hits this barrier and continuously activates the connections crossing A (in terms of $M_{ij} \phi(x_j)$), they are eventually deleted, which makes the value field rearrange and guide along another path crossing B. Now the limb hits the barrier at B and connections are deleted, which finally leads to a path that allows the limb to reach the goal. (See note 1 to access recordings of these experiments.)

Finally, since the stimulus kernel (see equation 4.2) was chosen as a gaussian, the stimulus can be also given in representations other than directly as the location y of the limb. For instance, Figure 8C displays the sensorimotor map learned for the plane when the input was given as a 40-dimensional range vector (d_1, \dots, d_{40}) , each $d_i \in [0, 1]$ for 40 different bearings. The only difference with the setup in section 5 was the choice of the kernel width: now we used $\sigma_S = 1$. The learned topology is slightly more dense close to the borders. This stems from the fact that the range vector changes more dramatically close to a wall since the visual angle under which the wall is seen (and thus the number of entries of the range vector affected by the wall) varies more. Anticipation and planning equally work for this stimulus representation. However, the model is not sufficient to handle ambiguous (partially observable) stimuli. For example, in the maze, there exist many positions with very similar range sensor profile. The sensorimotor map learned of the maze with only range sensor data would lead to an incorrect topology (cf. section 10).

9 Discussion

A key mechanism of the sensorimotor map is that motor activations modulate the lateral connection strengths and thereby induce anticipatory shifts of the activity peak on the sensorimotor map. This modulatory sensorimotor coupling encodes a model of the change of stimuli depending on the current motor activities. The mechanism attributes a specific role to the lateral connectivity, namely, motor-modulated anticipatory excitation, which differs significantly from previously proposed roles for lateral connections. However, we believe that the different views on the roles of lateral connections do not compete but complement each other; lateral connections may play different roles depending on the context and function of the respective neural representation. For instance, the role of lateral connections has been extensively discussed in the context of pure sensor representations, in particular, for the visual cortex (Miikkulainen, Bednar, Choe, & Sirosh, 2005). In such sensor representations, the function of lateral connections could be subsumed as either enforcing coherence or competition between laterally

connected units. The formation of topographic maps, columnar structure, or patterns of orientation-selective receptive fields can be explained on this basis (e.g., Bednar et al., 2002). Also the stabilization of noisy or occluded stimuli, or the disambiguation between contradicting stimuli can be modeled—for example, with standard neural field dynamics involving local excitatory and global inhibitory lateral connections (Erlhagen & Schöner, 2002). In the context of temporal signal representations, the function of lateral connections is to induce specific temporal dynamics, learned, for example, with a temporal Hebb rule (spike-time dependent plasticity; Dayan & Abbott, 2001). Self-organized temporal map models (Euliano & Principe, 1999; Somervuo, 1999; Wiemer, 2003; Varsta, 2002; Klemm & Alstrom, 2002) can learn to anticipate stimuli, for example, when a stimulus **B** always follows a stimulus **A**. The role we attributed to lateral connections naturally differs from these models since we consider a sensorimotor representation where anticipation needs to depend on the current motor activities and for which we proposed the modulatory sensorimotor coupling.

Long-term prediction, for example, path integration (see Etienne & Jeffery, 2004, for a review), is clearly related to the sensorimotor anticipation that we addressed with our model. Some existing models of path integration are based on one- or two-dimensional representational maps of position or head direction, and anticipation is realized by a motor-dependent translation of the activity pattern. For instance, in Hartmann and Wehner (1995) and Song and Wang (2005), the translational shift on a one-dimensional head direction representation is realized with two additional layers of inhibitory neurons—one for right and one for left movements—that are coupled to the motor system. Zhang (1996) achieves an exact translation of the activity pattern by coupling a derivative term in the dynamics, while Stringer, Rolls, Trappenberg, and Araujo (2002) induce translational shifts on a two-dimensional place field representation with a complex coupling of head direction units and forward velocity units into the lateral place field dynamics, in effect similar to our approach. None of these approaches addresses the problem of planning or the emission of motor signals based on the learned forward model. Although our model implements sensorimotor anticipation, it is in its current form limited with regard to the task of exact path integration: only the direction but not the magnitude of anticipatory shifts is guaranteed to be correlated with the true movement, as the experiments in section 6 demonstrate. However, future extensions of the model might solve this problem, for example, by a precise tuning of the parameter η that allows us to calibrate the magnitude of the anticipatory shift (see Figure 3).

