Biol Cybern (2005) 92: 427–437
DOI 10.1007/s00422-005-0569-8

ORIGINAL PAPER

Biological Cybernetics

James C. Houk

Agents of the mind

Received: 12 October 2004 / Accepted: 18 March 2005 / Published online: 24 May 2005

Abstract The higher order circuitry of the brain is comprised of a large-scale network of cerebral cortical areas that are individually regulated by loops through subcortical structures, particularly through the basal ganglia and cerebellum. These subcortical loops have powerful computational architectures. Using, as an example, the relatively well-understood processing that occurs in the cortical/basal ganglionic/cerebellar distributed processing module that generates voluntary motor commands, I postulate that a network of analogous agents is an appropriate framework for exploring the dynamics of the mind.

1 Introduction

Agent-based modeling has been a productive approach for illuminating emergent properties of complex systems (Riolo et al. 2001; Wilensky and Reisman 1998). The mind clearly is a complex system, and concepts about the mind's agents have been elusive to identify (Minsky 1986). Some investigators have considered cortical columns to be nodes of consciousness (Crick and Koch 2003). Others have focused more microscopically on single neurons, or even the individual synapses that transmit messages between neurons (Debiec et al. 2002). Instead, Freeman (2005) has focused more macroscopically on wave packets, each of which occupies a substantial fraction of each hemisphere. Here, I suggest that a more opportune choice for answering many questions about how the mind thinks and controls action may involve focusing on networks of anatomically defined assemblies that have been called distributed processing modules (Houk 2001). These **DPMs** are comprised of one area of the cerebral cortex together with its topographically specific subcortical loops through the basal ganglia and through the cerebellum, and, in some cases, through other subcortical structures.

What are the computational rules that DPMs, as agents of the mind, might obey, and what are they computing? Addressing this question will require blending quite diverse neurophysiological and neuroanatomical data, and abstraction of underlying computational principles. To facilitate this synthesis, I will begin with a short summary of the abstract properties that I will end up positing for DPM agents. Next, in Sect. 3, I will summarize the unique cellular properties of the principal neurons in the cerebellar cortex and in the striatum of the basal ganglia. The special computational properties of these neurons are one key reason for focusing on loops through cerebellum and basal ganglia in my choice of agents. Section 4 summarizes a second key reason for focusing on loops through cerebellum and basal ganglia, namely the consistent neuronal architecture characterizing each of these loops. This consistency serves to define a reasonable model for the architecture of the mind. Section 5 discusses signal-processing operations at the level of systems neuroscience. Since we know much more about the loops that regulate movement than we know about the loops that regulate thinking, I will focus initially on the signal-processing operations that underlie the generation of voluntary motor commands. Then, I will appeal to analogy to facilitate discussion of the signal-processing operations that may underlie thinking.

2 Abstract signal-processing operations Posited for each agent

The signal-processing operations posited for each DPM are illustrated diagrammatically in Fig. 1. The final result of all the computations in a given module will be a spatiotemporal pattern of activity in the module's set of output neurons. At any given moment, pattern formation in this output vector may or may not be initiated by spatiotemporal elements in one or several of the input vectors from other areas of cerebral cortex. Embodiment of potential initiation elements

J. C. Houk
Northwestern University, Feinberg School of Medicine,
303 E. Chicago Ave., M211 Chicago, IL 60611, USA
Tel.:+1-312-5038219
Fax.:+1-312-5035101
E-mail: j-houk@northwestern.edu

as a vectorial output from a population of cortical neurons is controlled by the set of facilitatory and suppressive outputs of the pattern classification operations in the loop through basal ganglia. This vectorial output is in response to the two inputs to the pattern classification operation – a vectorial input from the same cortical area and a reward propensity conveyed by dopamine neurons. The latter is currently thought to function as a scalar input, although the anatomy of the projection allows the possibility of it being a vector. In any case, the reward propensity input modulates pattern classification on two time scales. An immediate modulation is a multiplicative input that allows pattern classification to function in a manner analogous to a motivationally-modulated decision process. In contrast, the long-term modulation is a consolidation of synaptic strengths that occurs incrementally on a trial-by-trial basis. This pattern classification operation effectively decides on the relative correctness of alternative thoughts or actions based on its cortical input vector, and the reward propensity input modulates this result by motivational factors. Note that decisions can either be facilitated by the direct pathway (closed arrow) or suppressed by the indirect pathway (open arrow) through the basal ganglia. The embodiment step allows the results of this computation to select appropriate sets of cortical input so as to allow them to initiate pattern formation, or, if the result of pattern classification is strong enough, to actually initiate pattern formation even without receiving appropriate cortical input.

If pattern formation is initiated, the loop through the cerebellum starts to amplify and refine the spatiotemporal pattern in the cortical output vector. Regenerative recurrent feedback amplifies the intensity, duration and spatial extent of the output vector, whereas the refinement operation restrains the amplification process and sculpts it into a refined output vector. This output vector is transmitted to other areas of the cerebral cortex to be used in their modular operations. After the output vector has been used by these other DPMs, the refinement operation inhibits the amplification process to create a relatively quiescent output pattern in that DPM. Ultimately, the entire network of activated DPMs gives rise either to a skilled action or to the conclusion of a thought process. If an action results, climbing fibers detect sensorimotor errors providing delayed feedback to the refinement process in the cerebellar cortex. Traces of the computations that led to an erroneous output vector permit the delayed error information to guide adjustments in synaptic strength in a direction that progressively minimizes the chance of future errors. Errors in a purely thinking process, without an action, are probably also sensed by climbing fibers, but the physiology of such a process is quite unclear at present.

3 Special computational properties of the principal neurons in the cerebellum and basal ganglia

The neuronal architecture of the cerebellum is quite exceptional (Houk and Mugnaini 2002). As illustrated in Fig. 2, its principal neurons, the Purkinje cells, receive an order of magnitude larger number of inputs than do any other type of neuron in the brain, and its granule cells are more numerous than all remaining brain cells in combination (Tyrrell and Willshaw 1992). The incredible number of spinous synapses that axons of the granule cells (the parallel fibers in Fig. 2, right) form on Purkinje cells each exhibit a special form of synaptic plasticity — a type of long-term depression (LTD) (Ito 2001) that could give rise to a learning rule that is highly effective at reducing errors in network performance (Houk and Alford 1996; Barto et al. 1999). This powerful computational architecture of the cerebellum is analogous to that of a perceptron (Albus 1981).

The neuronal architecture of the basal ganglia is also exceptional (Houk 2001; Houk and Wise 1995). Its principal neurons, striatal medium spiny neurons (Fig. 3), receive a highly convergent input from 10 to 20 thousand different cortical neurons (Kincaid et al. 1998). The excitatory synapses made by the cortical inputs display a special form of long-term potentiation (LTP) that is modulated by input from dopamine neurons located in the substantia nigra pars compacta. The modulatory action occurs on two time scales, an immediate one (Nicola et al. 2000) that is implicated in motivational effects (Kawagoe et al. 1998) and a long-term one (Charpier and Deniau 1997) that appears to consolidate LTP (Houk et al. 1995; Schultz et al. 1995). Dopamine neurons signal predictions of upcoming rewards (Schultz 1998) and appear to mediate a powerful evaluation-based form of learning (Barto 1995).

