A Biologically Inspired Action Selection Algorithm Based on Principles of Neuromodulation

Jeffrey L. Krichmar
Department of Cognitive Sciences
Department of Computer Science
University of California, Irvine
Irvine, CA 92697-5100, USA
jkrichma@uci.edu

Abstract— The brain's neuromodulatory systems play a key role in regulating decision-making and responding to environmental challenges. Attending to the appropriate sensory signal, filtering out noise, changing moods, and selecting behavior are all influenced by these systems. We introduce a neural network for action selection that is based on principles of neuromodulatory systems. The algorithm, which was tested on an autonomous robot, demonstrates valuable features such as fluid switching of behavior, gating in important sensory events, and separating signal from noise.

Keywords – adaptive behavior; computational neuroscience; neuromodulation; neurorobots

I. INTRODUCTION

A general purpose algorithm, based on principles of the brain's neuromodulatory systems, is presented for action selection in robots. Neuromodulatory systems are present in all vertebrates and are critical for an animal to quickly assess the context of sensory input and take action [1]. Neuromodulators signal environmental changes to the nervous system and alter neuronal responses such that the organism can respond quickly and accurately to these changes.

Rather than present a neurobiologically detailed model of how the nervous system achieves this function through neuromodulation (see for example [2]), a general-purpose, but minimal model of neuromodulatory function is developed, which can be applied to robot control.

Although there have been great advances in autonomous robotics [3-6], the controllers of these machines are still very much tailored to specific missions and do not have the behavioral repertoire normally associated with that of biological organisms. Behavior-based robots neither learn from their experience nor adapt to environmental change [7]. Probabilistic robot controllers need an accurate model of their sensors and actuators [8]. Evolutionary robots are constrained by a fitness function specified by the designer [9]. Robots, which are controlled by reinforcement learning or machine learning, are

driven by reward expectation and do not address attention, novelty, and risk assessment [10].

A design based on principles of the neuromodulatory systems would provide a framework that would allow agents to operate autonomously, optimally explore their environment, and be decisive when environmental conditions call for action.

The vertebrate neuromodulatory systems play a key role in regulating decision-making and responding to environmental challenges [1]. In particular, the serotonergic (5-HT) system underlies control of stress, social interactions, and risk-taking behavior [11, 12]. The dopaminergic (DA) system has been implicated in the prediction of rewards and incentive salience or "wanting" [13, 14]. In a theory put forth recently by Boureau and Dayan [15], the serotonergic (5-HT) and dopaminergic (DA) systems oppose each other with respect to predicting punishment (5-HT) versus predicting reward (DA) along one dimension (labeled valence), and inhibition (5-HT) versus invigoration (DA) along another dimension (labeled action). Serotonin influences the amount of risk one is willing to take, the speed at which decisions are made, the impulsiveness of such decisions, and the suppression of actions when they are thought to lead toward a punishment or cost [16].

The cholinergic (ACh) and noradrenergic (NE) systems are thought to play important roles in attention and judging uncertainty [17]. Removal of ACh projections to the cortex impairs the ability to increase attentional effort [18]. NE neurons are sensitive to novel and salient objects in the environment [19] and task relevant stimuli that cannot be fully predicted, such as recognizing an unreliable or oddball stimulus [20]. Each of these neuromodulators triggers the brain's attention system, and depending on the immediacy of an environmental cue, can result in the organism concentrating on the highest priority challenge.

The nervous system responds to these high priority events through phasic neuromodulation, where sensory information and competition through inhibition is amplified relative to recurrent or associational information [1, 21, 22]. The result of this change in the relative weighting of information is to

sharpen responses to environmental input, increase the signal to noise ratio, and drive decisive responses in neural networks.

In the remainder of the paper, a minimal neural model is presented that captures the aspects of neuromodulation described above with the goal of developing a biologically inspired controller for robots. To demonstrate its capabilities, the algorithm will be used to control an autonomous robot in an office setting.

II. METHODS

A. Robot Control

Experiments were run on an iRobot Create equipped with an URG-04-LX laser range finder (Hokuyo Automatic Co. LTD.) and a System 76 netbook running the Ubuntu Linux operating system for computation (see Figure 1). The Matlab Toolbox for iRobot Create (http://www.usna.edu/Users/weapsys/esposito/roomba.matlab/) was used to interface with the robot. The neural simulation and robot control algorithm for iRobot Create was written in Matlab (MathWorks) and can be downloaded at: http://www.socsci.uci.edu/~jkrichma/krichmar_ijcnn2012_roomba_network.m
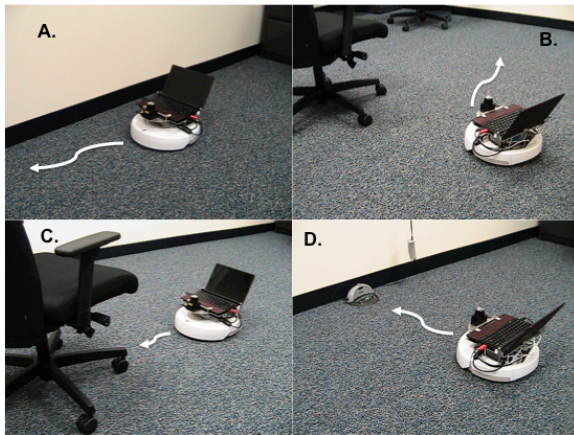

Figure 1. Experiments were run on an iRobot Create with a laser range finder and netbook for onboard computation. The action selection algorithm arbitrated between four behavior states. **A.** Wall Following. **B.** Open Field. **C.** Explore Object. **D.** Find Home.