In the context of machine learning, predictive forward models are typically learned as a function, for example, with a neural network (Jordan & Rumelhart, 1992; Wolpert, Ghahramani, & Jordan, 1995; Ghahramani, Wolpert, & Jordan, 1997; Wolpert, Ghahramani, & Flanagan, 2001). It is assumed that a goal trajectory is readily available such that the learned

function allows them to compute the motor signals necessary to follow this trajectory. A representational map of the state space is not formed. In contrast, some model-based reinforcement learning systems have addressed the self-organization of state space representations (Kröse & Eecen, 1994; Zimmer, 1996; Appl, 2000), for example, by using discrete fuzzy representations (e.g., the Fuzzy-ARTMAPs; Carpenter et al., 1992). However, these approaches do not propose a direct coupling of motor activities into a sensorimotor representation to realize anticipation and planning by neural dynamics on this representation; instead, they use the learned representation as an input to separate, standard reinforcement learning architectures.

10 Conclusion

The sensorimotor map we describe in this letter proposes a new mechanism of how motor signals can jointly be coupled with sensor signals on a sensorimotor representation. The immediate function of this sensorimotor map and the proposed modulatory sensorimotor coupling is the anticipation of the change of stimuli depending on the current motor activity.

Anticipation on its own is a fundamental ingredient of sensorimotor control, for example, to consolidate noisy sensorial information by fusing it with the temporal prediction or to bridge the inherent time lag of sensorial information. However, the ability to anticipate also provides the basic ingredient for planning. The forward model implicitly encoded by the sensorimotor map can be used to realize planning. We considered standard reinforcement learning techniques as a starting point and proposed a value dynamics on the sensorimotor map that performs basically the same computations as value iteration. For this to work in a neural systems framework, it is crucial that there exists a neural representation of the state space. The sensorimotor map provides such a representation.

The self-organization and learning processes that develop the sensorimotor map do not set principled constraints on the type of sensor and motor representations coupled to the map. However, a more general problem was not solved and remains a limiting factor. Also in our model, the self-organization of the sensorimotor representation is mainly sensor driven. This leads to problems, as when different states induce the same stimulus (partial observability, stimulus ambiguity), since the current self-organization process will not be able to grow separate units for the same stimulus. The self-sustaining and anticipatory dynamics of the sensorimotor map are able to disambiguate such states depending on the temporal context. But a prerequisite is the existence of multiple units associated with the same stimulus.

This leads us back to the challenge of understanding higher-level cognitive processes like internal simulation and planning, as mentioned in the context of Köhler's classic monkey experiments. The basic mechanisms of anticipation and planning proposed in this letter, in particular, the

action-dependent modulation of lateral interactions, might be transferable to such higher-level representations. An open question is how animals and humans are able to organize such higher-level abstract representations, which clearly are not purely sensor based but state abstractions that capture both the sensor context and its relevance for behavior.

Acknowledgments

I thank the German Research Foundation (DFG) for funding the Emmy Noether fellowship TO 409/1-1, allowing me to pursue this research.