The signal-processing operations that occur in these two subcortical loops are indeed powerful computationally, which is due in large part to the capacity for spiny neurons and Purkinje cells to classify complex spatial patterns in their arrays of cortical input in an efficient, adaptive manner (Houk and Wise 1995). The elaborate array of inhibitory collaterals formed by spiny neurons (colored red in Fig. 3) make pattern classification in the basal ganglia competitive (Houk 2005). Because the results of this classification project back to the same area of cortex, the cortical-basal ganglionic module has been considered (Beiser and Houk 1998) capable of implementing the powerful mathematical operation of "recursion" (Minsky 1963). In other words, the module uses the results of its pattern classification operation to update the cortical pattern that provides its own input.

A second specialization is that Purkinje cells and spiny neurons are provided with training information that promotes excellent "credit assignment" for learning. Credit assignment is the problem of getting the right training information to the right synapses (spatial credit assignment) at the right time (temporal credit assignment) in order to promote learning. Spatial credit assignment is promoted in the cerebellum by the precise alignment of its climbing fibers, which allows error signals to register with the particular Purkinje cells that can actually correct the errors (Houk and Barto 1992). Temporal credit assignment is promoted by a learning rule that compensates for time delays in the receipt of training information (Barto et al. 1999; Houk and Alford 1996; Houk and

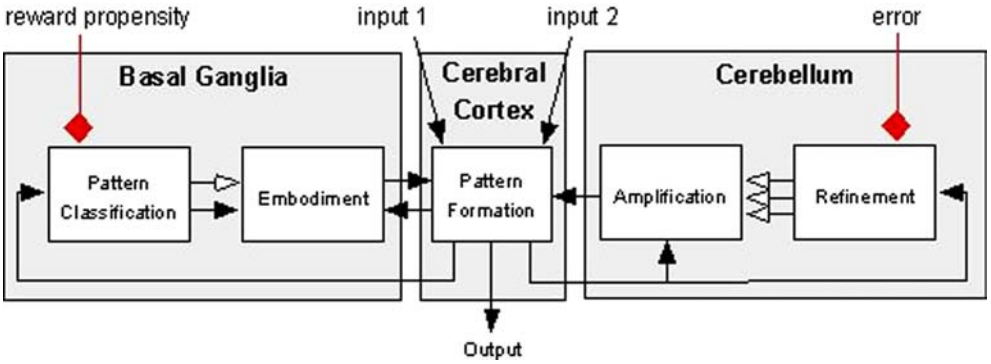

Fig. 1 The signal-processing operations posited for a distributed processing module (DPM). Net excitatory pathways are shown with closed arrows, net inhibitory pathways are shown with *open arrows*, and the *red diamonds* signify modulatory and training inputs

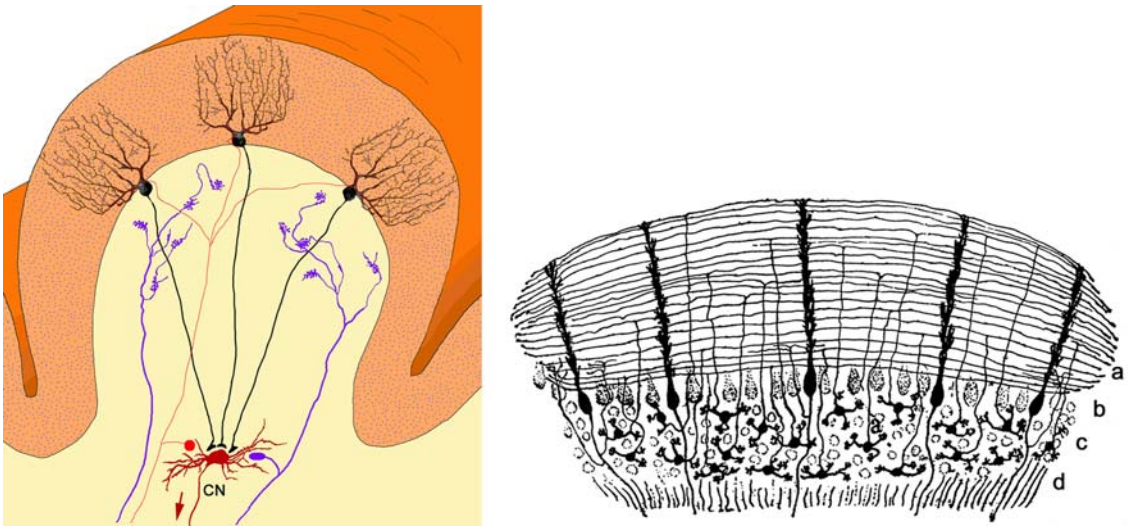

Fig. 2 The basic microcircuitry in the cerebellar cortex, illustrated in orthogonal planes. The right panel shows, in the transverse plane, a Cajal drawing of the cerebellar circuitry, and the left panel shows, in the parasagital plane, Mugnaini's Cajal-style drawing (Houk and Mugnaini 2002). In Cajal's drawing, *a* molecular layer *b* Purkinje cell layer *c* granular layer *d* white matter. In Mugnaini's drawing, *black* Pukinje cells, *purple* mossy fibers, *red* climbing fibers (*on top*) and cerebellar nucleus (*below*). The Purkinje cells have a large fan-shaped dendritic tree, which look radically different in the two planes; each Purkinje cell receives 200,000 parallel fiber inputs. The parallel fibers are the small axons of the enormous number of tiny granule cells in the granular layer of the cerebellum

Mugnaini 2002). In the striatum, spatial credit assignment could be promoted by the relatively small axonal fields of dopamine neurons (Wilson 1990), even though present data does not support this. Temporal credit assignment is clearly promoted by the remarkable ability of dopamine neurons to predict reward (Schultz 1998), and then to predict the predictions of reward, which is another example of recursion (Houk et al. 1995). Temporal credit assignment is critical because it allows the reinforcement of synaptic activities that promote those behaviors that are most likely to yield rewards well into the future.

The bistability that is present in the dendrites of Purkinje cells (Genet and Delord 2002) and in spiny neurons (Gruber et al. 2003) is a third feature that is computationally powerful. Bistable operations involve sharp thresholds between on- and off-states of their outputs. This is advantageous because it creates clean, state-dependent decision surfaces for distinguishing between appropriate and inappropriate patterns of convergent input (Gruber et al. 2003). Bistability should help Purkinje cells classify input patterns that are specifically correlated with good performance, and should help spiny neurons classify input patterns that are specifically correlated with reward likelihood.

Synaptic plasticity in cortical neurons appears to be guided mainly by local correlations between presynaptic and postsynaptic activity (LTP) (Bliss and Collingridge 1993). If a cortical neuron is repeatedly forced by its cerebellar and basal ganglionic inputs to fire in a particular manner, it should learn, through practice, intracortical associations capable of causing the neuron to respond in a rapid, direct manner whenever the same circumstances are repeated (Hua and Houk 1997).