Robot control was achieved through processing events and states. States were pre-canned behaviors and events were driven by sensory signals. An event could cause a switching of behavior states. The neural simulation, which is described below, arbitrated between incoming events and decided when to switch states.

The robot handled four events: 1) *Laser change.* A large change between laser scans in front of the robot triggered this event. Specifically, a laser change event occurred when there was 100 cm of change in the middle 100 degrees of laser scans. 2) *Battery low.* A random event proportionally related to the drop in battery level since its last charge. The more the battery level had dropped, the more likely the event would occur. 3) *Bump detected.* This event was triggered by the robot's bump sensors or if the laser detected an object closer than 33 cm. 4) *Dock beam.* This event was triggered if the robot detected the iRobot Create Dock Beam, which meant that the robot was within 500 cm of its docking station.

The robot switched between four behavior states: 1) *Wall Following* (Figure 1A). Wall following to the right or left was randomly chosen upon entering this state. Using its laser range finder, the robot would attempt to stay between 40 and 60 cm of a wall or other large object. If a collision with an object was detected, the robot rotated away from the wall (e.g., if following the right wall, the robot would rotate left). 2) *Open Field* (Figure 1B). The robot would drive toward the most open area of the environment, as judged by the laser range finder. If a collision with an object was detected, the robot would rotate clockwise. 3). *Explore Object* (Figure 1C). The robot would move toward the area of the environment that changed the most between laser scans. If a collision with an object was detected, the robot would rotate clockwise. 4) *Find Home* (Figure 1D). In this state, the robot would enter the iRobot Create's cover and dock mode. This caused the robot to perform a random search and an attempt to dock if a buoy from the iRobot Create's dock beam was detected. If the beam was detected, the robot approached the dock. If the robot got within the dock's force field or ten seconds passed without detecting a dock beam, the robot would back up and rotate 180 degrees. If the robot was within the force field, the battery level was set to fully charged.

B. Neural Simulation

Neuromodulatory systems receive sensory information and drive behavior by innervating downstream neural systems. Much of the information these systems receive is sub-cortical and their targets are subcortical as well [23-25]. The general framework of the present architecture is that sensory events can trigger neuromodulatory systems, which in turn drive behavior states. Although the architecture given in Figure 2 is specific to the present problem space, any combination of events and behavioral states could potentially be arbitrated by the present model of neuromodulatory systems.

In the present paper, the neural simulation consisted of four event neurons, each of which corresponded to one of the four events described above, four state neurons, each of which corresponded to one of the four states described above, and neuromodulatory neurons. There was one dopaminergic neuron (DA), one serotonergic neuron (5-HT), and four ACh/NE neurons, each of which corresponded to one of the four events described above. Figure 2 shows the connectivity of the network.

Initial simulations were carried out to set the weights and parameters given in the equations below. Weights were chosen such that the network demonstrated stable activity, and such that a phasic burst of neuromodulatory activity could efficiently drive action selection [1, 2]. State neurons connected to other state neurons with both excitatory (weight = 0.5) and inhibitory (weight = -1.0) connections. Neuromodulatory neurons selectively connected to state neurons with weights set at 5, event neurons selectively connected to neuromodulatory neurons with weights set at 1, event neurons connected to the corresponding ACh/NE neurons with weights set at 1, and each event neuron connected to all the state neurons with weights set at 1.

In the present simulation, laser changes signaled novelty or something potentially rewarding in the environment worth taking a risk to investigate, and thus triggered dopaminergic neurons (LaserChange→DA in Figure 2). A low battery or a dock beam signaled a need for low risk or harm aversive behavior, and thus triggered serotonergic neurons (LowBattery→5-HT and DockBeam→5-HT in Figure 2). A bump could signal either something interesting or noxious in the environment. Therefore, the bump event triggered DA in some simulations and 5-HT in others, as will be explained below. Future versions of the simulation architecture will investigate adaptive switching between DA and 5-HT for shared events. DA neurons triggered curiosity-seeking behaviors, such as OpenField and ExploreObject. 5-HT neurons triggered harm aversive behaviors, such as WallFollow and FindHome.
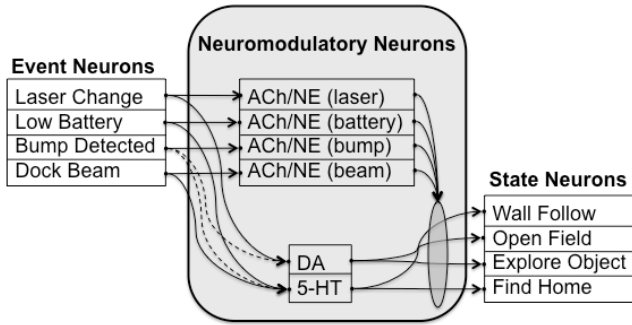

Figure 2. Architecture of the neural simulation. Arrows denote connections between neurons. The dashed lines denote connections that are not present in all experiments. For clarity, connections between state neurons are omitted.

