International Journal of Neural Systems, Vol. 11, No. 2 (2001) 125–143

OCKHAM'S RAZOR MODELING OF THE MATRISOME CHANNELS OF THE BASAL GANGLIA THALAMOCORTICAL LOOPS

ANDRÁS LŐRINCZ*,‡,§ and GYÖRGY HÉVÍZI*,†
*Department of Information Systems, Eötvös Loránd University,
Pázmány Péter sétány 1/D, Budapest, Hungary, H-1117
†Department of Theoretical Physics, University of Szeged,
Tisza Lajos körút 84-86, Szeged, Hungary, H-6720
‡E-mail: lorincz@valerie.inf.elte.hu

CSABA SZEPESVÁRI
Mindmaker Ltd., Konkoly-Thege 29-33, Budapest, Hungary, H-1121

Received 8 September 2000
Revised 8 January 2001
Accepted 8 January 2001

A functional model of the basal ganglia-thalamocortical (BTC) loops is described. In our modeling effort, we try to minimize the complexity of our starting hypotheses. For that reason, we call this type of modeling Ockham's razor modeling. We have the additional constraint that the starting assumptions should not contradict experimental findings about the brain. First assumption: The brain lacks direct representation of paths but represents directions (called speed fields in control theory). Then control should be concerned with speed-field tracking (SFT). Second assumption: Control signals are delivered upon differencing in competing parallel channels of the BTC loops. This is modeled by extending SFT with differencing that gives rise to the robust Static and Dynamic State (SDS) feedback-controlling scheme. Third assumption: Control signals are expressed in terms of a gelatinous medium surrounding the limbs. This is modeled by expressing parameters of motion in parameters of the external space. We show that corollaries of the model fit properties of the BTC loops. The SDS provides proper identification of motion related neuronal groups of the putamen. Local minima arise during the controlling process that works in external space. The model explains the presence of parallel channels as the means to avoiding such local minima. Stability conditions of the SDS predict that the initial phase of learning is mostly concerned with selection of sign for the inverse dynamics. The model provides a scalable controller. State description in external space instead of configurational space reduces the dimensionality problem. Falsifying experiment is suggested. Computer experiments demonstrate the feasibility of the approach. We argue that the resulting scheme has a straightforward connectionist representation exhibiting population coding and Hebbian learning properties.

1. Introduction

The curse of dimensionality problem[6] concerns the discretization in many dimensional spaces. This problem is apparent when the control of the human body is considered with its cca. 600 units to be controlled. If to each unit we limit the discretization to two (plus and minus) regions then the number of states is 2^{600} — too large for explorative learning. Moreover, control of the dynamical system may require the discretization of the phase space, i.e., the tangent bundle of the configurational space to form a position-and-direction-to-action (PDA) map.

§Author to whom correspondence should be addressed.

How could a PDA map satisfy the discretization problem? We note that direct representation of individual paths, instead of PDA maps, is even more demanding. However, it seems that the brain does not formulate signals to specify individual paths, but, instead, it formulates signals that specify the position of targets in the extrapersonal space. Differencing between the position of the target and the present position gives rise to direction. Neuronal signals of this kind can be found in the parietal association cortex.[73]

Interestingly, measurements on the basal ganglia (BG) indicate that planning and control is made in external space, in somatotopic coordinates. There is a gelatinous medium surrounding the limbs[24–26] that can be interpreted as the input to a distributed PDA map. This map may be modulated by the configuration of the limbs. Modulation may be accomplished by competition between matrisomes[29,17] controlled by medium spiny neurons of the putamen. Then the output of the BG should be modulated by the configuration. This suggestion meets experimental findings: It has been shown that the output of the basal ganglia is organized into multiple[22] barely-interacting channels.[50] We take this viewpoint and consider configurational modulation instead of direct discretization of the configurational space. Configurational modulation of the configurational space seems attractive owing to the relaxed need of computational units: The scheme that utilizes

- high resolution in the external space for precise positioning, and
- configurational modulation with low resolution for a crude model of the inverse dynamics separates the problems of high dimension and precision.

Such scheme may allow to use a crude, coarse coded representation of the configuration, such as the one found by Graziano *et al.*[24–26] The question arises, however, whether PDA maps can give rise to stable control or not. Alas, how could configurational modulation, the PDA map, be created?

PDA maps can be "translated"[18,44,63] into speed-field tracking[33] control problems. Speed-field tracking (SFT) may be formulated as a special case of trajectory tracking provided that the distinguished autonomous system is the non-linear plant itself controlled by an optimally designed state feedback control law. It then follows that the major difference between trajectory tracking and speed-field tracking is that the former is given as a function of time, whereas the latter is given as a function of state. Consequently, speed-field tracking is more robust against state perturbations. Moreover, speed-field tracking arises naturally in stationary optimal-control problems such as path planning tasks.[33] Speed-field tracking task has another advantage: the designer can incorporate several objectives into the form of speed-field to be tracked, and hence, extend the model's range of possibilities.

As we shall see, the working of the basal ganglia-thalamocortical (BTC) loops suggests that a differencing scheme is present in the BG. Thus, speed-field tracking and differencing can be combined in the model. It has been shown that this combination gives rise to robust control with ultimately uniform bounded performance for plants of every order.[61] In previous works, such combination of PDA maps and differencing schemes has been proposed as the model of the basal ganglia-thalamocortical loops.[43–45] Computational experiments were conducted in configurational space.

In this paper, we address the question how to create low-dimensional PDA map using information about the limb in the external space instead of full description in the configurational space. It will be argued that for configurational modulation a simple low-dimensional representation may satisfy the constraints of speed-field tracking. The novelty of our computational scheme is in the low-dimensional external space description that can be viewed as a gelatinous medium surrounding the limb. Computer simulations on multi-segment robotic arm will be used to demonstrate the idea. This solution to the control problem can be scaled. We shall argue that configurational modulation required for the low dimensional description may correspond to the clustering property of motor experience data.[39,38]

Various functional and computational models have been proposed that emphasize different aspects of the basal ganglia-thalamocortical interaction, such as the thalamic disinhibition (TDI) model[12] the electrotonic coupling model[8,9] and the winner-lose-all model.[5] A comprehensive review of experimental findings and models of various aspects of the basal ganglia (BG) can be found in a recent book.[32] The closest to our model is the TDI model. The

TDI model suggests that increasing (decreasing) the BG output shifts the motion towards hypokinetic (hyperkinetic) states. Although the model is capable of explaining several features of the BTC loops, it also predicts that lesion to the BG output results in dyskinesia, which is not the case. It has been argued that this problem may be reconciled on the basis of the voltage-dependent bistable properties of thalamic neurons[60] that give rise to a different mode for thalamic neurons if their GABAergic input is strongly reduced.[1] Comparison of the SDS model and other models of the basal ganglia can be found in Ref. 44.

There is a principled approach in our modeling efforts. There are two typical routes in brain modeling. The first of these approaches is called functional modeling. Taken to the extremes, functional modeling is similar to associations. A functional model makes use of a concept set borrowed from a different field. One example is the psycho-hydraulic model of Lorenz[42] that explains motivation of animals. We call this type of modeling top-down (TD) modeling. TD modeling has some danger. The trap is in the neglect of mapping the model onto the neuronal substrate. The opposite end of modeling is within the realm of computational neuroscience (CNS). CNS would like to avoid the above trap and deals with measured parameters of neurons, axons, synapses, and their interactions. CNS restricts itself to mathematical derivations from such quantities. We call CNS modeling bottom-up (BU) modeling. However, BU modeling is also subject to some danger. The trap emerges from the uncertainty of the measured values of parameters and thus the allowed range of parameter fitting. BU models can be highly complex and may have strong non-linearities. Non-linear dynamical system with restricted parameter range can still exhibit a broad range of behaviors and the generality of the results may be questioned in certain cases. We note that if the basic parameters of CNS models are compressed into simplifying assumptions then the borderline between TD and BU modeling vanishes. In this paper, we shall emphasize the importance of restrictions in modeling. We shall adapt Ockham's razor principle as our guiding tool. This principle has been a fruitful idea in mathematics, complexity theory and information theory (see, e.g., Ref. 10 and references therein). It says that out of the set of possible hypotheses the most likely is the simplest one, which is consistent with all observations. There are so many known facts about the brain that one may try to take advantage of this razor in model construction. Our recipe is the following: (1) Choose basic and widely accepted properties of the brain. (2) Select as few as possible of those. Call them starting hypotheses. (3) Try to derive other properties as consequences of those hypotheses. The derived properties are the "predictions" of your "model", irrespective whether those have been measured or not. If those have not been measured then falsifying experiments can be suggested. The main point is this: Even if such falsifying predictions have not been reached the probability that a model is related to reality is the highest for the one that makes the minimum assumptions. The set of possible hypotheses is not an ordered set and thus concepts like "minimum set" or "simplest set" may remain subject to debate. In what follows, we make a strong effort in minimizing the set of our starting hypotheses. Given the set we shall derive properties that have been measured and properties that have not. The latter subset of properties serves as falsifying predictions to our "model". The set of hypotheses should be consistent with all observations and thus it is crucial whether a model can be put into the form of a BU model. In other words, we consider CNS like modeling observation. In our view, *CNS is a non-linear functional of observations.* The model should be consistent with this special observation. Thus, the eventual test is in CNS modeling. We shall argue that our model could be translated into a BU model with population coding and Hebbian learning.

The paper is organized as follows. First, our selection of basic assumptions will be described. This section is followed by the description of two "consequences", such as the description of speed-field tracking and the description of the Static and Dynamic State (SDS) control scheme. Then the description of the control method in external space is given. Mathematical considerations and computer demonstrations are provided here. In the discussion section, we shall review the control architecture as a model of the basal ganglia-thalamocortical (BTC) loops. Conclusions are drawn in the last section. A more detailed description of the basal ganglia (BG) and the BTC loops are given in the Appendix.