4 Architecture of networks of DPMs

A good starting point for the development of theories about the function of the many loops through cerebellum and basal ganglia comes from an understanding of the connectivity of

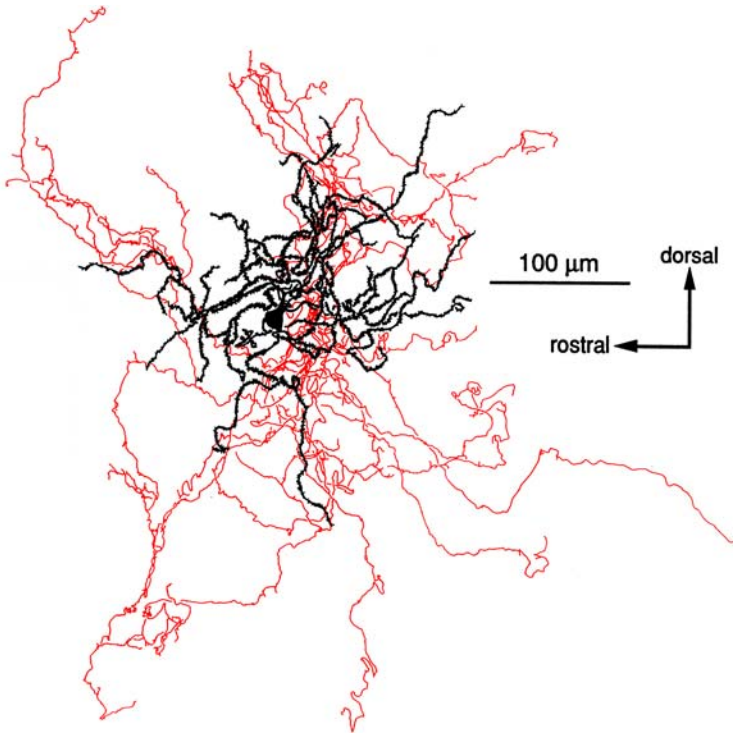

Fig. 3 A picture of a spiny neuron (Tepper et al. 2004), the principal neuron type in the basal ganglia loop. Spiny neurons are located in the striatum, which receives highly convergent input from the cerebral cortex. They have a stellate array of dendrites (*shown in black*) with about 20,000 spines that mark the synapses made by different inputs from a given area of the cerebral cortex. Spiny neurons learn to classify different spatial patterns present in this input. Each spiny neuron also has an elaborate array of collaterals (*shown in red*) that form inhibitory synapses on nearby spiny neurons, which enhance pattern classification through competition

cortical–subcortical loops that is being elaborated by Strick and his colleagues (Middleton and Strick 1997; Dum and Strick 2002). For each area of frontal cortex that has been investigated via retrograde transneuronal transport of viruses, it has been possible to identify a unique channel through basal ganglia that provides input to that cortical area via thalamus, and for most of these same cortical areas, a unique channel of cerebellar input has also been identified. On the afferent side of these loops, the most prominent input to a given basal ganglia channel derives from the same area of cerebral cortex that is targeted by the channel's output projections (Inase et al. 1996; Kelly and Strick 2003; Strick et al. 1995), and a similar organizational scheme appears to dominate each cerebellar channel (Kelly and Strick 2004; Schmahmann and Pandya 1997). These observations suggest that each area of cortex is innervated by a relatively private recurrent loop through the basal ganglia, and is often also innervated by a relatively private recurrent loop through the cerebellum.

Figure 4 summarizes an overview of these organizational features of the brain's signal-processing networks. Three areas of cerebral cortex are generically labeled B, C, and D, and I specifically include the primary motor cortex (M1). If these cortical areas are functionally related, we can anticipate reciprocally organized corticocortical connections between many of them (Felleman and Van Essen 1991; Goldman-Rakic 1988). These cortical-cortical linkages are shown by four bidirectional green arrows that reciprocally connect areas M1, B and C, and C with D, in Fig. 4. We can further anticipate that each area of cortex is regulated by a recurrent loop through the basal ganglia (bidirectional arrows that are red to reflect prominent inhibition), and frequently by a second subcortical loop passing through the cerebellum (bidirectional arrows that are blue to reflect the combination of excitation in the loop through cerebellar nuclei and inhibition in the loop through the cerebellar cortex).

A given area of cerebral cortex, together with its recurrent channels through basal ganglia and cerebellum, forms the entity that I refer to here as a distributed processing module (DPM). Note that the term distributed is being used in an anatomical sense; a given area of cortex has long-distance topographic connections with discrete regions of subcortical structures. The neocerebrum is comprised of a substantial number of these distributed modules, which communicate with each other in two ways. The predominant mode of intercommunication is by way of the cortical–cortical connections that have already been mentioned and are indicated by green arrows in Fig. 4. In addition, some functionally related areas of cerebral cortex project in a unidirectional manner as inputs to the channels through basal ganglia (Graybiel 1991; Inase et al. 1996; Yeterian and Van Hoesen 1978) and to the channels through cerebellum (Brodal and Bjaalie 1997; Schmahmann and Pandya 1997). Since the latter connections

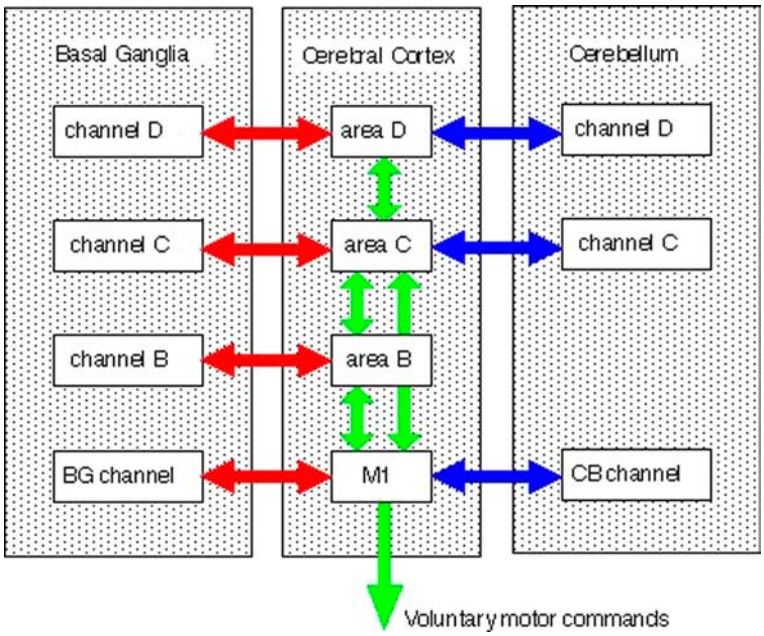

Fig. 4 Distributed modular architecture (Houk 2001). The diagram shows four areas of cerebral cortex, labeled M1, B, C and D, their subcortical loops through basal ganglia (*red*) and through cerebellum (*blue*), and their cortical-cortical pathways (*green*). A given area of cortex together with its subcortical loop(s) forms a *distributed processing module* and the different DPMs communicate with each other through the cortical–cortical connections. The green arrow connecting M1 (the primary motor cortex) with area C goes underneath area B

tend to duplicate the communications provided by cortical-cortical linkages, they are left out of Fig. 4 for convenience.

In Houk 2001, I focus on two examples of a DPM: the M1–DPM which regulates voluntary movement commands in M1, and a cognitive module that regulates working memories of sensory events in Brodman Area 46 of the prefrontal cortex. Many other areas of the cerebral cortex are similarly organized and presumed, by analogy, to utilize the same signal-processing mechanisms. I estimate that there may be on the order of a hundred DPMs in the human brain, comprising a much larger network than the tiny example illustrated in Fig. 4. Different behaviors and thought processes engage diverse subsets of this larger network of potential DPMs, and their individual operations need to be coordinated in some manner. Corticocortical, corticostriatal, and corticopontocerebellar projections each offer possible mechanisms for some coordination among modules. However, global neuromodulatory mechanisms that are mediated through monoamine release or through a cholinergic mechanism may contribute more importantly to the shaping of the collective behavior of the brain (Houk and Wise 1995; Doya 2002; Gruber et al. 2003).