A simulation cycle, t, occurred approximately once per second, which was roughly the time needed to read the robot's sensors, update the neural simulation, and send a motor command to the robot. The main limitation for cycle duration was Matlab handling of I/O. Future versions of the software will be written in C/C++ to speed up I/O and shorten simulation cycles.

Event neurons were binary and set to 1 when an event occurred, and 0 otherwise. All other neurons were governed by the following activation function, which kept neural activity between 0 and 1:

$$n(t) = \frac{1}{1 + e^{-gI(t)}} \qquad (1)$$

Where g was the gain of the function and I was the input to the neuron. Given the initial weights, gains and the baseline input in equation 2 were set such that the range of excitatory and inhibitory inputs to the neuron would cover the full range of the sigmoid curve. Therefore, the gain was set to 2 for state and neuromodulatory neurons, and 5 for ACh/NE neurons. Input to the neuron was based on pre-synaptic neural activity, previous neural activity, and neuromodulation:

$$I_j(t) = b + \sum_i c(t) n_i(t) w_{ij}(t) + p n_j(t-1) + nm(t) \qquad (2)$$

Where b was the baseline input set to -1.0 for DA and 5-HT neurons, -0.5 for ACh/NE neurons, and -1.0+rand (0.0,0.5) for state neurons, $c(t)$ was set to the sum of DA and 5-HT neuron

activity for inhibitory connections, otherwise $c(t)$ was set to 1.0 [1]. p was the persistence set to 0.25 for all neurons and $nm(t)$ was the neuromodulatory input into state neurons:

$$nm_i(t) = \sum_j \sum_k n_k(t) w_{ki}(t) AChNE_j(t) e_j(t) w_{ji}(t) \qquad (3)$$

Where $nm_i(t)$ is the neuromodulatory input into state neuron i, $n_k(t)$, is the activity of either the DA or 5-HT neuron, $w_{ki}(t)$ is the weight from neuromodulatory neuron k to state neuron i, $AChNE_j(t)$ and $e_j(t)$ are the activities of ACh/NE and event neurons corresponding to event j, and $w_{ji}(t)$ is the weight from event neuron j to state neuron i.

Short-term plasticity dictated the activity of the neuromodulatory neurons. Because cholinergic and noradrenergic neurons respond to novelty, short-term plasticity was such that ACh/NE neurons responded vigorously to rare events and quiescent to frequent events. Because dopaminergic and serotonergic neurons are sensitive to appetitive and aversive events, respectively, short-term plasticity for DA and 5-HT were set such that they were sensitized to salient events.

To capture this plasticity, Event to 5-HT, DA, and ACh/NE weights were set based on the following update rule:

$$w_{ij}(t) = \begin{cases} p w_{ij}(t-1) & if\ e_i = 1 \\ w_{ij}(t-1) + \dfrac{1 - w_{ij}(t-1)}{\tau} & otherwise \end{cases} \qquad (4)$$

Where i is the index of the event neuron, j is the index of the 5-HT, DA, or ACh/NE neuron, p is the amount of change in response to an event, and τ, which was set to 50, was a time constant that governed the rate at which weights returned to their original value. Weights from event neurons to 5-HT and DA neurons were facilitating, meaning that each event caused the weight to increase ($p = 1.1$). Weights from event neurons to ACh/NE neurons were depressing, meaning that each event caused the weight to decrease ($p = 0.1$). These rates were set based on the expected occurrence of events during a five-minute session of running the robot.

Action selection occurred after the neural activities and weight updates were calculated. A new behavior state was selected if the maximally active state neuron had activity greater than 0.67, which was set such that new actions would be selected roughly 2 to 3 times per minute.