2. Some Basic Properties of the Control Faculty of the Brain and Their Modeling Consequences

2.1. *Lack of trajectory representation*

There is no evidence of signals in the brain that would specify trajectories. It seems that the brain formulates signals that specify positions and directions of targets in extrapersonal space. Neuronal signals of this kind can be found, e.g., in the parietal association cortex.[73] The starting point of our model is that the set of possible trajectories is too large to be directly specified. Instead, a dynamic and distributed representation is created by the object under control (i.e., the limb), the target to be reached, and the obstacles to be avoided. This distributed representation is engaged in specifying positions and deriving directions from position differences.

2.1.1. *Speed-field formulation of the control task*

The control problem should be formulated in terms of directions pointing towards target positions. Such mapping from target state(s) to direction is called speed-field. The important feature of speed-field is that motion is not specified in time. The control task is called speed-field tracking (SFT). SFT formulation is flexible because it allows motions to speed up or to slow simply by scaling of the speed-field. The controller, in turn, maps state and direction to action. This map is called position-to-directional action (PDA) map. It has been shown that the control executed and the measured directional information together is sufficient to train PDA maps by Hebbian means. Such training gives rise to robust population coding.[18]

2.2. *Comparator function (differencing property) of the BG*

As is known, the corticostriatal projections are glutamatergic and excitatory whereas the striatopallidal and the pallidofugal connections are GABAergic and inhibitory. Within the motor circuitry of the BG, there are two major projections. The direct pathway projects from a certain subpopulation of striatal neurons in the putamen to the internal segment of the globus pallidus (GPi). With regard to the indirect pathway, one arm has its origin in a different population of striatal neurons, then proceeds through the external segment of the globus pallidus (GPe) and the subthalamic nucleus (STN) before reaching the GPi. The second arm of the indirect pathway is made up of GPe projections to the output nuclei themselves. All intrinsic projections are GABAergic and thus inhibitory, except the projections from the STN to the GPi, which are excitatory. The two main arms of the two pathways are illustrated in Fig. 1.

These pathways are complemented by the GPe–reticular thalamic nucleus (RTN)–ventrolateral thalamus pathway. Also, (i) most of the putamen-projecting sensorimotor areas send excitatory connections directly to the STN, and (ii) the STN sends glutamatergic connections back to the GPe and the putamen.

There appears to be a functional consistency among the various GPe and GPi projections.[1]

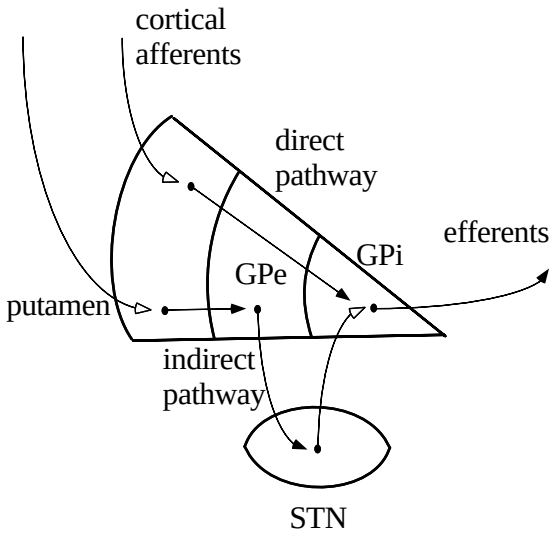


Fig. 1. Schematic representation of parts of the direct and the indirect pathways. The two main arms of the two pathways are illustrated in the figure. Both pathways are excited by corticostriatal projections. The part of the direct pathway shown in the figure inhibits the pallidofugal connections of the internal segment of the globus pallidus (GPi) and thus disinhibits the thalamocortical connections. The part of the indirect pathway shown in the figure inhibits the projections of the GPe and, in turn, disinhibits the excitatory projections of the STN to the GPi and thus reinforces the pallidofugal connections and, in turn, suppresses the thalamocortical projections. Open and filled arrows represent excitatory and inhibitory connections, respectively.

Activation of medium spiny neurons (MSNs) associated with either arm of the indirect pathway will tend to increase BG output by increasing neuronal activity at the level of the output nuclei. In one case the BG output is increased by disinhibiting the STN with its excitatory projections to the GPi and the substantia nigra pars reticulata (SNr), and in the other by directly disinhibiting the GPi and the SNr. In contrast, activation of MSNs associated with the direct pathway tends to decrease BG output by directly suppressing activity at the level of GPi and SNr. The net result is that cortically initiated activation of the (indirect) direct pathway will tend to (suppress) enhance re-entrant thalamocortical excitation by (increased) decreased inhibitory outflow from BG to thalamus.

The functional consistency of various GPe and GPi connections is further emphasized by the dopamine pathways. Dopaminergic input to the putamen consists of nigrostriatal projections that originate in the substantia nigra pars compacta (SNc). The cortical motor and premotor areas receive separate dopaminergic projections from the ventral tegmental area (VTA). At the network level dopamine has an inhibitory action on the striatal GABAergic cells projecting to the GPe giving rise to a decrease of the effect of the indirect pathway. It is an open question whether dopamine has a neutral or an excitatory action on the neurons projecting directly to the GPi and the SNr and if it gives rise to an increase of the effect of the direct pathway.[27]

2.2.1. *Speed field tracking and SDS control*

The comparator like function (the differencing property) found in the BG, which forms the basis of the TDI model will be assumed the general property of these nuclei here. We shall assume that most comparisons between intended and experienced parameters of motion are made in these nuclei. Below we show that this assumption when built on SFT provides the ultimately uniform bounded property. The related architecture is called the Static and Dynamic State (SDS) feedback control scheme.[61,62,65,43,45]

The SDS feedback architecture performs speed-field tracking. Let $D \subseteq \Re^n$ denote the domain of the system's state with the equation of motion given by

$$\mathbf{u} = \mathbf{A}(\mathbf{q})\dot{\mathbf{q}} + \mathbf{b}(\mathbf{q}) \tag{1}$$

where $\mathbf{q}$ is the state vector of the system, the dot denotes temporal derivation, and $\mathbf{u} \in \Re^m$ is the control vector. In case of a multi-segment idealized robotic arm, components of the state vector of the arm correspond to the angles and to the angular velocities of the arm. Accordingly, the temporal derivative of the state vector corresponds to the angular velocities and to the angular accelerations of the arm. For the sake of notational simplicity the dependence of $\mathbf{A}$ and $\mathbf{b}$ on $\mathbf{q}$ will from now on not be explicitly represented. Let us now assume that we have an estimate of the true inverse-dynamics function $\Phi(\mathbf{q}, \dot{\mathbf{q}}) = \mathbf{A}\dot{\mathbf{q}} + \mathbf{b}$ given by $\hat{\Phi}(\mathbf{q}, \dot{\mathbf{q}})$.

The SDS feedback control equations can then be written as

$$\mathbf{u} = \mathbf{u}_{\text{ff}}(\mathbf{q}, \dot{\mathbf{q}}, \mathbf{v}(\mathbf{q})) + \mathbf{w} \tag{2}$$

$$\dot{\mathbf{w}} = \Lambda(\hat{\Phi}(\mathbf{q}, \mathbf{v}(\mathbf{q})) - \hat{\Phi}(\mathbf{q}, \dot{\mathbf{q}}) \tag{3}$$

where $\mathbf{u}_{\text{ff}}$ is the so called feedforward controller (to be specified later), $\dot{\mathbf{w}} = \mathbf{u}_{\text{fb}}$ is the so called feedback controller that gives rise to the feedback control vector w upon temporal integration, $\Lambda > 0$ is the gain of feedback, and the desired motion is determined by a speed-field tracking task that prescribes the speed vector $\dot{\mathbf{q}}$ of the system as a function of the state vector:

$$\dot{\mathbf{q}} = \mathbf{v}(\mathbf{q}) \tag{4}$$

Below, we review some of the results on the SDS control scheme.[61,62,65]

If $\mathbf{M}$ is a real quadratic matrix ($\mathbf{M} : D \to \Re^{p \times p}$, $p > 0$) then let $\mathbf{M} > 0$ denote that $\mathbf{M}$ is positive definite ($\mathbf{M}$ is said to be positive definite uniformly over D iff for all $\mathbf{q} \in D$ the term $\mathbf{M}(\mathbf{q})$ is positive definite and there exists an $\varepsilon > 0$ such that $\lambda_{\min}\mathbf{M}(\mathbf{q}) > \varepsilon$ holds for all $\mathbf{q} \in D$). Similarly, if $\mathbf{M}$ is a matrix field over D, let $\mathbf{M} > 0$ denote that $\mathbf{M}$ is uniformly positive definite over D. Then the following theorem holds.

Theorem 1

Assume that the feedforward controller has the form

$$\mathbf{u}_{\text{ff}}(\mathbf{q}, \dot{\mathbf{q}}, \mathbf{v}) = \hat{\Phi}(\mathbf{q}, \mathbf{v}(\mathbf{q})) - \hat{\Phi}(\mathbf{q}, \dot{\mathbf{q}}) \tag{5}$$

(which is the same as the input of the feedback integrator). Further, assume the following:

(i) $\hat{\Phi}(\mathbf{q}, \dot{\mathbf{q}}) = \hat{\mathbf{A}}\dot{\mathbf{q}} + \hat{\mathbf{b}}(\mathbf{q})$