References

- Abbott, L. (1991). Realistic synaptic inputs for network models. *Network: Computation in Neural Systems*, 2, 245–258.
- Amari, S. (1977). Dynamics of patterns formation in lateral-inhibition type neural fields. *Biological Cybernetics*, 27, 77–87.
- Appl, M. (2000). *Model-based reinforcement learning in continuous environments*. Unpublished doctoral dissertation, Institut für Informatik, Technische Universität München.
- Bagchi, S., Biswas, G., & Kawamura, K. (2000). Task planning under uncertainty using a spreading activation network. *IEEE Transactions on Systems, Man and Cybernetics A*, 30, 639–650.
- Bednar, J. A., Kelkar, A., & Miikkulainen, R. (2002). Modeling large cortical networks with growing self-organizing maps. *Neurocomputing*, 44–46, 315–321.
- Bergener, T., Bruckhoff, C., Dahm, P., Janßen, H., Joublin, F., Menzner, R., Steinhage, A., & von Seelen, W. (1999). Complex behavior by means of dynamical systems for an anthropomorphic robot. *Neural Networks*, 12, 1087–1099.
- Bertsekas, D., & Tsitsiklis, J. (1996). *Neuro-dynamic programming*. Nashua, NH: Athena Scientific.
- Bishop, C., Hinton, G., & Strachan, I. (1997). GTM through time. In *Proc. of IEE Fifth Int. Conf. on Artificial Neural Networks* (pp. 111–116). London: IEE.
- Carpenter, G., Grossberg, S., Markuzon, N., Reynolds, J., & Rosen, D. (1992). Fuzzy ARTMAP: A neural network architecture for incremental supervised learning of analog multidimensional maps. *IEEE Transactions on Neural Networks*, 5, 698–713.
- Dahm, P., Bruckhoff, C., & Joublin, F. (1998). A neural field approach to robot motion control. In *Proc. of 1998 IEEE Int. Conf. on Systems, Man, and Cybernetics (SMC 1998)* (pp. 3460–3465). New York: IEEE.
- Dayan, P., & Abbott, L. (2001). *Theoretical neuroscience*. Cambridge, MA: MIT Press.
- Erlhagen, W., & Schöner, G. (2002). Dynamic field theory of movement preparation. *Psychological Review*, 109, 545–572.
- Etienne, A. S., & Jeffery, K. J. (2004). Path integration in mammals. *Hippocampus*, 14, 180–192.
- Euliano, N., & Principe, J. (1999). A spatio-temporal memory based on SOMs with activity diffusion. In *Workshop on Self-Organizing Maps* (pp. 253–266). Helsinki.

- Fritzke, B. (1995). A growing neural gas network learns topologies. In G. Tesauero, D. Touretzky, & T. Leen (Eds.), *Advances in neural information processing systems*, 7 (pp. 625–632). Cambridge, MA: MIT Press.
- Gersho, A., & Gray, R. (1991). *Vector quantization and signal compression*. Boston: Kluwer.
- Ghahramani, Z., Wolpert, D. M., & Jordan, M. I. (1997). Computational models of sensorimotor integration. In P. Morasso & V. Sanguineti (Eds.), *Self-organization, computational maps and motor control* (pp. 117–147). Dordrecht: Elsevier.
- Grush, R. (2004). The emulation theory of representation: Motor control, imagery, and perception. *Behavioral and Brain Sciences*, 27, 377–396.
- Hartmann, G., & Wehner, R. (1995). The ant's path integration system: A neural architecture. *Biological Cybernetics*, 73, 483–497.
- Hesslow, G. (2002). Conscious thought as simulation of behaviour and perception. *Trends in Cognitive Sciences*, 6, 242–247.
- Iossifidis, I. & Steinhage, A. (2001). Controlling an 8 DOF manipulator by means of neural fields. In A. Halme, R. Chatila, & E. Prassler (Eds.), *Int. Conf. on Field and Service Robotics (FSR 2001)*. Helsinki.
- Jordan, M., Rumelhart, D. (1992). Forward models: Supervised learning with a distal teacher. *Cognitive Science*, 16, 307–354.
- Klemm, K., & Alstrom, P. (2002). Emergence of memory. *Europhysics Letters*, 59, 662.
- Köhler, W. (1917). *Intelligenzprüfungen am menschenaffen*. (3rd ed.). Berlin: Springer.
- Kohonen, T. (1995). *Self-organizing maps*. Berlin: Springer.
- Kröse, B., & Eecen, M. (1994). A self-organizing representation of sensor space for mobile robot navigation. In *Proc. of Int. Conf. on Intelligent Robots and Systems (IROS 1994)* (pp. 9–14). New York: IEEE.
- Majors, M., & Richards, R. (1997). *Comparing model-free and model-based reinforcement learning*. (Tech. Rep. No. CUED/F-INENG/TR.286). Cambridge: Cambridge University Engineering Department.
- Mel, B. W., (1990). *The sigma-pi column: A model of associative learning in cerebral cortex*. (Tech. Rep. CNS Memo 6). Pasadena: Computation and Neural Systems Program, California Institute of Technology.
- Mel, B. W., & Koch, C. (1990). Sigma-pi learning: On radial basis functions and cortical associative learning. In D. S. Touretzky (Ed.), *Advances in neural information processing systems*, 2 (pp. 474–481). San Mateo, CA: Morgan Kaufmann.
- Miikkulainen, R., Bednar, J. A., Choe, Y., & Sirosh, J. (2005). *Computational maps in the visual cortex*. Berlin: Springer.
- Phillips, W., & Singer, W. (1997). In the search of common foundations for cortical computation. *Behavioral and Brain Sciences*, 20, 657–722.
- Rose, P., & Scott, S. (2003). Sensory-motor control: A long-awaited behavioral correlate of presynaptic inhibition. *Nature Neuroscience*, 12, 1309–1316.
- Schöner, G., & Dose, M. (1992). A dynamical system approach to task level system integration used to plan and control autonomous vehicle motion. *Robotics and Autonomous Systems*, 10, 253–267.
- Schöner, G., Dose, M., & Engels, C. (1995). Dynamics of behavior: Theory and applications for autonomous robot architectures. *Robotics and Autonomous Systems*, 16, 213–245.