III. RESULTS

The robot was run in a series of experiments in an office having several desks, chairs, a door and a window (Figure 1). Each experiment lasted approximately 5 minutes. Some simulations were slightly longer to allow the robot to complete a behavioral state at the end of a trial. The robot was run under three experimental conditions: 1) Bumps were treated as potentially harmful by connecting Bump event neurons to the 5-HT neuron. 2) Bumps were treated as novel and interesting by connecting Bump event neurons to the DA neuron. 3) The second condition was repeated with the ACh/NE neurons always active. This condition tested the effectiveness of the ACh/NE neuromodulation on action selection. All conditions were run five times.
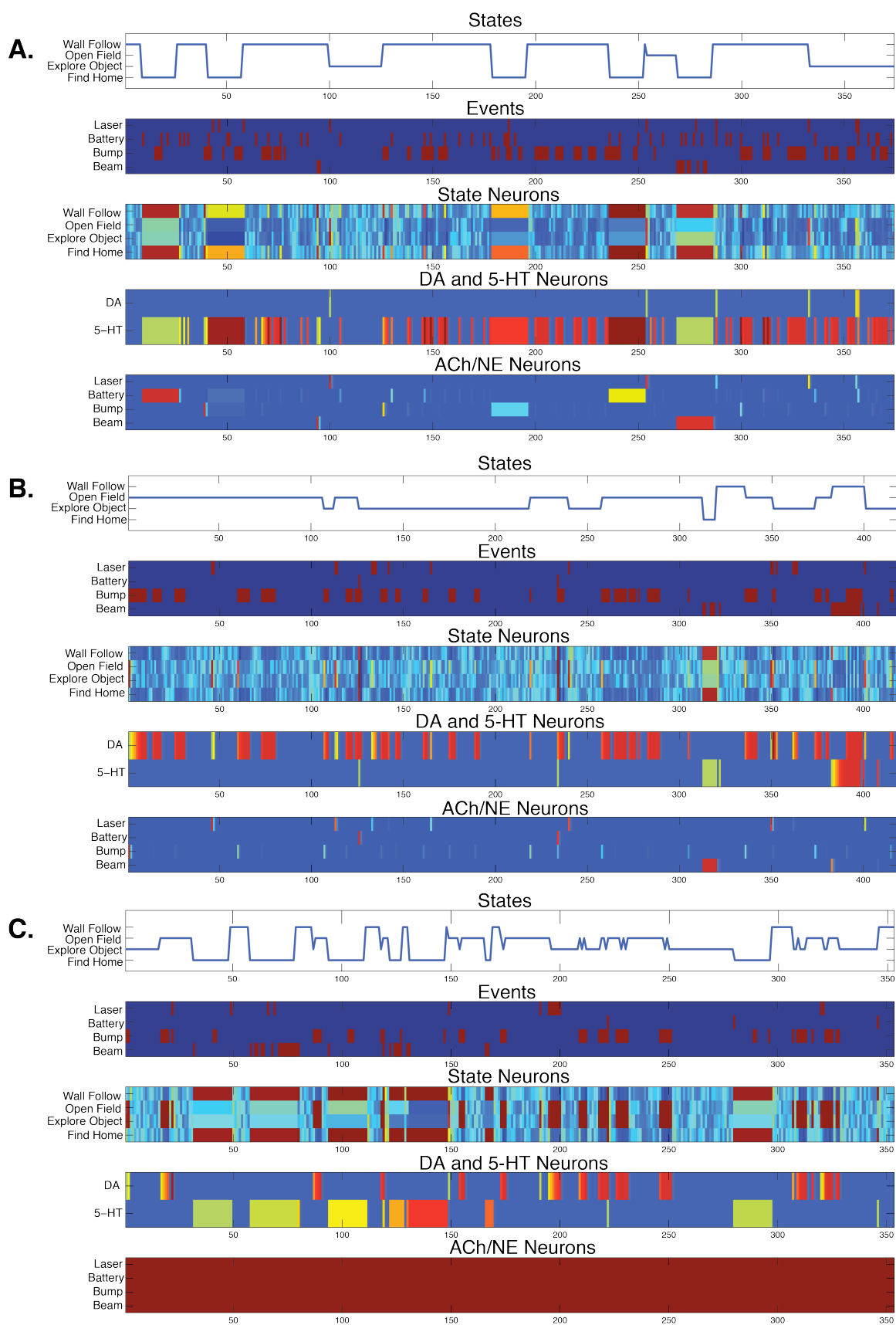
Figure 3. State and neural responses during representative trials. Neural activity ranged from quiescent (dark blue) to maximally active (bright red). The x-axis shows the simulation cycles, which took roughly 1 second per cycle. **A.** Bump event neurons connected to the 5-HT neuron. **B.** Bump event neurons connected to the DA neuron. **C.** Bump event neurons connected to the DA neurons and ACh/NE neurons were set to be maximally active.

A. Action Selection and Neuronal Response

Figure 3 shows representative trials from the three experimental conditions. The top chart shows the behavioral states over time. The activity of event neurons, state neurons, and neuromodulatory neurons are shown underneath the behavioral state chart. Neuronal response is color coded from dark blue (0 or quiescent) to bright red (1 or maximally active).

1) Harm aversive responses

Figure 3A shows an experiment in which the neural simulation had a connection from Bump events to the 5-HT neuron. Because connections from Event neurons to the 5-HT neuron were plastic and facilitative (Eqn. 4), each event caused an increase in serotonergic activity during the trial. This resulted in the robot having a tendency to stay in the harm aversive states, such as WallFollow or FindHome. Note that when the robot was in the FindHome state, neurons were not updated because the robot was running the pre-canned iRobot Create Cover and Dock algorithm.