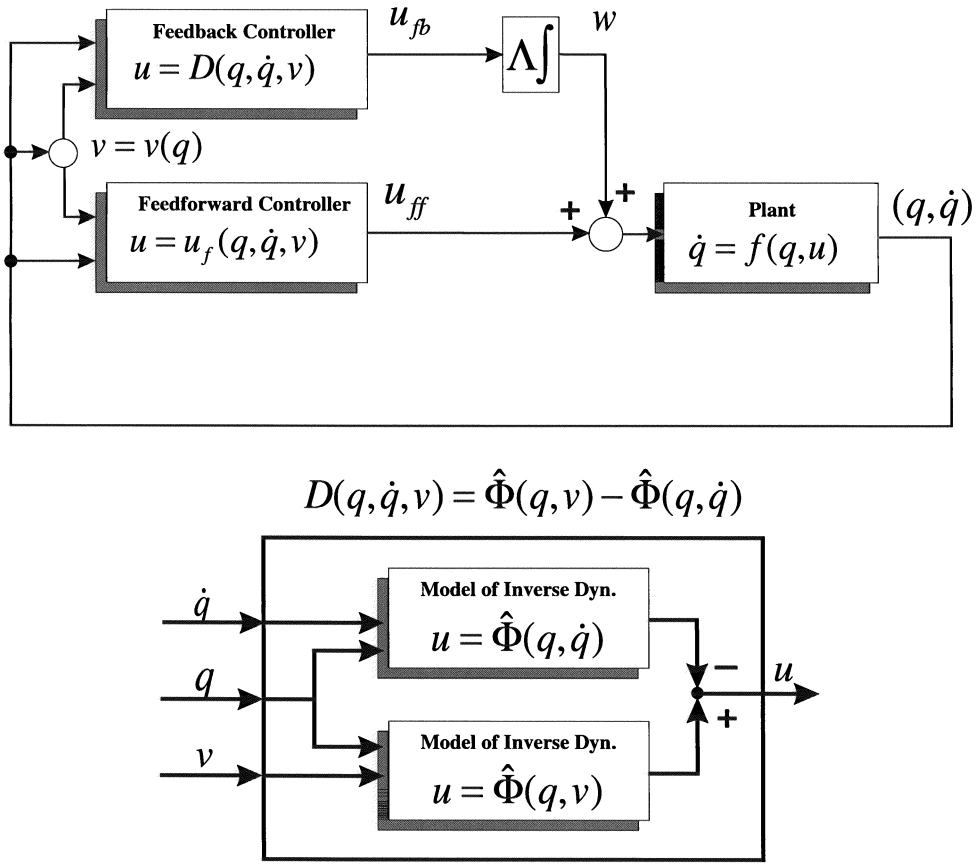


Fig. 2. Compensatory control. The scheme utilizes identical feedforward and feedback controllers built from estimates of the approximate inverse dynamics. The feedforward and feedback controllers provide control vectors $\mathbf{u}_{\rm ff}$ and $\mathbf{u}_{\rm fb}$, respectively. The feedback control vector is time integrated and the resulting control vector is added to the feedforward control vector. The state equation of the system gives rise to changes in the state of the system subject to its actual state, the control vector, and the dynamics of the system. The identical feedforward and feedback controllers are detailed in the lower part of the figure. According to the scheme each controller comprises two copies of the model of the inverse dynamics. These models are inputted by the "desired" and the "experienced" speeds, respectively. The outputs of these models undergo differencing.

(ii) $\mathbf{A}^T\mathbf{A}$, $\mathbf{A}^T\hat{\mathbf{A}}$ and $\hat{\mathbf{A}}^T\hat{\mathbf{A}}$ are uniformly positive definite over D (sign-properness condition)

(iii) $\mathbf{A}$, $\mathbf{v}$, $\mathbf{b}$ are bounded and have uniformly bounded derivatives w.r.t. $\mathbf{q}$ over D

Then for all Λ the error of tracking $\mathbf{v}(\mathbf{q})$, $\mathbf{e} = \mathbf{v}(\mathbf{q}) - \dot{\mathbf{q}}$ is ultimately uniformly bounded (UUB) and, further, the eventual bound b of the tracking-error can be made arbitrarily small. More specifically $b = O(1/\Lambda)$ and the eventual bound for the time reaching $\|e\| < b$ is independent of Λ. The proof of this theorem relies on a Liapunov-function approach. Firstly the relation

$$(\mathbf{A} + \hat{\mathbf{A}})\mathbf{e} = \mathbf{A}\mathbf{v}(\mathbf{q}) + \mathbf{b} - \mathbf{w} \qquad (6)$$

can be employed to show that $L = \frac{1}{2}\mathbf{e}^T[(\mathbf{A} + \hat{\mathbf{A}})^T(\mathbf{A}+\hat{\mathbf{A}})]\mathbf{e}$ is an appropriate semi-Liapunov function. Then the proof is complete.[65,a] The control scheme is depicted in Fig. 2.

The subtle point of this theorem is the requirement that $\hat{\mathbf{A}}^{\mathbf{T}}\hat{\mathbf{A}}$ should be uniformly positive

[a]We note that Eqs. (3) and (5) can be seen as integrals; Eq. (3) is an integral with a Dirac's delta function and Eq. (5) is an integral a constant function as temporal kernels. The theorem can be generalized to other, e.g., exponential temporal kernels.

definite over D. The theorem gives rise to semi-global stability result. Notice too, that the particular form of the feedforward and feedback controllers makes it unnecessary to build an estimate of $\mathbf{b}$.

An important property of this theorem is that in the error equation (Eq. (6)) the r.h.s. does not depend on the approximated inverse-dynamics. This feature can be exploited to show that proof of the theorem is not affected if $\hat{\mathbf{A}}$ and $\hat{\mathbf{b}}$ vary in time but the conditions of the theorem (most importantly the sign-properness condition) remain valid at every instant. Thus, we get the following important corollary:

Corollary 1

Suppose that the conditions of the theorem hold and suppose that $\hat{\mathbf{A}} = \mathbf{A}(t)$ and $\hat{\mathbf{b}} = \mathbf{b}(t)$. Next assume that $\mathbf{A^T}\hat{\mathbf{A}}(t)$ and $\hat{\mathbf{A}}^{\mathbf{T}}\hat{\mathbf{A}}(t)$ are uniformly positive-definite over D for all $t > 0$, and that $\hat{\mathbf{A}}(t)$ is bounded. Then the conclusions of the theorem still hold.

The lower bound of Λ is inversely proportional to $\inf_{t,\mathbf{q}} \lambda_{\min}(\hat{\mathbf{A}}^{\mathbf{T}}(q)\hat{\mathbf{A}}(\mathbf{q}, t)$ and $\inf_{t,\mathbf{q}} \lambda_{\min}(\hat{\mathbf{A}}^{\mathbf{T}}(\mathbf{q}, t)$ $\hat{\mathbf{A}}(\mathbf{q}, t)$ and proportional to $\sup_{t,\mathbf{q}}(\|\mathbf{A}(\mathbf{q}, t)\| + \|\mathbf{A}(\mathbf{q}, t)\|^2)$. The uniform positive-definiteness conditions of the corollary follow, e.g. $\hat{\mathbf{A}}$ is bounded away from singularities uniformly over D: an assumption often required in adaptive control.[57] It is clear too, that the stability result does not depend on the specific adaptation mechanism utilized, which is a rare condition in adaptive control theory. However, one has to ensure that the sign-properness conditions are obeyed.

2.3. *Gelatinous medium around the limbs*

It has been shown by Graziano and Gross that the visual space is represented by bimodal neurons as a gelatinous medium surrounding the body that deforms in a topology-preserving fashion whenever the head rotates or the limbs move.[24] Such a map gives the location of the stimulus with respect to the body surface in somatotopic coordinates. Based on these findings one may assume that the configuration of the limb (or, maybe, the whole body) is available to the neurons of the putamen.

2.3.1. *Control using parameters of external space*

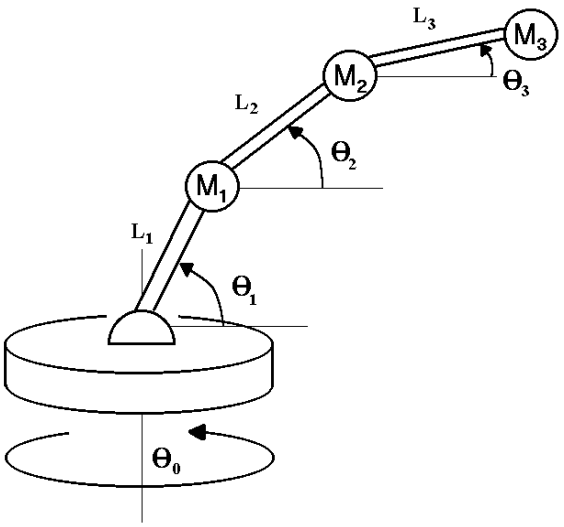
The problem with controlling of a multi-segment robotic arm is that controllers act on joint angles (in non-linear fashion). In most cases, however, motion tasks are not formulated in terms of joint angles, but in terms of starting and target positions of the end-effector. Moreover, a given configuration can be described by considering each joint as individual end-effectors. It is thus tempting to effectively shortcut the learning of the inverse kinematics and to learn the inverse dynamics using the low-dimensional metric of the external space instead of the high-dimensional metric of the configurational space. We shall address this problem as follows:

Traditional formulation of the control problem of an n-joint robotic arm starts from the following equation (see, e.g., Ref. 57 and references therein):

$$\mathbf{u} = \mathbf{M}(\mathbf{\Theta})\ddot{\mathbf{\Theta}} + \mathbf{C}(\mathbf{\Theta}, \dot{\mathbf{\Theta}}) \tag{7}$$

Here $\mathbf{\Theta} = (\Theta_0, \Theta_1, \ldots, \Theta_{n-1})$ is the vector of angular positions of the robot (Θ_0 is the angular position of the robot base axis, $\Theta_1, \ldots, \Theta_{n-1}$ are the angular elevation of the arm-segments above horizontal), $\mathbf{u} = (u_0, u_1, \ldots, u_{n-1})$ is the torque vector of actuators (u_1 denotes the torque of the base, $u_2, \ldots, u_n$ denote the torque actuators of the respective arm segments), $\mathbf{M}(\mathbf{\Theta})$ is the inertia matrix, $\mathbf{C}(\mathbf{\Theta}, \dot{\mathbf{\Theta}})$ represents Coriolis, centripetal forces, gravity loading, and friction terms (see Fig. 3). Note that this equation has the form of Eq. (1) with $\mathbf{q}^T = [\mathbf{\Theta}^T, \dot{\mathbf{\Theta}}^T]$, $\mathbf{A}(\mathbf{q}) = [\mathbf{0}, \mathbf{M}(\mathbf{\Theta})]$ and $\mathbf{b}(\mathbf{q}) = \mathbf{C}(\mathbf{\Theta}, \dot{\mathbf{\Theta}})$. The additive term $\mathbf{b}(\mathbf{q})$ is, however, not bounded if $\|\dot{\mathbf{\Theta}}\|$ goes to infinity. Because the system is a mechanical system *a priori* boundedness of $\dot{\mathbf{\Theta}}$ is granted provided that loss is present and that the controls remain bounded. Prescribed bounds can be achieved, for example, by variants of the σ-modification scheme or a projection method. The SDS scheme has been applied to the control problem of a robotic arm in configurational space and demonstrated robustness.[44,65]

Turning to coordinates in external space, let $\mathbf{x}(\mathbf{\Theta}) \in \Re^3$ denote the Cartesian coordinates of the end-effector in the three-dimensional space. Let us fix an arbitrary target position in the 3D space, let us denote it by $\mathbf{x}_f$. The task is to control the arm such that $\mathbf{x}(\mathbf{\Theta}(t)) \to \mathbf{x}_f$ as $t \to \infty$ and at the same time to have $\dot{\mathbf{\Theta}} \to 0$.