5 Signal-processing operations in DPMs

I begin here with the M1–DPM which links the primary motor cortex (M1) with its loops through the basal ganglia and cerebellum. The voluntary motor commands that control limb movements actually originate in both M1 and in the magnocellular division of the red nucleus (RNm). Generally speaking, voluntary commands are bursts of discharge that preceed movement by about 100 ms; burst frequency codes movement velocity and burst duration codes movement duration (Gibson et al. 1985). These relationships to velocity probably exist because velocity normally correlates with the degree of muscle activation—in actuality, the muscle activation variable more accurately represents how voluntary commands are coded (Miller and Houk 1995). The specific subcortical pathways of the M1–DPM and its relationship to the RNm are summarized in the bottom part of Fig. 5, using closed arrows to mark routes that are predominantly excitatory and open arrows to mark routes that are predominantly inhibitory (Hoover and Strick 1999).

There are two principal signal-processing operations involved in the generation of voluntary commands. First, there is some mechanism for the selection and initiation of a command (Requin et al. 1988; Sakai et al. 2000). Second, the fact that voluntary commands are graded so as to reflect the direction, speed, and size of a movement (Gibson et al. 1985; Lamarre and Spidalieri 1983) indicates that there is a mechanism for regulating command intensity and duration. The prominence of recurrent connections from M1 through both basal ganglia and cerebellum back to M1 favors their collective participation in these operations, along with intracortical processing. The classical symptom in patients with Parkinson's disease, suffering from a disorder of the basal ganglia, is a difficulty in initiating movements, particularly when there is no strong sensory cue for triggering them (Benecke et al. 1987). This deficit has been traced to abnormally

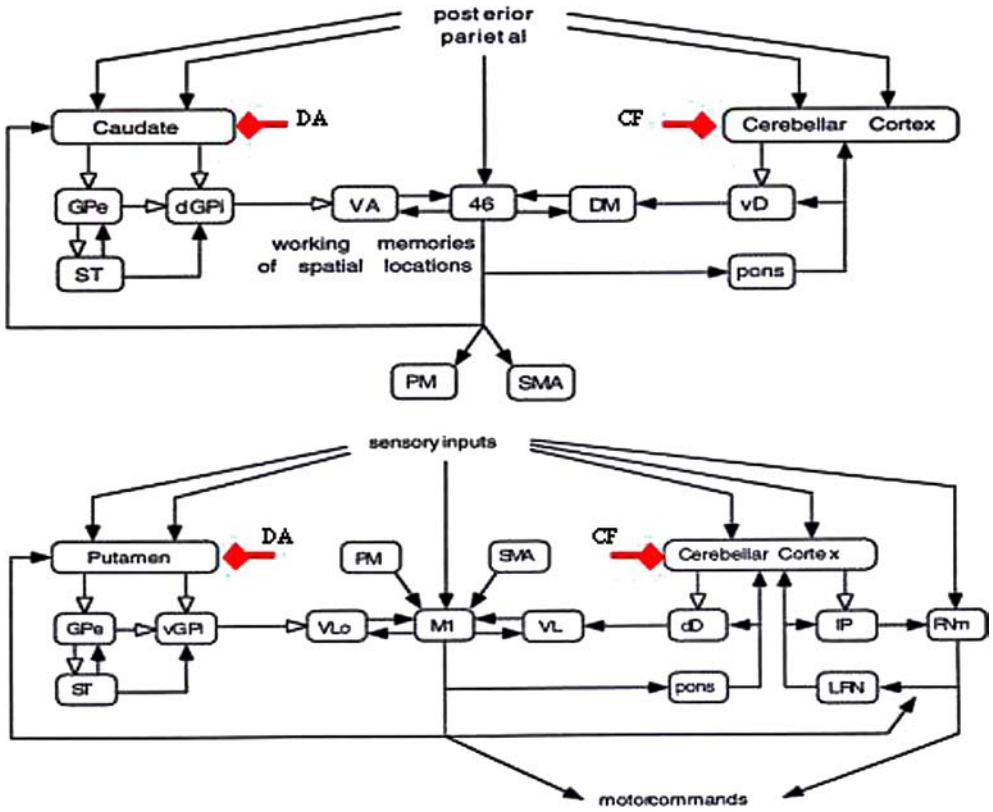

Fig. 5 Detailed architecture of the M1–DPM (*bottom*) and the BA46–DPM (*top*); note their striking similarities. Predominately excitatory pathways are shown with closed arrows, predominately inhibitory pathways are shown with open arrows, and the red diamonds signify training and modulatory inputs (*DA* dopamine; *CF* climbing fiber). **Bottom** For the DPM subserving primary motor cortex (*M1*), the subcortical channel through the basal ganglia includes a zone of the putamen, a ventral zone of the internal globus pallidus (*vGPi*), a zone of external globus pallidus (*GPe*) and subthalamus (*ST*). This channel loops back to M1 through ventral lateral thalamus pars oralis (*VLo*). The subcortical channel through the cerebellum includes a portion of the pontine nucleus (*pons*), a dorsal zone of the dentate nucleus (*dD*) and a zone of cerebellar cortex. This channel loops back to M1 through ventral lateral thalamus (*VL*). The magnocellular division of the red nucleus (*RNm*) also contributes voluntary motor commands using a loop from lateral reticular nucleus (*LRN*) and nucleus interpositus (IP). The other cortical areas are premotor (*PM*) and supplementary motor area (*SMA*). **Top** The architecture of the distributed processing module subserving BA46 is virtually identical. The subcortical channel through the basal ganglia goes through a zone of caudate instead of putamen, and a dorsal zone of the internal globus pallidus (*dGPi*) instead of vGPi. This channel loops back to BA46 through ventral anterior thalamus (*VA*) instead of VLo. The subcortical channel through the cerebellum via the pons passes through a ventral zone of the dentate nucleus (*vD*) instead of dD, and it passes through a different part of the cerebellar cortex. This channel loops back to BA46 through dorsomedial thalamus (*DM*) instead of VL

high discharge in output cells of the basal ganglia (Wichmann et al. 1999), the neurons in the ventral zone of the internal globus pallidus (vGPi; see Fig. 5). Excessive inhibitory input to pars oralis of the ventrolateral nucleus of the thalamus (VLo) from these neurons seems to impede the initiation of the M1 bursts that command movement segments. This suggests that the M1 cortical-basal ganglionic module may be especially important in regulating command embodiment (Fig. 1). In contrast, a classical symptom in patients with damage to the cerebellum is dysmetria, a failure to regulate the direction, velocity, and endpoint of movement (Holmes 1939). This suggests that the M1 cortical-cerebellar module may be especially important in regulating the intensity and duration of voluntary commands. The magnocellular division of the RNm is also recurrently connected with the cerebellum and operates through a mechanism analogous to that for M1 (Miller and Houk 1995).

The upper part of Fig. 5 summarizes the specific subcortical pathways of the DPM subserving Brodman Area 46 of the prefrontal cortex, the BA46-DPM. Note that its basal ganglia loop goes through a dorsal region in the globus pallidus (dGPi), as opposed to a ventral region (vGPi) that participates in the M1–DPM (Fig. 5, top versus bottom). The BA46–DPM loops back to cortex via the VA thalamus, as opposed to the VLo region that subserves the M1–DPM. Similarly, the loops through cerebellum are topographically distinct. The BA46–DPM targets a ventral region of the dentate nucleus (vD) whereas the M1–DPM targets a dorsal region of dentate (dD), and the pathways through thalamus are also distinct. The reader is referred to Kelly and Strick (2003, 2004) for a more complete description of the topography of these loops.