The ACh/NE neurons responded to change or uncertainty in the environment. This was achieved through the plastic, depressive connections from Event neurons to ACh/NE neurons (Eqn. 4). ACh/NE filtered out events that occurred frequently and responded strongly to unexpected events. For example, the battery event occurred throughout the trial. ACh/NE responded to the event initially (cycle 10 in Figure 3A) resulting in a switch to the FindHome state. However, the system ignored the event until cycle 230, when a battery event occurred after a long absence. The spike in ACh/NE activity resulted in a spike in State neuron activity for WallFollow and FindHome, resulting in the robot selecting the FindHome behavioral state. Close examination of Figure 3 shows how this gating of important or novel events can result in swift and decisive action selection. In contrast, uninteresting or frequent events are automatically ignored.

2) Curiosity and risk-taking responses

Changing the connection from the Bump event such that it triggered DA activity caused the robot to spend more time exploring objects and moving toward the center of the environment (Figure 3B). DA activity was much higher in the Figure 3B trial than in the Figure 3A trial. This resulted in the robot choosing the curiosity seeking OpenField and ExploreObject states more often. ACh/NE activity had the same gating effect as before. For example, a rare laser event at cycle 350 resulted in a spike of activity in the ACh/NE and ExploreObject neurons causing the robot to select the ExploreObject behavioral state (Figure 3B). In contrast, the ACh/NE neurons habituated to the constant Bump events from cycles 100 to 200 causing the robot to ignore these events for the most part.
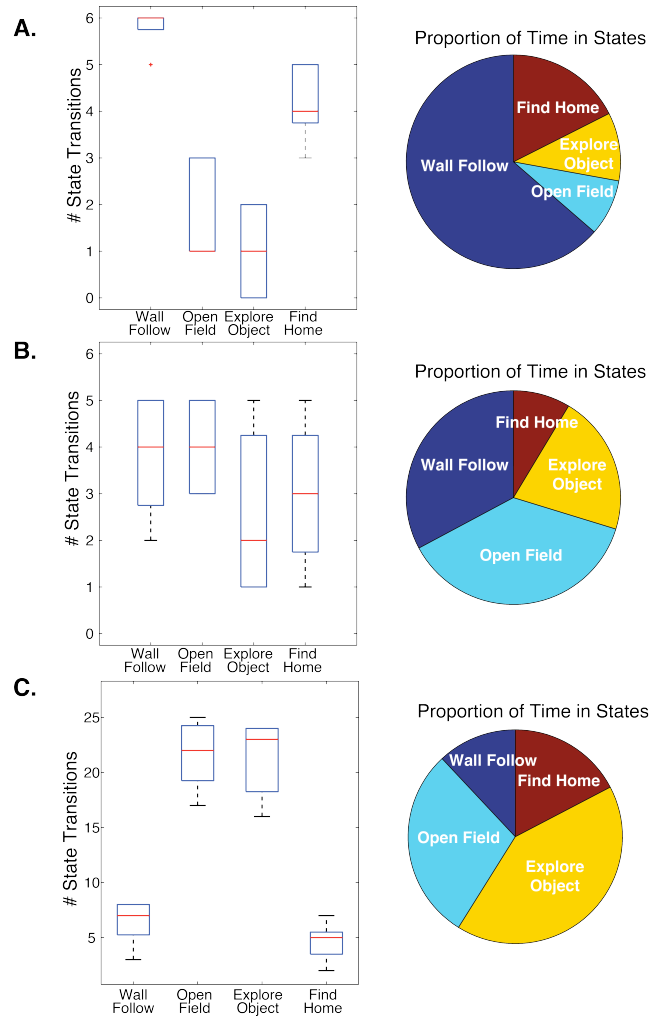

Figure 4. Behavioral state selection. Boxplots on the left show the number of transitions to a state. Red lines denote the median, blue box denotes 25th to 75th percentile, and whiskers show the remaining data. Charts on the right show the proportion of time spent in a behavioral state. Each condition had five trials. **A.** Bump event neurons connected to the 5-HT neuron. **B.** Bump event neurons connected to the DA neuron. **C.** Bump event neurons connected to the DA neuron when ACh/NE neurons were always maximally active.

B. Behavioral Performance

1) Harm aversive behavior

When there was more serotonergic activity, the robot tended to select harm aversive behaviors such as WallFollow and FindHome (Figure 4A). The increase in serotonergic activity was due to connections from the Battery, Beam, and Bump event neurons to the 5-HT neuron. There was a significant increase in the amount of time spent in WallFollow and FindHome by robots in this condition compared to when Bump events triggered DA activity ($p < 0.01$, Wilcoxon rank sum test, and compare Figure 4A to 4B). Wall following and staying near the nest is typical of rodents when in an unfamiliar environment or when they are highly anxious [26, 27].