L_1, L_2, L_3: lengths of the arm segments
M_1, M_2, M_3: point masses between the two parts of the arm-segments
J: rotational inertia of base
Θ_0: (the angular position of manipulator base axis, Θ_i ($i = 1, 2, 3$): the angular elevation of the arm segments above horizontal (rotational inertia of base is not shown)

Fig. 3. Control signal in different configurations.

Consider the control

$$\begin{aligned} \mathbf{u} &= \mathbf{w} \\ \dot{\mathbf{w}} &= \Lambda\{\alpha(\mathbf{G}(\mathbf{x}(\boldsymbol{\Theta}), \mathbf{x}_f) - \dot{\boldsymbol{\Theta}}) - \ddot{\boldsymbol{\Theta}}\} \end{aligned} \tag{8}$$

Here $\Lambda > 0$, $\mathbf{G}(\mathbf{x}(\Theta), \mathbf{x}_f) = \mathbf{f}(|\mathbf{x}(\boldsymbol{\Theta}) - \mathbf{x}_f|)\mathbf{S}(\mathbf{x}(\boldsymbol{\Theta}), \mathbf{x}_f)$ where $f : \Re_0^+ \to \Re_0^+$ is some monotonous function of its argument satisfying $f(\mathbf{0}) = 0$ and $f(\mathbf{x}) > 0$ if $\mathbf{x} > 0$, and for each component of $\mathbf{S}$, $S_i(\mathbf{x}(\boldsymbol{\Theta}), \mathbf{x}_f) \subset \{0, 1\}$ ($1 \leqq i \leqq n$) determines the desired direction of motion for the ith joint, which is computed locally at the joint. This local computation at the joint is a local decision whether sign "+" or sign "−" (i.e., an increase or a decrease of the angle) will move the end-effector closer to $\mathbf{x}_f$. This is a simple local decision at each joint angle that can be learned by exploration. The learner can perform an infinitesimal movement at a given joint angle (keeping all the other joint angles fixed) and decide if it corresponded to an increase (or decrease) of the distance to the end-effector. Such input–output pairs can be used to approximate function $\mathbf{S}$ by artificial neural network means, e.g., by support vector methods. In our computer runs, mapping $\mathbf{S}$ was computed at each step according to Fig. 4(a): The end-effector and the center of rotation of the joint were connected by the base axis. Decision concerned incrementing or decrementing the joint angle in order to decrease the angle between the joint axis and the base axis.

Proposition 1

Define $\varepsilon(t)$ by

$$\dot{\boldsymbol{\Theta}} = \varepsilon(t) + \mathbf{G}(\mathbf{x}(\boldsymbol{\Theta}), \mathbf{x}_f) \tag{9}$$

then $\varepsilon(t)$ is UUB.

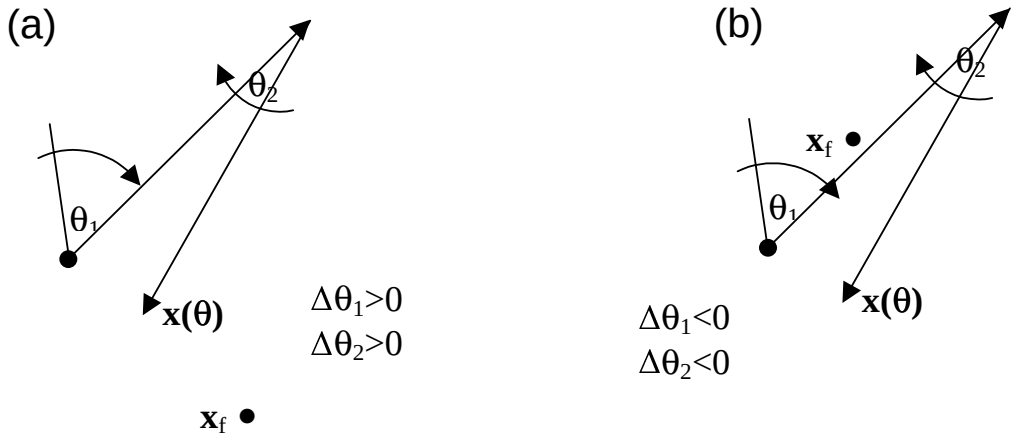


Fig. 4. Control signal in different configurations. (a) The end-effector $\mathbf{x}(\boldsymbol{\Theta})$ can be moved closer to the desired point $\mathbf{x}_f$ by increasing both angles. The arm will not be trapped. (b) The end-effector $\mathbf{x}(\boldsymbol{\Theta})$ can be moved closer to the desired point $\mathbf{x}_f$ by decreasing both angles. The arm will be trapped. However, if angle Θ_1 is kept constant at the early stage of motion then the same method will avoid the trap.

We prove the proposition by means of the theorem. In our system we let $\mathbf{v}(\mathbf{q}) = [\mathbf{0}, \mathbf{G}(\mathbf{q})]^T$, $\hat{\mathbf{A}}(\mathbf{q}) = [\mathbf{0}, \alpha\mathbf{I}]$, where $\mathbf{I}$ is the identity matrix, $\mathbf{G}(\mathbf{q}) = \mathbf{G}(\mathbf{x}(\mathbf{\Theta}), \mathbf{x}_f)$. Λ and α will be modified in the computer simulations. The controlled system takes the inverse form $\mathbf{A}(\mathbf{q}) = [\mathbf{0}, \mathbf{M}(\mathbf{\Theta})]$. In order to prove the UUB property of $\varepsilon(t)$ we need that $\hat{\mathbf{A}}(\mathbf{q})^T\mathbf{A}(\mathbf{q})$ be uniformly positive definite. However, $\mathbf{A}(\mathbf{q})^T\mathbf{A}(\mathbf{q}) = \mathbf{M}(\mathbf{\Theta})$ being uniformly positive definite, which was an assumption on the controlled arm.

The proposition has the important property for the case of speed field tracking: the motion of the controlled arm can be thought of as the perturbation of the system having the following non-linear equation

$$\dot{\mathbf{\Theta}} = \mathbf{G}(\mathbf{x}(\mathbf{\Theta}), \mathbf{x}_f) \tag{10}$$

In turn, it is satisfactory to determine the properties of Eq. (10) to control the system. The equilibrium points of Eq. (10) satisfy $\mathbf{G}(\mathbf{x}(\mathbf{\Theta}), \mathbf{x}_f) = 0$. Function $\mathbf{G}(.,.)$ is designed such that $\mathbf{G}(\mathbf{x}(\mathbf{\Theta}), \mathbf{x}_f) \neq 0$ unless $\mathbf{x}(\mathbf{\Theta}) = \mathbf{x}_f$ and thus the system stops if and only if the arm reaches the desired point $\mathbf{x}_f$.

By construction of $\mathbf{G}$, the distance $|\mathbf{x}(\mathbf{\Theta}) - \mathbf{x}_f|$ decreases monotonically as it can be shown by using infinitesimal analysis. However, it is not guaranteed that this distance goes to zero because the dynamics has non-repelling equilibrium points satisfying $|\mathbf{x}(\mathbf{\Theta}) - \mathbf{x}_f| > 0$. Consider, for example, the case depicted in Fig. 4(b). A configuration is depicted in the domain of attraction of an equilibrium point with $\mathbf{x}(\mathbf{\Theta}) \neq \mathbf{x}_f$. The arm can be trapped in such configurations.

The trap is a consequence of our requirement that each control component should decrease the distance between the actual position and the desired position. The trap can be avoided by requiring only for a set of the control components to act accordingly. Consider that the joints of Fig. 4(b) are connected in a rigid manner to each other. In turn, angle Θ_2 cannot be modified. Under this condition and using our simple output feedback control scheme the arm will reach a position where angle Θ_2 can be released again without the arm being locked in a local minimum. In other words, we need configurational modulation supervising the output control feedback scheme in a selective manner. The selective supervising architecture should learn the subset of angles that can be released for a given pair of starting and target position. Similar problem arises in the case of obstacle avoidance.

3. Computer Simulations

In the computer simulations, a four-segment robotic arm (Fig. 3) was controlled using Eqs. (2) and (3). For the simulations the Robotics Toolbox designed by P. I. Corke (*http://asgard.mlb.dmt.csiro.au/staff/PeterCorke.html* see also *http://cswww.essex.ac.uk/Research/tuuv/simulators.html*) was used. Mapping $\mathbf{S}$ was computed at each step according to Fig. 4(a). The sign of the control component corresponding to a joint angle was determined. Decision concerned whether incrementing or decrementing the joint angle in rest would decrease the angle between the joint axis and the base axis.

Note that a four segment robotic arm offers many possible solutions to most tasks. The controller makes use of parameters measured in external space local to each segment and its relation to the end-effector. However, the controller as described does not use information about distant segments or information about overall the number of segments. In turn, the controller can be scaled in the number of segments. We assume that the parameterization involving the direction of the arm segment and the relative position between end-effector and joint is provided by the "gelatinous medium" surrounding the limb segments.

We had performed three runs with different parameter settings and different inverse dynamics. The parameters of the computer simulations are given in Table 1.