Single unit studies in BA46 have revealed the three functional components of response illustrated in Fig. 6 (Fuster 1997; Goldman-Rakic et al. 1990). (Individual neurons often

show mixtures of two of these components, presumably mediated by intracortical processing.) The monkeys were trained on a working memory task in which a visual cue (Instruction) was briefly presented, and, after a memory period lasting a few seconds, a Go signal was given. The monkey had to remember the location of the visual cue and make a movement to that location. The abbreviated descriptions in the following three paragraphs are supplemented and referenced in Houk (2001).

The Cue-Related component of single unit discharge appears at a fixed latency after cue presentation and is like a sensory response, except that its presence and magnitude depends on motivation. This dependence suggests that it is mediated by transmission through the basal ganglia where the motivational component derives from dopamine-dependent neuromodulation (DA-labeled diamonds in Fig. 5). While dopamine neurons also innervate the cortical targets of the loop, that projection is an order of magnitude less prominent than is the innervation of the striatum (the striatum includes both caudate and putamen). The reward propensity signal in Fig. 1 is an abstract representation of dopamine neuromodulation.

The sustained component of discharge spans the memory period (Fig. 6). When it is blocked, the monkey performs at chance, leading to the conclusion that this sustained discharge is the neural correlate of a working memory. Except for the fact that the duration is longer than is a motor command, the sustained component resembles motor commands recorded in M1. This is consistent with my hypothesis that it is mainly mediated by positive feedback in the loop through vD of the cerebellum, in analogy with the role of positive feedback in the M1–DPM (Holdefer et al. 2000; Houk 2005; Houk et al. 1993; Hua and Houk 1997). There is also evidence (reviewed in Houk (1997)) that this working memory is shaped by inhibitory input from the cerebellar cortex, the Refinement operation in Fig. 1.

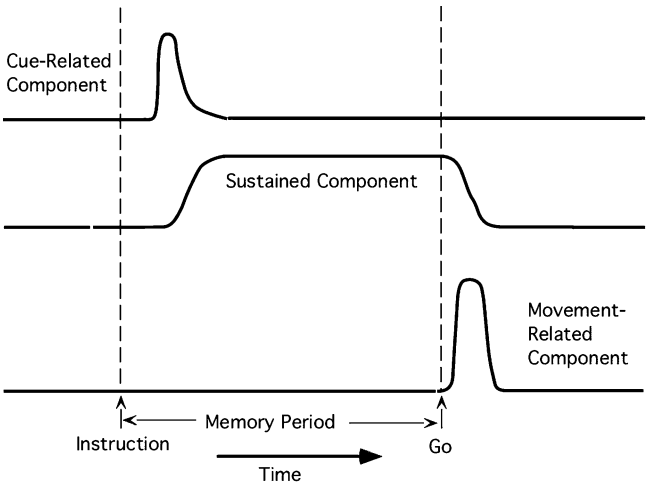

Fig. 6 Components of BA46 discharge in a working memory task. A brief presentation of the instruction elicits cue-related components, after which sustained components (the neural correlates of a working memory) begin and continue until the Go instruction is delivered, and that is followed by movement related components

The movement-related component in Fig. 6 closely resembles motor commands that are recorded from primary motor cortex. Most likely that's where these components derive from, being relayed to BA46 via corticocortical projections. Most neurophysiologists believe that this signal is used to turn off the sustained memory-related component. This makes sense since the working memory has finished being used to help generate the motor command.

Time scales of operation are different in different DPMs, each of which subserves different manifestations of the temporal operations of the brain in regulating diverse behaviors (Fuster 2001). Generally speaking, the duration of an action is shorter than the duration of a thought. For example, if Fig. 6 were being used to illustrate a voluntary motor command in M1, the sustained component would be a train of action potentials having a duration of a few hundred milliseconds, as opposed to the several second duration of the memory period in a typical working memory task. As another example, consider how a working memory for a more complicated action is translated into a sequence of movements (Tanji 2001). Neurons in the supplementary motor area (SMA in Fig. 5) generate preparatory signals that begin when one movement command is completed and continue until the next movement command is initiated. The duration of these preparatory signals is typically shorter than the working me mory of the sequence and longer than an individual motor command. Thus, as one moves up the hierarchy from motor cortex to premotor areas to prefrontal cortex, time scales of operation tend to get slower, although there can be exceptions to this rule.

The time scales of operation in the subcortical loops are also interesting to consider. Sensory cues for eliciting a motor command can be quite brief, say 10 ms. Assuming that a particular sensory cue is embodied by the loop through the basal ganglia, the motor cortex then initiates the amplification process in the loop through the cerebellar nucleus. Although conduction times around this loop amount only to a few milliseconds, the amplification process proceeds more slowly. Amplification depends on temporal summation of two types of excitatory synaptic potentials, a fast AMPA-mediated one and a slow NMDA-mediated one. The latter is a particularly important driving force for the sustained component of the motor command (Jiang et al. 2002). Furthermore, the loop through the cerebellar nucleus is actually a whole array of microscopic loops (Hua and Houk 1997). Positive feedback in one microscopic loop has to spread to many additional loops in order to recruit the large population of motor commands that is required to produce a movement (Houk et al. 1993). This occurs over a time course of 10's of ms, and the amplified command is then sustained for one hundred to several hundred milliseconds. The Purkinje cells learn to predict when to turn off the sustained activity in order for the movement to terminate accurately on the target (Barto et al. 1999).

6 Discussion

In this article, I posit that subcortical loops through the basal ganglia are specialized for the selection and initiation (embodiment) of patterns of population activity that encode incomplete thoughts, which then need to be amplified and refined by loops through the cerebellum. The analogy with the motor system is the initiation of an incomplete motor command in M1 by its loop through the basal ganglia and the subsequent amplification and refinement of M1 population activity by its loop through the cerebellum. The populations of neural activity representing thoughts need to be amplified and refined in order to represent accurate thoughts, just as the population of M1 neurons needs to be amplified and refined in order to represent accurate movement commands. Both the thought selection operations and the amplification and refinement operations may proceed in cycles of processing in their loops through basal ganglia and cerebellum. This analogy helps one contemplate and interpret the sustained memory-related discharge (and fMRI activity) that is observed in the prefrontal cortex in working memory tasks (Fuster 1997; Goldman-Rakic et al. 1990). The sustained activity terminates after the working memory is utilized, analogous to the termination of a movement command after it has directed the limb toward its desired endpoint. Analogy may also help one interpret the complex responses that have been observed in parietal and temporal cortex in other cognitive and social tasks (Fuster 2001; Perrett et al. 1990).

The selection of an appropriate precursor of an incomplete thought is considered to be a difficult problem, equivalent to a search through a very large database containing many alternatives. However, instead of following a sequential course, this search is conceived as a parallel, recursive process. The signal-processing that mediates a parallel search is presumed to be competitive pattern classification in the striatum of the basal ganglia, coupled with a recursion-like operation mediated by cortical-basal ganglionic loops. The analogy with the voluntary motor system would be the selection of a tentative movement command, one that warrants consideration given the context and sensory events. There is also a need to select appropriate submovements, in order to correct an erroneous primary movement (Novak et al. 2002).