- Somervuo, P. (1999). Time topology for the self-organizing map. In *Proc. of Int. Joint Conf. on Neural Networks (IJCNN 1999)* (pp. 1900–1905). New York: IEEE.
- Song, P., & Wang, X.-J. (2005). Angular path integration by moving “hill of activity”: A spiking neuron model without recurrent excitation of the head-direction system. *J. Neuroscience*, *25*, 1002–1014.
- Stringer, S. M., Rolls, E. T., Trappenberg, T. P., & Araujo, I. E. T. de. (2002). Self-organizing continuous attractor networks and path integration: Two-dimensional models of place cells. *Network*, *13*, 429–446.
- Sutton, R., & Barto, A. (1998). *Reinforcement learning*. Cambridge, MA: MIT Press.
- Toussaint, M. (2004). Learning a world model and planning with a self-organizing, dynamic neural system. In S. Thrun, L. K. Saul, & B. Schölkopf (Eds.), *Advances in neural information processing systems*, *16* (NIPS 2003) (pp. 929–936). Cambridge, MA: MIT Press.
- Varsta, M. (2002). *Self-organizing maps in sequence processing*. Unpublished doctoral dissertation, Helsinki University of Technology.
- von der Malsburg, C. (1973). Self-organization of orientation-sensitive cells in the striate cortex. *Kybernetik*, *15*, 85–100.
- Wiemer, J. (2003). The time-organized map algorithm: Extending the self-organizing map to spatiotemporal signals. *Neural Computation*, *15*, 1143–1171.
- Willshaw, D. J., & von der Malsburg, C. (1976). How patterned neural connections can be set up by selforganization. *Proceedings of the Royal Society of London*, *B194*, 431–445.
- Wolpert, D., Ghahramani, Z., & Flanagan, J. (2001). Perspectives and problems in motor learning. *Trends in Cognitive Science*, *5*, 487–494.
- Wolpert, D., Ghahramani, Z., & Jordan, M. (1995). An internal model for sensorimotor integration. *Science*, *269*, 1880–1882.
- Zhang, K. (1996). Representation of spatial orientation by the intrinsic dynamics of the head-direction cell ensemble: A theory. *J. Neuroscience*, *16*, 2112–2126.
- Zimmer, U. (1996). Robust world-modelling and navigation in a real world. *Neuro-Computing*, *13*, 247–260.