2) Curiosity and risk-taking behavior

The robot was more likely to explore its environment through OpenField and ExploreObject behaviors when its

dopaminergic system was active (Figure 4B). The increase in dopaminergic activity, during this condition, was due to connections from the Laser, and Bump event neurons to the DA neuron. There was a significant increase in the amount of time spent in the OpenField state by robots in this condition ($p < 0.001$, Wilcoxon rank sum test, and compare Figure 4A to 4B). As rodents become more familiar with their surroundings, they will tend to move toward the center of the environment and will explore novel objects [26]. Moreover, dopamine activity has been shown to be involved in novelty seeking [25, 28] and incentive salience [13].

3) Distracted behavior

In agreement with empirical findings, an impaired ACh/NE system resulted in distracted behavior [29]. In the trial shown in Figure 3C, the ACh/NE neurons were always set to 1. The result was more action selection and the inability to ignore unimportant events. Figure 4C illustrates that there was a significant increase in switching to OpenField and ExploreObject states, as well as the WallFollow and FindHome states (note the change in the y-axis scale) when ACh/NE was impaired ($p < 0.01$, Wilcoxon rank sum test). In general, an overactive ACh/NE system results in the robot responding to nearly every event with the corresponding behavioral state.

IV. DISCUSSION

An algorithm was introduced for action selection and decision-making based on principles of neuromodulatory signaling in the brain. Similar to classic robot control algorithms, such as subsumption architecture [30] and behavior-based schemas [31], the algorithm presented here automatically arbitrates between actions based on current sensory input. However, the present algorithm has the ability to adapt to changes in the environment by: 1) increasing sensitivity to sensory inputs, 2) responding to unexpected or rare events, and 3) habituating or ignoring uninteresting events.

A. Increasing Sensitivity and Shifting Behavior

By changing the internal meaning of a sensory event (e.g., bump triggers DA instead of 5-HT), the robot switches from harm aversive to curious behavior (compare Figure 4A to 4B). This behavioral shift is brought about by an increased sensitivity of a specific neuromodulatory system to environmental events. The increased sensitivity triggers activity in the appropriate state neurons causing a new behavior to be selected. Similar to a biological system, this shift is swift and decisive [1, 23]. In the present work, the designer of the algorithm chose the connectivity based on the desired robot function. In the future, it will be of interest to have this switch in connectivity occur automatically due to experience dependent learning.

Similar to the robot behavior shown here, rodents switch from risk averse to exploratory behavior as they become familiar with an environment [26]. Interestingly, these systems affect similar functionality in humans. For instance, variation of the serotonergic system influences social anxiety [32]. Variations in the dopaminergic system have been shown to affect risk-taking during gambling, the ability to filter out noise, and cognitive flexibility [33, 34].

B. Responding to Unexpected Events

Comparable to biological nervous systems, the simulated ACh/NE neurons responded to unexpected events with a sharp, phasic burst of activity causing the appropriate neuromodulatory and state neurons to respond [1, 23]. For example, the neural responses to a rare event such as the beam event (see cycle 270 of Figure 3A) caused the ACh/NE system to gate in this event to the system such that it could be handled properly.

The ACh/NE system is known to be critical for appropriately allocating attention [17, 18]. When the ACh/NE system was impaired in the algorithm, the robot lost its ability to filter out noise and responded to any incoming sensory event (Figure 3C and 4C). Such distracted behavior, when these systems are impaired, is in agreement with empirical findings [29].

C. Ignoring Uninteresting Events

In addition to gating in important events, the ACh/NE system filtered out unimportant sensory events. The cholinergic and noradrenergic neurons respond strongly to novel and salient objects in the environment [17, 19]. The algorithm demonstrated this selectivity. For example, since the robot was operating in a crowded office, the bump event was being triggered constantly (see Figure 3). Initially, the robot acknowledged this event, but over time, the robot learned to ignore this event so that it could respond appropriately to other, more meaningful events. The ACh/NE system was necessary for this filtering as could be seen in Figure 3C where the robot responded to these frequent bump events resulting in constant switching of behavior when the ACh/NE system was impaired.

D. Related Work

While there have been many models of action selection, the present work addresses how principles of neuromodulation could control autonomous robot behavior. Briefly, some related work is reviewed here. In the field of autonomous agents, "affect" has been used to shape the behavior of both simulated and robotic agents. For example, Blanchard and Cañamero examined trade-offs between exploration and exploitation based on the notions of well-being and affect [35]. They define "affect" as the immediate or instinctive evaluation of a situation (positiveness or negativeness). In their experiments, the robot's well-being was related to the agent's internal value judgment of its distance to a box, and its affect was related to the agent's evaluation of safety based on its familiarity with the objects it was sensing. The robot's behavior was dictated by a dynamical system and the affect term modulates the robot's motivation to continue. Similar to affect is the notion of comfort or safety, which has also been proposed to influence exploration behavior in robots [36]. Affect has been used in evolutionary algorithms to develop exploration/exploitation strategies in dynamic choice trials [37], and affect has been embedded into the reinforcement-learning algorithm where reward is based on the happiness and sadness of the agent [38]. In the autonomous agent models discussed above, brain-inspired terms such as anticipation,