I further posit that loops through the cerebellum are specialized for the amplification, elaboration, and refinement of incomplete thoughts originally initiated by loops through the basal ganglia. These perfection processes are conceived as progressive amplifications and sculptings of spatial patterns of sustained network activity, regulated by Purkinje cells in the cerebellar hemispheres. Initial incomplete thoughts may be little more than vague hints, in which case their amplification, refinement, and ultimate perfection, could be a very difficult problem. The neuronal architecture of the cerebellar cortex is well suited for these difficult functions (Houk and Mugnaini 2002). The analogy with the voluntary motor system would be the amplification and refinement of an incomplete movement command, sculpting it into a refined command for moving the hand accurately to the desired endpoint.

DPMs are powerful computational agents – they shape the activity of any given cortical population into a spatiotemporal pattern that is useful for controlling an action or for thinking about how to control other areas of the cerebral cortex to achieve particular goals. Functional imaging has revealed how different networks of cortical and subcortical areas become active when human subjects engage in problem solving tasks such as the Tower of London (Dagher and Owen 1999), or when subjects process words (Petersen and Fiez 1993), or attempt to solve difficult peg puzzles (Kim et al. 1994), and in many other cognitive tasks.

The influential book on the application of artificial intelligence principles to human behavior published by Minsky in 1986 [The Society of Mind] needs to be translated into language that neuroscientists use to talk about the brain. Prior progress in that direction has been minimal (but see (Doya 1999; Houk and Wise 1995; Houk 2001)). The conceptual and computational models posited in this paper may provide a useful global framework for advancing our understanding of cognitive brain function, and perhaps even for rethinking about the mind-brain problem. Arrays of up to 100 DPM agents may provide a useful approach for exploring the complex dynamics of the mind.

Recursion is a powerful computational operation that Chomsky and his colleagues (Hauser et al. 2002) consider to underlie the rich, expressive and open-ended power of human language (also cf. Elman 2004). Cortical-basal ganglionic modules appear to possess the capacity for a recursion-like operation, namely an ability to use the results of its pattern classification operation to update the cortical pattern that provides its own input. Iteration also does this, but in a subtly different way. Independent of this distinction, the recursion-like operation performed by cortical-basal ganglionic loops gives these modules the capacity to deal effectively with serial order in analyzing events and in controlling behavior (Beiser and Houk 1998).

What property of mathematical recursion do cortical-basal ganglionic modules lack? Thinking of recursion as a subroutine-like capability, an important feature they lack is the ability to call themselves, putting their computation to use recursively on different time scales. This may not be a severe limitation for the DPM architecture outlined in Fig. 4. The capacity to execute the same sequence of computational steps is reproduced about a hundred times across the cerebral cortex, once in each DPM. The ability to hold a computational result in working memory while calling upon, through cortical–cortical connectivity, another DPM to perform its computation and report back, is analogous to a recursive subroutine call. As an example, consider the serial order working memory signals of the Beiser and Houk (1998) model of the BA46-DPM. Different items (and their serial order) in a sequence of k targets are encoded by sustained discharge in different subsets of neurons. Their translation into the preparatory signals that were mentioned near the end of Sect.

5 would be analogous to successive subroutine calls to the supplementary motor area (SMA). After SMA neurons encode the preparatory cue for the next target in the sequence, they send this result not only to M1 for execution, but also back to BA46 as a potential cue for terminating the working memory of that item. The system would then go on to the next item in the sequence in a recursive fashion. My colleagues and I intend to explore this hypothesis more fully in the future.

Because of the limited number of copies of DPMs, this form of recursion is not infinitely deep as it is in true recursion. Instead, it represents a limited form of recursion, which is not much different than the limits imposed by computational recursion, due to the finite memory capacity in any given computer. There has been a large expansion of the cerebral cortex, and its associated DPMs, in evolution. Crow and colleagues have made the case for an evolutionary link between the origin of language and the etiology of schizophrenia (Berlim et al. 2003). This is interesting because people with schizophrenia show a major deficit in their capacity for dealing with serial order, which has been attributed to a deficit in the cortical-basal ganglionic portion of the BA46–DPM (Fraser et al. 2004). I would further attribute this to a problem with the competitive pattern classification operation discussed earlier in this paper. The model of spiny neurons relies on inhibitory collaterals (colored red in Fig. 3) to make the classification of its cortical input vector competitive. It is interesting that the neurotransmitter for this inhibition is GABA, and there is a modified expression of the GABAb receptor in schizophrenia (Enna and Bowery 2004). Based on the above logic, this is a likely gene contributing to the multigenic inheritance of schizophrenia (Freedman et al. 2001). The former authors have identified another gene which causes altered expression of a nicotinic receptor that is prevalent in the loops through the cerebellar nucleus. Altered transmission in these loops is implicated in the cognitive dysmetria of schizophrenia (Andreasen 1999).

Agent-based modeling (Axelrod 1997) may indeed be a productive approach for illuminating emergent properties of the mind. The large expansion of the cerebral cortex and its associated DPMs in man may help to explain *universal grammar* (Hauser et al. 2002) as an emergent property of large arrays of DPMs. A DPM-based model of language needs to be developed to test this idea.

Computational brain dynamics which results in the formation of global state transitions called "wave packets" is the topic of another article in this issue (Freeman 2005). Freeman (2003) has put forth a neurobiological theory of meaning in perception that attributes these broad-spectrum aperiodic time-locked oscillations as capturing moments when subjects are engaged in repeated categorizations of input. Indeed, the occurrences of wave packets do correlate with moments of insight in problem solving tasks (Ohl et al. 2001). Many neuroscientists today (cf. Seth et al. 2004) are focusing on various high-frequency synchronizations as solutions to the binding problem, or what sometimes is called the sensory fusion problem. In essence, this is the difficulty of putting together different modalities of sensation to form holistic perceptions. The model of the mind proposed here approaches this problem in a complementary fashion. Each DPM receives input from about seven other modules that can each be processing different modalities of sensation and/or thought (Felleman and Van Essen 1991). The module's subcortical loops are given the job of putting all these features together to form a coherent output vector. A network of simultaneously active DPMs functioning in this manner could assemble an overall plan for action, which seems analogous to the formation of a holistic perception, a Gestalt.

What might explain the dramatic occurrences of wave packets across the cortical circuitry that Freeman observes? The brain's neuromodulatory systems, which have widespread projections across the brain, have been shown to cause rather dramatic transitions in the cellular properties of neurons. One example is the induction of bistability and nonlinear amplification in striatal spiny neurons that is produced by bursts of dopamine neuron discharge (Gruber et al. 2003). These nonlinear transitions taking place across a large network of simultaneously active DPMs might underlie the electroencephalographic events recorded from the surface of the brain when subjects have insight (Ohl et al. 2001). One test of this idea might be to interface a DPM model of the mind solving a complex problem with a detailed cellular-level model of its interacting cortical populations. This and other possibilities need to be explored in the future.

7 Conclusions

This paper presents a model of the architecture of the mind. I posit a set of mathematical operations that might be implemented by one cortical area in combination with its subcortical loops through the basal ganglia and through the cerebellum, comprising one DPM. The model of a cortical-basal ganglionic loop performs competitive pattern classification to enable the initiation of a tentative pattern of cortical activity that represents an incomplete thought. The model of a cortical-cerebellar loop amplifies and refines the enabled pattern in order to shape it into a mature pattern, one that represents a better thought. The model of intracortical circuitry learns to do its processing faster and more accurately through practice. Since all of the DPMs have the same neuronal architecture, each one of them should perform these same signal-processing operations on its particular set of about seven input vectors in order to form its particular output vector. That output is then shared with about seven other areas of cortex. A network of DPM agents may be an appropriate architecture for exploring the dynamics of the mind.