Comparisons are made with a similar scheme utilizing "perfect" inverse dynamics. "Perfect" inverse dynamics means that to each joint the desired directions were specified according to Fig. 4(a), control components to each joint were computed using the perfect inverse dynamics without the feedforward controller and were utilized simultaneously. We have studied the angles of the joint versus time (Fig. 5), the angular velocities of the joint versus time (Fig. 6), the coordinates of the end-effector versus time (Fig. 7), and the end-effector trajectories (Fig. 8). The computer simulations demonstrate the robustness of this crude controller. The larger the values of parameters α and Λ, the better the SDS architecture using external space controller

Table 1. Parameters of the computer simulations. Masses: 4.0 kg, length of segments: 0.5 m. Motion is performed in the $y = 0$ plane. In all cases the starting angles were set to zero that corresponds to a horizontal arm positioned in the xz plane. The coordinates of the target position were $x = 0.5$, $y = 0.0$ and $z = 0.5$. The choice of function $f : f(|\mathbf{x}(\mathbf{\Theta}) - \mathbf{x}_f|) = |\mathbf{x}(\mathbf{\Theta}) - \mathbf{x}_f|$. Controller using Eq. (8) deals with parameters of the external space. Equations (3) and (5) correspond to "perfect" inverse dynamics. This means that to each joint the desired directions were specified according to Fig. 4(a), control components to each joint were computed using the perfect inverse dynamics and were utilized simultaneously.

Λ	α	Controller using Eq. (8)	Controller using Eqs. (3) and (5) (with $\hat{\Phi} = \Phi$)
30	100		+
30	100	+	
200	600	+	

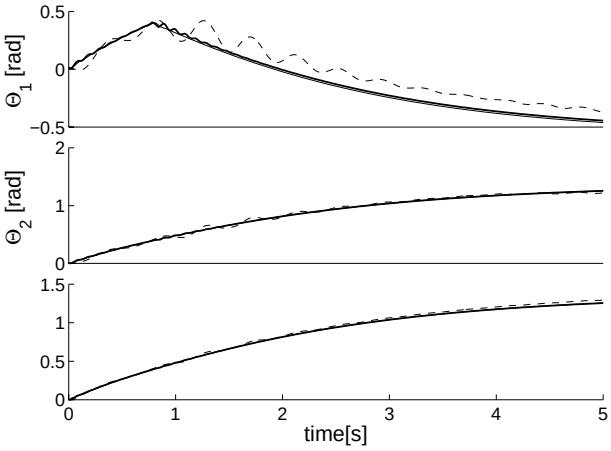


Fig. 5. Joint angles versus time. Thin solid line: Controller using SDS scheme and perfect inverse dynamics. Dashed and bold (wavy) lines: Controller using SDS scheme external space parameters. SDS parameters for dashed and for thin solid lines: $\alpha = 100$, $\Lambda = 30$, for bold line: $\alpha = 600$, $\Lambda = 200$.

approximates the SDS architecture using perfect inverse dynamics. Limitations are dictated by noise considerations. The proper choice of signs of the computer experiments approximates straight paths only during the second part of the motion. It can be seen that the SDS scheme and configurational modulation can provide straight paths. It is clear, however, that straight paths are not warranted by the SDS scheme.

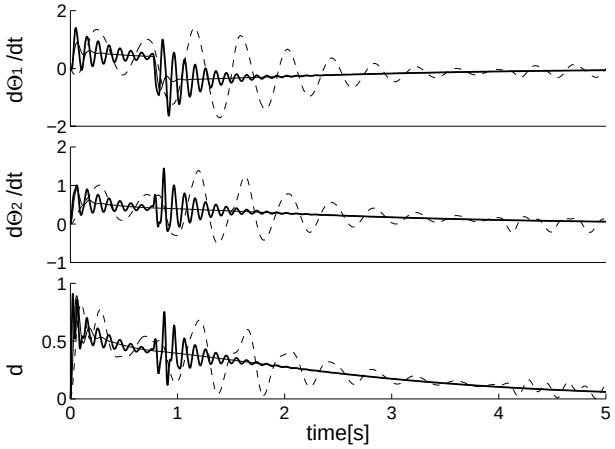


Fig. 6. Joint angular velocities versus time. Thin solid line: Controller using SDS scheme and perfect inverse dynamics. Dashed and bold (wavy) lines: Controller using SDS scheme external space parameters. SDS parameters for dashed and thin solid lines: $\alpha = 100$, $\Lambda = 30$, for bold line: $\alpha = 600$, $\Lambda = 200$.

4. Discussion

4.1. *Main properties of the basal ganglia-thalamocortical loops*

Over recent years a considerable amount of research has been carried out to characterize the structure and the functional organization of the BG that comprise the caudate nucleus, the putamen, the ventral striatum — the latter three forming the striatum — the globus pallidus, the substantia nigra, and the subthalamic nucleus. The BG have several puzzling features, including robust parallel input–output organization[4] and high convergence ratios from input to output.[54,55,71,53,19,34] The main parallel channels of the BG are the skeletomotor, the oculomotor, the associative, and the limbic channels. The present work concerns the skeletomotor channels; these receive inputs from most of the cerebral sensorimotor regions, including the supplementary motor area (SMA) and the motor cortex (MC). The output of the BG returns to the same regions via the thalamus and thus the BG form one of the stations of the basal ganglia-thalamocortical loops. The loops are further divided into functionally segregated and highly specific parallel channels.[29] This high specificity concerns both the initiating and the receiving cortical areas, these areas show a degree of overlap and it seems that the information undergoes remapping within the putamen.[17,22]

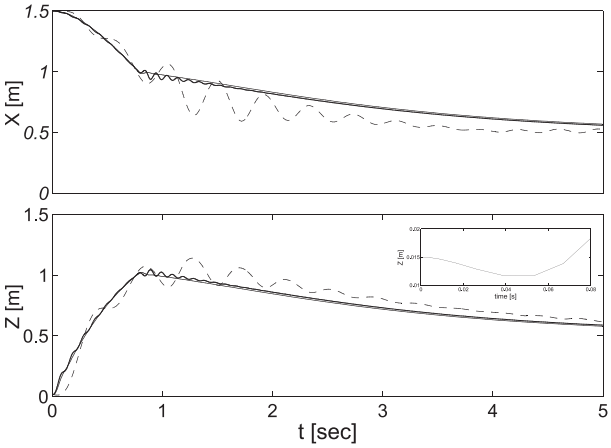


Fig. 7. Coordinates of end-effector versus time. Thin solid line: Controller using SDS scheme and perfect inverse dynamics. Dashed and bold (wavy) lines: Controller using SDS scheme external space parameters. SDS parameters for dashed and thin solid lines: $\alpha = 100$, $\Lambda = 30$, for bold line: $\alpha = 600$, $\Lambda = 200$. The inset depicts the initial perturbation of the controller caused by the weight of the arm observable for the dashed line.

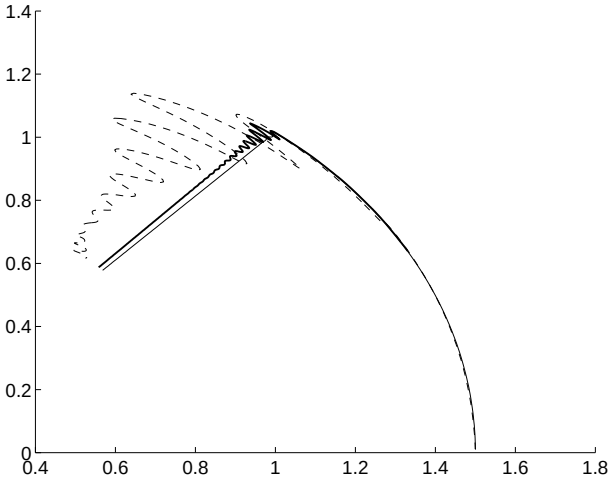


Fig. 8. End-effector trajectories. Thin solid line: Controller using SDS scheme and perfect inverse dynamics. Dashed and bold (wavy) lines: Controller using SDS scheme external space parameters. SDS parameters for dashed and thin solid lines: $\alpha = 100$, $\Lambda = 30$, for bold line: $\alpha = 600$, $\Lambda = 100$.

The parallel organization is a prominent feature of the BG. Experiments in which anatomical tracers were placed into monkey brains to label simultaneously anterograde sensorimotor inputs to the striatum and retrograde striatal outputs to the pallidum showed that labeled input fiber clusters can overlap clusters of backward labeled projection neurons quite precisely.[17,22] The emerging picture is that the information is dispersed to the distributed modules, the matrisomes, in the striatum,[47,13,59] but it can be brought together again at the next stage of processing, i.e., at the pallidum. This pattern has been observed both for striatal projections to the GPe and for projections to the GPi.[22] The organization of the BG can thus be viewed as a family of reentrant loops that are organized in parallel. The output of the BG returns to parts of those same cortical fields by way of specific BG recipient zones in the dorsal thalamus; that is, the outputs of the matrisomes are re-collected according to the recipient zones that correspond to the cortical fields. The convergence of direct and indirect projections to the GPi is estimated as 1000:1.[54,55,71,53,19] The convergence is manifested by the wide dendritic arborization of pallidal and nigral neurons oriented at right angles to the incoming striatal axons with small number of contacts between the striatal axon and a single pallidal cell. Transneuronal retrograde tracing indicates that the BG-thalamocortical loops are organized to form multiple segregated output channels.[29] The connectivity structure thus looks like an extensive input–output remapping with high convergence ratios towards the GPi.[55,71,53,19] It is a constraint or prediction of the SDS scheme that only different re-entrant loops can correspond to different the sign-proper feedback channels. This view is supported by the findings of Nini *et al.*[50] that neurons of the globus pallidus do not show correlated activity.

Finally, we note that the BG, that exhibits modular structure for control, has another type of modular structure: the striosomes.[23,21,20,30] According to suggestions in the literature, the matrisomes of the BG are trained by these striosomes. It has been argued that striosomes perform differencing between expected and experienced rewards for reinforcement learning.[5,30,49,31]

4.2. *Neuronal classification according to the SDS scheme*

Before discussing neuronal classification in terms of the SDS scheme, we ought to consider the sign-properness property of the scheme. The structural constraint that $\hat{\mathbf{A}}(\mathbf{q}) > 0$ should apply everywhere

is a demanding requirement for the control architecture. Let us consider the following examples.