affect, emotions, hormones, and modulation are simulated to facilitate action selection and exploration/exploitation trade-offs. However, these models do not specifically address the mechanism by which the nervous system gives rise to behaviors, such as action selection, exploration, and exploitation. The framework presented here, provides a neural description of how neuromodulation of neural circuitry could account for such adaptive behavior. A possible advantage of the present approach is that it provides a model that can be directly tested against animal models; both in its behavioral response and in its neuronal response. This synergy between empirical and simulated data, which can lead to improvements in the model and predictions in the modeled organism, is a goal of computational neuroscience.

In the field of computational neuroscience, theoretical models have been proposed on neuromodulation, but they have not considered all of the neuromodulatory systems and their interactions with cortical and subcortical areas. The phasic response of the dopamine system has been proposed to signal temporal difference error [14]. Following this idea, the phasic response of dopamine has been modeled to shape action selection and reward anticipation behavior in neurorobots [39-42]. Our previous model took into consideration the phasic aspects of dopaminergic and serotonergic neuromodulation [2]. This model postulated, similar to a model of noradrenergic neuromodulation [23], that phasic neuromodulation causes an organism to be more decisive, whereas a lack of phasic response would result in more arbitrary action selection. However, to our knowledge, few if any researchers have developed a model that includes the cholinergic, noradrenergic, and serotonergic systems in one robot controller.

V. CONCLUSION AND FUTURE DIRECTIONS

A biologically inspired algorithm was introduced for action selection and controlling robot behavior. The algorithm shows several important features for autonomous robot control in general, such as, fluid switching of behavior, gating in important sensory events, and separating signal from noise. The basic algorithm and architecture (Figure 2) can be tailored to many robot control problems. In the future, an algorithm to adaptively gate shared events, such as the Bump event, between different modulators will be developed. One possible implementation would be to set a comfort level for the agent such that when the agent is anxious, a bump is noxious and triggers 5-HT, and when the agent is curious, a bump could be considered something to explore. Mice have been known to switch between these comfort levels when exploring environments [26].

The present architecture did not take into consideration long-term memory and goal-directed signals from cortical areas. In the future, the architecture will be expanded such that predictive signals from the frontal cortex could modulate the neuromodulators and train the neuromodulators to predict events and take actions before those events occur. The prefrontal cortex and anterior cingulate cortex are known to have strong projections to all the neuromodulatory systems [23, 43]. In addition, future versions of the architecture will have a layer that can build up a memory of events and their

location. In general, the present architecture can be thought of an action selection module that can fluidly switch between behavioral states and could be added onto many different control systems.

REFERENCES

[1] J. L. Krichmar, "The Neuromodulatory System – A Framework for Survival and Adaptive Behavior in a Challenging World," *Adaptive Behavior,* vol. 16, pp. 385-399, 2008.

[2] B. R. Cox and J. L. Krichmar, "Neuromodulation as a Robot Controller: A Brain Inspired Design Strategy for Controlling Autonomous Robots," *IEEE Robotics & Automation Magazine,* vol. 16, pp. 72-80, 2009.

[3] A. Cho, "Robotics. Robotic cars tackle crosstown traffic--and not one another," *Science,* vol. 318, pp. 1060-1, Nov 16 2007.

[4] W. W. Gibbs, "From finish to start. Was the grand challenge robot race in March the fiasco it appeared to be? Hardly, argues William "Red" Whittaker. The annual event is pushing mobile robotics to get real," *Sci Am,* vol. 291, pp. 33-4, Aug 2004.

[5] S. Squyres, *Roving Mars : Spirit, Opportunity, and the Exploration of the Red Planet*. New York, NY: Hyperion, 2005.

[6] B. Yenne, *Attack of the Drones*. St. Paul, MN: Zenith Press, 2004.

[7] J. Jones and D. Roth, *Robot Programming : A Practical Guide to Behavior-Based Robotics*. New York, NY: McGraw-Hill, 2003.

[8] S. Thrun, W. Burgarde, and D. Fox, *Probabilistic Robotics*. Cambridge, MA: The MIT Press, 2005.

[9] S. Nolfi and D. Floreano, *Evolutionary robotics: The biology, intelligence, and technology fo self-organizing machines.* Cambridge, MA: MIT Press, 2000.

[10] R. S. Sutton and A. G. Barto, *Reinforcement Learning: An Introduction*. Cambridge, MA: MIT Press, 1998.

[11] R. Cools, A. C. Roberts, and T. W. Robbins, "Serotoninergic regulation of emotional and behavioural control processes," *Trends Cogn Sci,* vol. 12, pp. 31-40, Jan 2008.

[12] M. J. Millan, "The neurobiology and control of anxious states," *Prog Neurobiol,* vol. 70, pp. 83-244, Jun 2003.