Acknowledgements The author is grateful to Andrew Barto, Peter Strick, Alon Fishbach, David Fraser and Jun Wang for much helpful advice, and to many other colleagues for reading and commenting on the manuscript. This work was supported by grants NS44837 and NS44383 from the National Institute of Neurological Disorders and Stroke.

References

Albus JS (1981) Brains, behavior and robotics. Byte Books Petersborough, NH

Andreasen NC (1999) A unitary model of schizophrenia: bleuler's "fragmented phrene" as schizencephaly. Arch Gen Psychiatry 56:781–787

Axelrod R (1997) Advancing the art of simulation in the social sciences. In: Conte R, Hegselmann R, Terna P (eds) Simulating Social Phenomena. Springer, Berlin Heidelberg New York, pp 21–40

Barto AG (1995) Adaptive critics and the basal ganglia. In: Houk JC, Davis JL, Beiser DG (eds) Models of Information Processing in the Basal Ganglia. MIT Press, Cambridge, pp 215–232

Barto AG, Fagg AH, Sitkoff N, Houk JC (1999) A cerebellar model of timing and prediction in the control of reaching. Neural Comput 11:565–594

Beiser DG, Houk JC (1998) Model of cortical-basal ganglionic processing: encoding the serial order of sensory events. J Neurophysiol 79:3168–3188

Benecke R, Rothwell JC, Dick JPR, Day DL, Marsden CD (1987) Disturbance of sequential movements in patients with Parkinson's disease. Brain 110:361–379

Berlim MT, Mattevi BS, Belmonte-de-Abreu P, Crow TJ (2003) The etiology of schzophrenia and the origin of language: Overrview of a theory. Compr Psychiatry 44:7–14

Bliss TVP, Collingridge GL (1993) A synaptic model of memory: long-term potentiation in the hippocampus. Nature 361:31–39

Brodal P, Bjaalie JG (1997) Salient anatomic features of the cortico-ponto-cerebellar pathway. In: Zeeuw CID, Strata P, Voogd J (eds) Progress in Brain Research. Elsevier Amsterdam, pp 227–249

Charpier S, Deniau JM (1997) In vivo activity-dependent plasticity at cortico-striatal connections: evidence for physiological long-term potentiation. Proc Natl Acad Sci 94:7036–7040

Crick F, Koch C (2003) A framework for conciousness. Nat Neurosci 6:119–126

Dagher A, Owen AM, et al (1999) Mapping the network for planning: a correlational PET activation study with the Tower of London task. Brain 122:1973–1987

Debiec J, LeDoux JE, Nadler K (2002) Cellular and systems reconsolidation in the hippocampus. Neuron 36:527–538

Doya K (2002) Metalearning and neuromodulation. Neural Networks 15:495–506

Doya K (1999) What are the computations of the cerebellum, the basal ganglia and the cerebral cortex Neural Netw 12:961–974

Dum RP, Strick PL (2002) An unfolded map of the cerebellar dentate nucleus and its projections to the cerebral cortex. J. Neurophysiol. 89:634–639

Elman JL (2004) An alternative view of the mental lexicon. Trends Cogn Sci 8:301–306

Enna SJ, Bowery NG (2004) GABAb receptor alterations as indicators of physiological and pharmacological function. Biochemi Pharmacol 68:1541–1548

Felleman DJ, Van Essen DC (1991) Distributed hierarchical processing in the primate cerebral cortex. Cerebral Cortex 1:1–47

Fraser D, Park S, Clark G, Yohanna D, Houk JC (2004) Spatial serial order processing in schizophrenia. Schizophrenia Res 70:203–213

Freedman R, Leonard S, Oliney A, Kaufmann CA, Malaspina D, Cloninger CR, Svrakic D, Faraone SV, Tsuang MT (2001) Evidence for the multigenic inheritance of schizophrenia. Am J Med Genet (Neuropsychiatric Genetics) 105:794–800

Freeman WJ (2005) Self organizing criticality in scale-free neocortical dyanamics. Biol Cybern (this issue)

Freeman WJ (2003) A neurobiological theory of meaning in perception. Part I: Information and meaning in nonconvergent and nonlocal brain dynamics. Int J Bifurcation Chaos 13:2493–2511

Fuster JM (2001) The Prefrontal cortex – an update: time is of the essence. Neruon 30:319–333

Fuster JM (1997) The prefrontal cortex and the temporal organization of behavior. In: Sakata H, Mikami A, FusterJM (eds) The association cortex - structure and function, Harwood Academic Publishers, The Netherlands, pp 15–31

Genet S, Delord B (2002) A biophysical model of nonlinear dynamics underlying plateau potentials and calcium spikes in Purkinje cell dendrites. J Neurophysiol 88:2430–2444

Gibson AR, Houk JC, Kohlerman NJ (1985) Magnocellular red nucleus activity during different types of limb movement in the macaque monkey. J Physiol [Lond] 358:527–549

Goldman-Rakic PS (1988) Topography of cognition: parallel distributed networks in primate association cortex. Ann Rev Neurosci 11:137–156

Goldman-Rakic PS, Funahashi S, Bruce CJ (1990) Neocortical memory circuits. In: The brain: cold spring harbor symposia on quantitative biology. Cold Spring Harbor Laboratory Press, New York

Graybiel AM (1991) Basal ganglia – input, neural activity, and relation to the cortex. Curr Opin Neurobiol 1:644–651

Gruber AJ, Solla SA, Surmeier DJ, Houk JC (2003) Modulation of striatal single units by expected reward: a spiny neuron model displaying dopamine-induced bistability. J Neurophysiol 90:1095–1114

Gruber AJ, Solla SA, Houk JC (2003) Dopamine induced bistability enhances signal-processing in spiny neurons. NIPS

Hauser MD, Chomsky N, Tecumseh FW (20002) The faculty of language: what it is, who has it, and how did it evolve. Science 298:1569–1579

Holdefer RN, Miller LE, Chen LL, Houk JC (2000) Functional connectivity between cerebellum and primary motor cortex in the awake monkey. J Neurophysiol 84:585–590

Holdefer RN, Houk JC, Miller LE (2005) Movement-related discharge in the cerebellar nuclei persists after local injections of GABAa antagonists. J Neurophysiol 93:35–43

Holmes G (1939) The cerebellum of man. (The Hughlings Jackson Memorial Lecture). Brain 62:1–30

Hoover JE, Strick PL (1999) The organization of cerebellar and basal ganglia outputs to Primary motor cortex as revealed by retrograde transneuronal transport of herpes simplex virus type 1. J Neuroscience 19:1446–2463

Houk JC, Fraser D, Fishbach A, Roy SA, Simo LS, Bastianen C, Fansler-Wald D, Miller LE, Reber PJ, Botvinick M (2005) Action selection in subcortical loops through basal ganglia. Proceedings of the 1st International Workshop on Modeling Natural Action Selection, Edinburgh, July 2005

Houk JC (2001) Neurophysiology of frontal-subcortical loops. In: Lichter DG, Cummings JL (eds) Frontal-subcortical circuits in psychiatry and neurology. Guilford Publications, New York, pp 92–113