Example 1

Consider the problem of mirror writing when the trajectory to be followed by the pen is observed through a mirror. The ordinary approximate inverse dynamics tuned for writing without any mirror will not be sign-proper in this case and the motion will diverge in an exponential fashion. If the components of the approximate inverse dynamics are modified accordingly and the mirror writing becomes sign-proper, then writing without a mirror becomes divergent. It seems reasonable to divide the task space and utilize independent "channels" for the two different tasks. In this case, the approximate inverse dynamics should have two independent channels each activated upon recognition of the appropriate task.

Example 2

Consider the sudden task of catching a ball that will pass just above our hand. The control signal to be launched depends on several factors, including the present configuration of the body, the constraints of possible motions as well as the actual speed of parts of the body — including the hand. As it has been shown,[65] the actual speed can be neglected in the control problem. Body configurations, however, cannot be neglected. Raising the hand requires, for example, different shoulder motions depending on whether our body is in a vertical or horizontal position, etc. If any of the control components destroys the sign-properness requirements of the theorem in any domain of $\Re^n$ then the SDS scheme requires a separate sign-proper channel for control.

Example 3

Configurations that can be trapped and that cannot be trapped in local minima should be distinguished. Distinction may require considering some joints as rigid in different time domain of motion execution.

Example 4

Even if the controller avoids local minima, still, the path may be curved as it is demonstrated in Figs. 7 and 8. Learning may also concern the selection of sign combinations that can give rise to straight paths for given positions of the end-effector and the target. We note that the direction between the end effector and the target is known in external space. Thus, the learning of straight motion between points concerns the learning of the PDA map. The SDS scheme requires the approximate learning of a position and direction to infinitesimal angular modification mapping to generate UUB control outputs. Sign-proper solutions may be found by rendering one or more joint angles rigid. All four examples show that the SDS scheme, although robust and globally stable, still requires a complex setting for proper control: different tasks and different configurations may require different parallel channels for control.

Now, we shall identify the neuronal groups suggested/predicted by the SDS scheme.

- Desired acceleration may be expressed as the difference between the desired direction and the experienced direction.
- Feedforward control vector can be computed by means of subtracting the estimate of the "experienced control vector" from the estimate of the "desired control vector".
- Feedback control vector is the time-integrated value of the feedforward control vector.

In turn, it is possible to characterize the model neurons of the control areas that prepare the signals, which undergo differencing. For learned and unperturbed motions, one has the following classes:

(a) Model neurons representing the desired acceleration will have higher firing rates in the so-called preparatory phase, i.e., before motion is initiated, if motion closely follows the planned direction upon initiation.
(b) The firing of model neurons representing experienced direction is motion related.
(c) Model neurons that represent the desired direction have mixed characteristics.
(d) Model neurons that represent the output of the approximate inverse dynamics, i.e., neurons that express the components of the control vector, behave as "muscle like" neurons and exhibit motion related firing.

This classification of model neurons is in good agreement with the experimental findings[68,67,3,11,46] and classification.[3,11] However, for perturbed

motions the SDS model predicts other classes too. In case of perturbed motion "muscle like" neurons related to the desired and experienced control components will be different. In this respect, the SDS model predicts that "muscle like" neurons corresponding to the experienced control components and "muscle like" neurons corresponding to the desired control components will appear together. This prediction is a falsifying test to the model.

4.3. *Scaling and learning properties*

The described output feedback controller has the following properties:

(i) The controller requires the discretization of the external space instead of the configurational space and thus it saves discretizing computational units. Moreover, the number of neurons does not scale exponentially with the number of joints.

(ii) The controller requires configurational modulation to determine which joints are rigid and which joints are not rigid at a given moment of a given task. The configurational modulator may require as many as $2k$ different states where k is the number of joints, because any joint can be either rigid or non-rigid. In most cases, however, a much smaller number may be sufficient. To each actual configuration and to each state of the configurational modulator there belongs a volume in 3D that can be reached by the output feedback controller. One is interested in the smallest number of states of the configurational modulator that can cover the reachable space. This learning problem, together with the learning problems of self-avoidance and obstacle avoidance need further considerations. We make three notes here:

- the learning problem is inherently a selective learning problem,
- different arm designs may be sensitive to this problem in different ways,
- if the arm is connected to some base (e.g., the body) then the problem of keeping parts standing still while moving other parts is also related to the learning problem of the configurational modulator, because different parts are coupled.

(iii) Control parameters concern the external space.[b] Parameters of speed-fields can be derived by differencing.

(iv) If the configurational modulator is trained then the rest of the controller is robust.

4.4. *Action-perception loop and the hierarchy of controllers*

The model that has been described integrates action and perception into a loop, because it is an output feedback controller. This feature is further emphasized by the selective aspect of configurational modulator. The configurational modulator can be viewed as a switching unit that switches between modules. For each module, one can design a mapping that represents a single valued and smooth switching inverse kinematics.[39,38] According to footnote b, the inverse kinematics is sufficient to approximate the inverse dynamics to control the arm provided that the resulting feedback is sign-proper. Then control will give rise to straight motion of the end-effector, a feature that the present model exhibits only in special cases.

The emerging hierarchy of training signals within the SDS scheme is then as follows:

- there is the system of primary reinforcement that trains the striosomal modules to predict future reinforcements by means of the dopaminergic neurons of the SNc;
- then, in turn, the dopaminergic neurons of the striosomal modules train (reinforce) the matrisomes to select sign-proper and trap avoiding configurations for feedback correction in accordance with psychophysical findings;[28]
- sign-proper and trap avoiding inverse kinematics can be considered as a set of modules,[39,38]
- modules can be combined into higher-order modules.[39,38,35]

The striosome model based on differencing[58] performs reinforcement learning. This differencing can

[b]As it has been shown[14] the actual speed can be neglected owing to the fact that one can use the "simplified inverse-dynamics" $\hat{\Phi}_0(\boldsymbol{\Theta}, \ddot{\boldsymbol{\Theta}}) = \hat{\boldsymbol{\Phi}}(\boldsymbol{\Theta}, 0, \ddot{\boldsymbol{\Theta}})$, where $\hat{\boldsymbol{\Phi}}(\boldsymbol{\Theta}, \dot{\boldsymbol{\Theta}}, \ddot{\boldsymbol{\Theta}}) = \hat{\mathbf{M}}(\boldsymbol{\Theta})\ddot{\boldsymbol{\Theta}} + \hat{\mathbf{C}}(\boldsymbol{\Theta}, \dot{\boldsymbol{\Theta}})$. This is so, because $\dot{\boldsymbol{\Theta}} = 0$ in $\hat{\Psi}$ corresponds to an additive perturbation ($\mathbf{M}$ does not depend on $\dot{\boldsymbol{\Theta}}$) and this additive perturbation can be compensated without any additional requirements.

be reorganized to the form of difference between "expected" and "experienced" instantaneous rewards. It has been suggested that the basal ganglia is engaged in the differencing between planned (desired, expected) and experienced parameters of motion and cognition.[45]

The properties of the SDS controller are in line with the findings of Thelen and co-workers (Ref. 69 and references therein) about the learning of motion of infants. Thelen and co-workers have studied the first attempts at reaching of infants. They have shown that these attempts demonstrate considerable variations in kinematics, kinetics and muscle patterning. The common element found in the highly variable moves was that the infants eventually got their hands to the toy. The impression was that initially the infants worked on the shaping of the force dynamics, and that further improvements in kinematics were subsequent to this primary parameterizations.[69] If this primary parameterizations corresponds to the separation of sign-proper SDS feedback channels then it may be considered as a selective process.[14]

Experimental and theoretical investigations lead to the stochastic residual search[39,38] that can be viewed as a modular approach to primary parameterizations. In this algorithm at each step, a "single control unit" is selected which could drive the system towards the target. This single control unit drives the system approximately in the proper direction. In our context, a single control unit corresponds to one of the sign-proper parameterizations for a given task. Control is not executed completely, but it is cut short. The remaining path to the target represents the next step of the algorithm. An interesting finding of stochastic residual search is that configurational clusters emerge during learning by experience. Control that drives the system between configurational clusters can be seen as modules.[39] Modules, or macros, are considered as the means to reduce search problems of exponential complexity to linear. The developing and utilizing modules have been the subject of extensive research.[35,48,52,56] The concept of module development seems to us to bridge the gap between motor control and cognition.

4.5. *Considerations on translation of the model into bottom-up models*

We have shown in previous publications[18,61,43,45] that SFT gives rise to population code. Directions are represented by the population and their activity average is robust. Moreover, the same citations show that the learning of the approximate inverse dynamics can be performed by Hebbian means. This is easy to see. The key is in the inverse dynamics: we learn to associate actions to directions. However, if an action is taken and the direction of motion is measured then this pair is just the perfect training example. That is, no matter what level the inverse dynamics is at, learning concerns mapping between measured values. The organization of computational units can take advantage of topographical organizations.[18] Learning of topographical connections can be accomplished by Hebbian means.[64,66] Detailed description of the speed-field tracking architecture that can perform reactive path planning is described in Ref. 18. The BTC loops and the architecture of the BG can be turned into BU modeling alike to the model of Berns and Sejnowski.[5]

5. Conclusions

Few basic assumptions were used in the derivation of a functional model of the basal ganglia-thalamocortical loops. We list the assumptions and their immediate consequences below. Assumptions are numbered, whereas consequences are itemized:

(1) There is no evidence of representation of trajectories in the brain. Speed-field representation has been found in the brain.

- One is, in turn, restricted to speed-field representation and then control concerns speed-field tracking.
- Speed-field tracking needs knowledge about the inverse dynamics. Approximate inverse dynamics can give rise to instabilities and a robust extension of speed-field tracking is needed.
- Speed-field tracking can be made robust by means of re-entrant temporal integration of error between planned/desired and experienced parameters of motion. The resulting speed-field tracking scheme is ultimately uniformly bounded, that is, it exhibits a global stability property. Re-entrant integration of comparator-like

signals in some loops, such as the BTC loops, follows.
- Learning is fast, because crude sign-proper inverse dynamics suffices for control.