Houk JC (1997) On the role of the cerebellum and basal ganglia in cognitive signal-processing. Prog Brain Res 114:543–552

Houk JC, Adams JL, Barto AG (1995) A model of how the basal ganglia generates and uses neural signals that predict reinforcement. In: Houk JC, Davis JL, Beiser DG (eds) Models of information processing in the basal ganglia. MIT Press, Cambridge, pp 249–274

Houk JC, Barto AG (1992) Distributed sensorimotor learning. In: Stelmach GE, Requin J (eds) Tutorials in motor behavior II. Elsevier, Amsterdam, pp 71–100

Houk JC, Keifer J, Barto AG (1993) Distributed motor commands in the limb premotor network. Trends Neurosci 16:27–33

Houk JC, Mugnaini E (2002) Cerebellum. In: Squire LR, Bloom FE Roberts ea JL (eds) Fundamental neuroscience. Academic New York

Houk JC, Alford S (1996) Computational significance of the cellular mechanisms for synaptic plasticity in Purkinje cells. Behav Brain Sci 19:457–461

Houk JC, Wise SP (1995) Distributed modular architectures linking basal ganglia, cerebellum, and cerebral cortex: their role in planning and controlling action. Cerebral Cortex 5:95–110

Hua SE, Houk JC (1997) Cerebellar guidance of premotor network development and sensorimotor learning. Learn Mem 4:63–76

Inase M, Sakai ST, Tanji J (1996) Overlapping corticostriatal projections from the supplementary motor area and the primary motor cortex in the macaque monkey: an anterograde double labeling study. J Comp Neurol 373:283–296

Jiang MC, Alheid GF, Nunzi MG, Houk JC (2002) Cerebellar input to magnocellular neurons in the red nucleus of the mouse:

Synaptic analysis in horizontal brain slices incorporating cerebellorubal pathways. Neuroscience 110:105–121
Ito M (2001) Cerebellar long-term depression: Characterization, signal transduction, and functional roles. Physiol Rev 81:1143–1195
Kawagoe R, Takikawa Y, Hikosaka (1998) Expectation of reward modulates cognitive signals in the basal ganglia. Nat Neurosci 1:411–416
Kelly RM, Strick PL (2003) Macroarchiteccture of basal ganglia loops with the cerebral cortex: use of rabies virus to reveal mutisynaptic circuits. Prog Brain Res 143:441–451
Kelly RM, Strick PL (2004) Cerebellar loops with motor cortex and prefrontal cortex of a non-human primate. J Neuroscience
Kim SG, Ugurbil K, Strick PL (1994) Activation of a cerebellar output nucleus during cognitive processing. Science 265:949–951
Kincaid AE, Zheng T, Wilson CJ (1998) Connectivity and Convergence of single corticostriatal axons. J Neuroscience 18:4722–4731
Lamarre Y, Spidalieri G (1983) Fast ballistic arm movements triggered by visual, auditory, and somesthetic stimuli in the monkey. I. Activity of percentral cortical neurons. J Neurophysiol 50:1343–1358
Middleton FA, Strick PL (1997) New concepts about the organization of basal ganglia output. In: Obeso JA, DeLong MR, Ohye C, Marsden CD (eds) The basal ganglia and new surgical approaches for parkinson's disease, advances in neurology, Lippincott-Raven Publishers, Philadelphia, pp 57–68
Miller LE, Houk JC (1995) Motor co-ordinates in primate red nucleus: Preferential relation to muscle activation versus kinematic variables. J Physiol Lond 488:533–548
Minsky M (1986) The Society of mind. Simon and Schuster, New York
Minsky M (1963) Steps toward artificial intelligence. In: Feigenbaum EA, Feldman J (eds) Computers and thought. McGraw-Hill, NYC
Nicola SM, Surmeier DJ, Malenka RC (2000) Dopminergic modulation of neuronal excitability in the striatum and nucleus accumbens. Ann Rev Neurosci 23:185–215
Novak KE, Miller LE, Houk JC (2002) The use of overlapping submovements in the control of rapid hand movements. Exp Brain Res 144:351–364
Ohl FW, Scheich H, Freeman WJ (2001) Change in pattern of ongoing cortical activity with auditory category learning. Nature 412:733–736
Perrett MH, Harries HH, Mistlin JJ, Hietanen JK, Benson JJ, Bevan R, Thomas S, Oram MW, Ortega J, Brierley K (1990) Social signals analyzed at the single cell level: someone is looking at me, something touched me, something moved! Int J Comp Psychol 4:25–55
Petersen SE, Fiez JA (1993) The processing of single words studied with positron emission tomography. Annu Rev Neurosci 16:509–530
Requin J, Riehle A, Seal J (1988) Neuronal activity and information processing in motor control: From stages to continuous flow. Biol Psychol 26:179–198
Riolo RL, Cohen MD, Axelrod R (2001) Evolution of cooperation without reciprocity. Nature 414:441–443
Sakai K, Hikosaka O, Takino R, Miyauchi S, Nielsen M, Tamada T (2000) What and When: Paralled and convergent processing in motor control. J Neurosci 20:2691–2700
Schmahmann JD, Pandya DN (1997) Anatomic organization of the basilar pontine projections from prefrontal cortices in rhesus monkey. J Neurosci 17:438–458
Schultz W (1998) Predictive reward signal of dopamine neurons. J Neurophysiol 80:1–27
Schultz W, Romo R, Ljungberg T, Mirenowicz J, Hollerman JR, Dickinson A (1995) Reward-related signals carried by dopamine neurons. In: Houk JC, Davis JL, Beiser DG (eds) Models of Information Processing in the Basal Ganglia. MIT Press,Cambridge, pp 233–248
Seth AK, McKinstry JL, Edelman GM, Krichmar JL (2004) Visual binding through reentrant connectivity and dynamic synchronization in a brain-based device. Cerebral Cortex 14:1185–99
Strick PL, Dum RP, Picard N (1995) Macro-organization of the circuits connecting the basal ganglia with the cortical motor areas. In: Houk JC, Davis JL, Beiser DG (eds) Models of Information Processing in the Basal Ganglia. MIT Press, Cambridge, pp 117–130
Tanji J (2001) Sequential organization of multiple movements: involvement of cortical motor areas. Ann Rev Neurosci 24:631–651
Tepper JM, Koos T, Wilson CJ (2004) GABAergic microcircuits in the neostriatum. Trends Neurosci 27:662–669
Tyrrell T, Willshaw D (1992) Cerebellar cortex: its simulation and the relevance of Marr's theory. Phil Trans Royal Soc Lond B 336:239–257
Wichmann T, Bergman H, Starr PA, Subramanian T, Watts RL, DeLong MR (1999) Comparison of MPTP-induced changes in spontaneous neuronal discharge in the internal pallidal segment and in the substantia nigra pars reticulata in primates. Exp Brain Res 125:397–409
Wilensky U, Reisman K (1998) Learning biology through constructing and testing computational theories – an embodied modeling approach. In: Bar-Yam Y (ed) Proceedings of the second international conference on complex systems. New England Complex Systems Institute. Nashua
Wilson CJ (1990) Basal Ganglia In: The synaptic organization of the brain. Oxford University Press, NY, pp 279–316
Yeterian EH, Van Hoesen GW (1978) Cortico-striate projections in the rhesus monkey: the organization of certain cortico-caudate connections. Brain Res 139:43–63