(2) We have assumed that every type of motion related error (i.e., the difference between planned and executed, or desired and experienced parameters of motion) is computed in the matrisome system of the basal ganglia.
- Proper identification of neuronal groups was found. Falsifying experiment on neuronal groups was suggested. Falsifying experiment concerns perturbed conditions.

(3) There is a gelatinous media surrounding the limb that can be understood as a rough coordinate system relative to the limbs. We have assumed that the BTC loop is utilizing such description to deal with the curse of dimensionality problem.
- The controller can be stuck in local minima.
- These local minima can be avoided by selecting one or more joints and rendering those inflexible for particular configurations and target positions.
- Configurational modulation of parallel channels is required for feedback.
- Parallel channels can be modulated by context (consider, e.g., the case of mirror writing).
- Learning concerns external space and thus the dimensionality of the learning problem is limited.
- Motion can be made straight by selecting from sign-proper direction-to-control vector mappings.
- The SDS controller can serve the step-by-step learning of
  - controlling with semi-global stability properties,
  - inverse kinematics and controlling along straight paths,
  - feedforward controlling.
- The SDS controller smoothly turns itself off as the tuning of the feedforward controller proceeds.
- The sign switching requirements of the SDS controller seem to fit well the novel module concept of reinforcement learning methods.
- Discretization of the 2-dimensional surface of the robotic arm allows to consider each discretization point as an end-effector. Learning at neighboring points can take advantage of neighbor training methods.[37]

It has been argued that the skeletomotor control loops of the basal ganglia, thalamus and the cortex that deal with signals that specify the positions of targets in the extrapersonal space[73] can be functionally modeled by the SDS architecture. Similarly to the functional reinforcement learning model of the striosome system of the basal ganglia,[58] the SDS controller can be thought of as a model of the matrisome system of the same nuclei. These concepts together form a unifying view of the basal ganglia. According to this unifying view, the basal ganglia are engaged in differencing between "desired" (planned, expected) and "experienced" parameters of cognition and motion.

Appendix 1. Outline of the organization of basal ganglia

The basal ganglia have as their main components, the caudate and lentiform nuclei, both of which are large subcortical structures.[27] The lentiform nucleus comprises of two parts, the putamen and the globus pallidus (GP), these differ largely in histological structure and anatomic relationships. The putamen — the lateral part of the lentiform nucleus — is similar to the caudate nucleus, and the two structures together form the striatum.

Also included in the BG are the substantia nigra (SN) and the subthalamic nucleus (STN). The SN is to be found on the dorsal side of the basis pedunculi and extends throughout the mesencephalon from the rostral border of the pons into the subthalamic area. In its cell rich part, the substantia nigra pars compacta (SNc), dopamine is synthesized. The dopamine-synthesizing cells are also to be found in the VTA too. The cell-sparse part of the SN, the substantia nigra pars reticulata (SNr), is located ventrally and adjacent to the dopaminergic cell groups of the SNc. The SNr and the GP are the output structures of the BG. The STN influences motor activities primarily through its prominent projections to these two output structures. The STN

receives direct input from motor regions in the cerebral cortex.

The major input to the BG is from the cerebral cortex, and the whole cortical mantle is involved in this highly organized projection system to the striatum. The somatosensory and motor cortices project somatotopically to the putamen.[36,40,41] The putamen also receives direct projections from parietal area 7b.[70,7] The cortical association regions in the frontal, temporal, and parietal lobes project to the caudate nucleus.

GABAergic MSN projections are the most abundant neurons in the striatum.[72] Only the axons of the MSNs reach the two output structures of the BG, the GP and the SN, through the direct and the indirect pathways. These projection neurons are divided into two subpopulations: one projecting to the external segment of the globus pallidus (GPe), the other projecting either to the internal segment of the globus pallidus (GPi) or to the SNr.[51]

The GPi and the SNr, that project to the thalamus and the brain stem, form the major output structures of BG. The pathways to the thalamus are topographically organized: the GPi projects to the ventral anterior–ventral lateral (VA–VL) complex and the ventral pallidum primarily to the mediodorsal thalamic nucleus (MD). The SNr projects to the VA–VL complex and MD nuclei. The VA–VL complex and MD, in turn, project to motor-premotor and prefrontal cortical areas. Other projection sites include the superior colliculus and the reticular formation in the mesopontine tegmentum. The GPe sends inhibitory, feedforward projections to the reticular thalamic nucleus (RTN). The RTN sends GABAergic projections to the BG recipient nuclei of the ventrolateral thalamus.

The nigrostriatal dopaminergic pathway forms one of the side loops of the BG connection system. The dopaminergic cells of the SNc and the VTA project in a more or less topographic fashion to the caudate and the putamen.

The most important side loops involve the STN; these are inputted from the GPe and influence the GPi and the SNr. These side loops form the indirect pathway (see e.g. Ref. 2 and references therein) in contrast to the direct pathway which is formed by direct striatopallidal MSN projections.

All five major structures are functionally subdivided into skeletomotor, oculomotor, associative, and limbic territories based on their physiological properties and their interconnections with cortical and thalamic territories having the same functions.[1]

The striatum has a modular structure with two types of modules, the striosomes[23,21,20] and the matrisomes.[47,13,59] Both the striosomes and the matrisomes have interconnectivity within the modules, but not between them. The tonically active neurons (TANs) of BG tend to lie at striosome–matrisome boundaries.[50] The striosomes are neurochemically specialized patchy input–output zones that tend to collect inputs related to the limbic system and to project to the dopamine-containing SNc. In contrast, the matrisomes receive sensorimotor and associative inputs and project to the output nuclei of the BG.

Any given matrisome receives overlapping inputs from the same body-part representation in different sub-areas of the sensorimotor cortex, so that several sorts of information relevant to that body part converge. Experiments in which anatomical tracers were placed into monkey brains to label simultaneously anterograde sensorimotor inputs to the striatum and retrograde striatal outputs to the pallidum showed that labeled input fibre clusters can overlap clusters of backward labeled projection neurons quite precisely.[15–17] The emerging picture is that the information is dispersed to distributed modules in the striatum, but it can be brought together again at the next stage of processing, i.e., at the pallidum. This pattern has been observed both for striatal projections to the GPe and for projections to the GPi.[22]

Acknowledgments

This work was supported by OTKA (Grants T17100, T32712), and the US-Hungarian Joint Fund (Grant No. 519).

References

1. G. E. Alexander 1995, "Basal ganglia," in *The Handbook of Brain Theory and Neural Networks*, ed. M. A. Arbib (Bradford Books/MIT Press, Cambridge, MA), pp. 139–144.
2. G. E. Alexander and M. D. Crutcher 1990, "Functional architecture of basal ganglia circuits: Neural substrates of parallel processing," *Trends in Neuroscience* **13**, 226–271.
3. G. E. Alexander and M. D. Crutcher 1990, "Preparation for movement: Neural representations of

intended direction in three motor areas of the monkey," *Journal of Neurophysiology* **64**, 133–150.
4. G. E. Alexander and M. R. DeLong and P. L. Strick 1990, "Parallel organization of functionally segregated circuits linking basal ganglia and cortex," *Ann. Rev. Neuroscience* **9**, 357–381.
5. G. S. Berns and T. J. Sejnowski 1995, "How the basal ganglia make decisions," in *The Neurobiology of Decision Making* (Springer-Verlag, Heidelberg).
6. N. Bernstein 1967, *The Coordination and Regulation of Movements* (Pergamon Press, Oxford).
7. C. Cavada and P. S. Goldman-Rakic 1991, "Topographic segregation of corticostriatal projections from posterior parietal subdivisions in the macaque monkey," *Neuroscience* **42**, 683–696.
8. C. I. Connolly and J. B. Burns 1993, "A model for the functioning of the striatum," *Biological Cybernetics* **68**, 535–544.
9. C. I. Connolly and J. B. Burns 1993, "A new striatal model and its relationship to basal ganglia diseases," *Neuroscience Research* **16**, 271–274.
10. T. Cover and J. Thomas 1991, *Elements of Information Theory* (John Wiley and Sons, New York, USA).
11. M. D. Crutcher and G. E. Alexander 1990, "Movement-related neuronal activity selectively coding either direction or muscle pattern in three motor areas of the monkey," *Journal of Neurophysiology* **64**, 151–163.
12. M. R. DeLong 1990, "Primate models of movement disorders of basal ganglia origin," *Trends in Neuroscience* **13**, 281–285.
13. M. Desban, C. Gauchy, M. L. Kernel, M J. Besson and J. Glowinski 1989, "Three-dimensional organization of the striosomal compartment and patchy distribution of striatonigral projections in the matrix of the cat caudate nucleus," *Neuroscience* **29**, 551–566.
14. G. M. Edelman 1987, *Neural Darwinism: The Theory of Neuronal Group Selection* (Basic Books, New York).
15. A. W. Flaherty and A. M. Graybiel 1991, "Corticostriatal transformations in the primate somatosensory system. Projections from physiologically mapped body-parts representation," *Journal of Neurophysiology* **66**, 1249–1263.
16. A. W. Flaherty and A. M. Graybiel 1993, "Output architecture of the primate putamen," *Journal of Neuroscience* **13**, 3222–3237.
17. A. W. Flaherty and A. M. Graybiel 1994, "Input-output organization of the sensorimotor striatum in the squirrel monkey," *Journal of Neuroscience* **14**, 599–610.
18. T. Fomin, T. Rozgonyi, C. Szepesvári and A. Lőrincz 1997, "Self-organizing multi-resolution grid for motion planning and control," *Int. J. Neural Systems*, in press.
19. C. Francois, G. Percheron and G. Yelnik 1995, "Three-dimensional tracing of individual axons following biocytin injection into the striatum of macaques," in *V-th Meeting International Basal Ganglia Society* (Plenum Press, New York), p. 19,
20. C. R. Gerfen 1992, "The neostriatal mosaic: Multiple levels of compartmental organization," *Trends in Neuroscience* **15**, 133–139.
21. A. M. Graybiel 1990, "Neurotransmitters and neuromodulators in the basal ganglia," *Trends in Neuroscience* **13**, 244–254.
22. A. M. Graybiel, T. Aosaki, A. W. Flaherty and M. Kimura 1994, "The basal ganglia and adaptive motor control," *Science* **265**, 1826–1831.
23. A. M. Graybiel and C. W. Ragsdale Jr. 1978, "Histochemically distinct compartments in the striatum of human, monkeys, and cat demonstrated by acetylthiocholinesteras staining," *Proc. Natl. Acad. Sci. USA* **75**, 5723–5726.
24. M. S. A. Graziano and C. G. Gross 1993, "A bimodal map of space: Somatosensory receptive fields in the macaque putamen with corresponding visual receptive fields," *Experimental Brain Research* **97**, 96–109.
25. C. G. Gross and M. S. A. Graziano 1995, "Multiple representations of space in the brain," *The Neuroscientist* **1**, 43–50.
26. C. G. Gross, G. S. Yap and M. S. A. Graziano 1994, "Coding of visual space by premotor neurons," *Science* **266**, 1054–1057.
27. L. Heimer 1995, *The Human Brain and the Spinal Cord* (Springer-Verlag, New York).
28. E. A. Henish and T. Flash 1992, "A computational mechanism to account for averaged modified hand trajectories," in *Advances in Neural Information Processing Systems 4, Part IX* (Morgan Kaufmann, San Mateo, CA), pp. 619–626.
29. J. E. Hoover and P. L. Strick 1993, "Multiple output channels in the basal ganglia," *Science* **259**, 819–821.
30. J. C. Houk 1992, "Learning in modular networks," in *Proceedings of the Seventh Yale Workshop on Adaptive and Learning Systems* (Center for Systems Science, New Haven CO), pp. 80–84.
31. J. C. Houk, J. L. Adams and A. G. Barto 1995, "A model of how the basal ganglia generate and use neural signals that predict reinforcement," in *Models of Information Processing in the Basal Ganglia*, eds. J. C. Houk, J. L. Davis and D. G. Beiser (MIT Press, Cambridge, MA), pp. 249–270.
32. J. C. Houk, J. L. Davis and D. G. Beiser 1995, *Models of Information Processing in the Basal Ganglia* (MIT Press, Cambridge, MA).
33. Y. K. Hwang and N. Ahuja 1992, "Gross motion planning — a survey," *ACM Computing Surveys* **24**(3), 219–291.
34. D. Joel and I. Weiner 1997, "The connections of the primate subthalamic nucleus: Indirect pathways and the open-interconnected scheme of basal ganglia-thalamocortical circuitry," *Brain Res. Rev.* **23**, 62–78.

35. Zs. Kalmár, Cs. Szepesvári and A. Lőrincz 1998, "Module-based reinforcement learning: Experiments with a real robot," *Machine Learning* **31**, 55–85.
36. J. M. Kemp and T. P. S. Powell 1970, "The corticostriate projection in the monkey," *Brain* **93**, 525–546.
37. T. Kohonen 1982, "Self-organized formation of topologically correct feature maps," *Biol. Cyber.* **43**, 59–69.
38. M. Kositsky 1998, *Motor Learning and Skill Acquisition by Sequences of Elementary Actions*, Ph.D. thesis, Department of Applied Mathematics and Computer Science, The Weizmann Institute of Science, Israel. http://www-anw.cs.umass.edu/~kositsky/phdThesis/phdThesis.html.
39. M. Kositsky, T. Flash and S. Ullman 1998, "A cluster memory model for learning sequential activities," *NIPS Workshop on Movement Primitives: Building Blocks for Learning Motor Control*, Breckenridge, CO. http://www-anw.cs.umass.edu/~kositsky/Publications.html.
40. H. Kunzle 1977, "Projections from the primary somatosensory cortex to basal ganglia and thalamus in the monkey," *Experimental Brain Research* **30**, 481–492.
41. S. L. Liles and B. V. Updyke 1985, "Projection of the digit and wrist area of precentral gyrus to the putamen: Relation between topography and physiological properties of neurons in the putamen," *Brain Research* **339**, 245–255.
42. K. Lorenz 1970 *Studies in Animal and Human Behavior*, Vol. I–II (Harvard University Press, Cambridge, MA).
43. A. Lőrincz 1997, "Neurocontrol III: Temporal differencing models of the basal ganglia-thalamocortical loops," *Neural Network World*, pp. 43–72.
44. A. Lőrincz 1997, "Static and dynamic state feedback control model of basal ganglia-thalamocortical loops," *Int. J. Neural Systems* **8**, 339–357.
45. A. Lőrincz 1998, "Basal ganglia perform differencing between 'desired' and 'experienced' parameters," *Computational Neuroscience, Trends in Research 1997* (Plenum Press, New York), pp. 77–82.
46. J. T. Lurito, T. Georgakopoulos and A. P. Georgopoulos 1991, "Cognitive spatial-motor processes," *Experimental Brain Research* **87**, 562–580.
47. R. Malach and A. M. Graybiel 1986, "Mosaic architecture of the somatic sensory-recipient sector of the cat's striatum," *Journal of Neuroscience* **12**, 3436–3458.
48. M. J. Mataric 1997, "Behavior-based control: Examples from navigation, learning, and group behavior," *J. Exp. Theo. Artificial Intelligence* **9**, 323–336.
49. P. R. Montague, P. Dayan and T. J. Sejnowski 1994, "Foraging in and uncertain environment using predictive Hebbian learning," in *Neural Information Processing Systems 6*, eds. J. D. Cowan, G. Tesauro and J. Alspector, (Morgan Kaufmann, San Francisco CA), pp. 598–605.
50. A. Nini, A. Feingold, H. Slovin and H. Bergman 1995, "Neurons in the globus pallidus do not show correlated activity in the normal monkey, but phase-locked oscillations appear in the MPTP model of Parkinsonism," *Journal of Neurophysiology* **74**, 1800–1805.
51. A. Parent 1990, "Extrinsic connections of the basal ganglia," *Trends in Neuroscience* **13**, 254–258.
52. R. Parr and S. Russell 1998, *Advances of Neural Information Processing Systems*, Vol. 10, chapter Reinforcement Learning with Hierarchies of Machines, ed. M. I. Jordan, M. J. Kearns and S. A. Solla (MIT Press, Cambridge), pp. 1043–1049.
53. G. Percheron and M. Filion 1991, "Parallel processing in the basal ganglia: Up to a point," *Trends in Neuroscience* **14**, 55–56.
54. G. Percheron, G. Yelnik and C. Francois 1984, *The Basal Ganglia* (Plenum Press, New York), pp. 87–105.
55. G. Percheron, G. Yelnik and C. Francois 1984, "A Golgi analysis of the primate globus pallidus: III. Spatial organization of the striato-pallidal complex," *Journal of Neural Computation* **227**, 214–227.
56. D. Precup and R. S. Sutton 1998, *Advances of Neural Information Processing Systems*, Vol. 10, chapter Multi-Time Models for Temporally Abstract Planning, ed. M. I. Jordan, M. J. Kearns and S. A. Solla (MIT Press, Cambridge), pp. 1050–1056.
57. S. Sastry and M. Bodson 1989, *Adaptive Control — Stability, Convergence and Robustness* (Prentice Hall, Englewood Cliffs, New Jersey).
58. W. Schultz, P. Dayan and R. R. Montague 1997, "A neural substrate of prediction and reward," *Science* **275**, 1593–1599.
59. L. D. Selemon and P. S. Goldman-Rakic 1990, "Topographic intermingling of striatonigral and striatopallidal neurons in rhesus monkeys," *Journal of Comp. Neurol.* **297**, 359–376.
60. M. Steriade and R. R. Llinas 1988, "The functional states of the thalamus and the associated neuronal interplay," *Physiological Review* **68**, 649–742.
61. C. Szepesvári, S. Cimmer and A. Lőrincz 1997, "Neurocontroller using dynamic state feedback for compensatory control," *Neural Networks* **10**, 1691–1708.
62. C. Szepesvári and A. Lőrincz 1997, "Robust control using inverse dynamics neurocontrollers nonlinear analysis," *Methods and Applications* **30**, 1669–1676.
63. C. Szepesvári and A. Lőrincz 1998, "An integrated architecture for motion-control and path-planning," *Journal of Robotic Systems* **15**, 1–15.
64. Cs. Szepesvári, L. Balázs and A. Lőrincz 1994, "Topology learning solved by extended objects: a neural network model," *Neural Computation* **6**, 439–456.
65. Cs. Szepesvári and A. Lőrincz 1997, "Approximate inverse-dynamics based robust control using

static and dynamic feedback," *Applications of Neural Adaptive Control Theory*, Vol. II (World Scientific, Singapore), pp. 151–179.
66. C. Szepesvári and A. Lőrincz 1996, "Approximate geometry representation and sensory fusion," *Neurocomputing* **12**, 267–287.
67. J. Tanji and K. Kurata 1985, "Contrasting neuronal activity in supplementary and precentral motor cortex of monkeys. I. Responses to instruction determining motor responses to forthcoming signals of different modalities," *Journal of Neurophysiology* **53**, 129–141.
68. J. Tanji, K. Tanuguchi and T. Saga 1980, "Supplementary motor area: Neuronal response to motor instructions," *Journal of Neurophysiology* **43**, 60–68.
69. E. Thelen and D. Corbetta 1994, "Exploration and selection in the early acquisition of skill," *International Review of Neurobiology* **37**, 75–102.
70. J. T. Weber and T. C. T. Yin 1984, "Subcortical projections of the inferior parietal cortex (area 7) in the stump-tailed monkey," *J. Comp. Neurol.* **224**, 206–230.
71. C. J. Wilson 1990, "Basal ganglia," in *The Synaptic Organization of the Brain* (Oxford University Press, New York), pp. 279–316.
72. C. J. Wilson 1992, "Dendritic morphology, inward rectification, and the functional properties of neostriatal neurons," in *Single Neuron Computation* (Academic Press, Boston), pp. 141–171.
73. D. Zipser and R. A. Andersen 1988, "A backpropagation programmed network that simulates response properties of a subset of posterior parietal neurons," *Nature* **331**, 679–684.