[13] K. C. Berridge, "Motivation concepts in behavioral neuroscience," *Physiol Behav,* vol. 81, pp. 179-209, Apr 2004.

[14] W. Schultz, P. Dayan, and P. R. Montague, "A neural substrate of prediction and reward," *Science,* vol. 275, pp. 1593-9, Mar 14 1997.

[15] Y. L. Boureau and P. Dayan, "Opponency revisited: competition and cooperation between dopamine and

serotonin," *Neuropsychopharmacology,* vol. 36, pp. 74-97, Jan 2011.

[16] R. Cools, K. Nakamura, and N. D. Daw, "Serotonin and dopamine: unifying affective, activational, and decision functions," *Neuropsychopharmacology,* vol. 36, pp. 98-113, Jan 2011.

[17] A. J. Yu and P. Dayan, "Uncertainty, neuromodulation, and attention," *Neuron,* vol. 46, pp. 681-92, May 19 2005.

[18] D. J. Bucci, P. C. Holland, and M. Gallagher, "Removal of cholinergic input to rat posterior parietal cortex disrupts incremental processing of conditioned stimuli," *J Neurosci,* vol. 18, pp. 8038-46, Oct 1 1998.

[19] A. Vankov, A. Herve-Minvielle, and S. J. Sara, "Response to novelty and its rapid habituation in locus coeruleus neurons of the freely exploring rat," *Eur J Neurosci,* vol. 7, pp. 1180-7, Jun 1 1995.

[20] G. Aston-Jones, J. Rajkowski, P. Kubiak, and T. Alexinsky, "Locus coeruleus neurons in monkey are selectively activated by attended cues in a vigilance task," *J Neurosci,* vol. 14, pp. 4467-80, Jul 1994.

[21] M. E. Hasselmo and J. McGaughy, "High acetylcholine levels set circuit dynamics for attention and encoding and low acetylcholine levels set dynamics for consolidation," *Prog Brain Res,* vol. 145, pp. 207-31, 2004.

[22] M. Kobayashi, K. Imamura, T. Sugai, N. Onoda, M. Yamamoto, S. Komai, and Y. Watanabe, "Selective suppression of horizontal propagation in rat visual cortex by norepinephrine," *Eur J Neurosci,* vol. 12, pp. 264-72, Jan 2000.

[23] G. Aston-Jones and J. D. Cohen, "An integrative theory of locus coeruleus-norepinephrine function: adaptive gain and optimal performance," *Annu Rev Neurosci,* vol. 28, pp. 403-50, 2005.

[24] M. Goard and Y. Dan, "Basal forebrain activation enhances cortical coding of natural scenes," *Nat Neurosci,* Oct 4 2009.

[25] P. Redgrave and K. Gurney, "The short-latency dopamine signal: a role in discovering novel actions?," *Nat Rev Neurosci,* vol. 7, pp. 967-75, Dec 2006.

[26] E. Fonio, Y. Benjamini, and I. Golani, "Freedom of movement and the stability of its unfolding in free exploration of mice," *Proc Natl Acad Sci U S A,* vol. 106, pp. 21335-40, Dec 15 2009.

[27] L. K. Heisler, H. M. Chu, T. J. Brennan, J. A. Danao, P. Bajwa, L. H. Parsons, and L. H. Tecott, "Elevated anxiety and antidepressant-like responses in serotonin 5-HT1A receptor mutant mice," *Proc Natl Acad Sci U S A,* vol. 95, pp. 15049-54, Dec 8 1998.

[28] E. S. Bromberg-Martin, M. Matsumoto, and O. Hikosaka, "Dopamine in motivational control: rewarding, aversive, and alerting," *Neuron,* vol. 68, pp. 815-34, Dec 9 2010.

[29] M. Sarter, W. J. Gehring, and R. Kozak, "More attention must be paid: the neurobiology of attentional effort," *Brain Res Rev,* vol. 51, pp. 145-60, Aug 2006.

[30] R. A. Brooks, "Intelligence Without Representation," *Artificial Intelligence,* vol. 47, pp. 139-159, 1991.

[31] R. C. Arkin, *Behavior-Based Robotics (Intelligent Robotics and Autonomous Agents)*. Cambridge: The MIT Press, 1998.

[32] A. Caspi, K. Sugden, T. E. Moffitt, A. Taylor, I. W. Craig, H. Harrington, J. McClay, J. Mill, J. Martin, A. Braithwaite, and R. Poulton, "Influence of life stress on depression: moderation by a polymorphism in the 5-HTT gene," *Science,* vol. 301, pp. 386-9, Jul 18 2003.

[33] P. Roussos, S. G. Giakoumaki, S. Pavlakis, and P. Bitsios, "Planning, decision-making and the COMT rs4818 polymorphism in healthy males," *Neuropsychologia,* vol. 46, pp. 757-63, Jan 31 2008.

[34] G. Winterer and D. R. Weinberger, "Genes, dopamine and cortical signal-to-noise ratio in schizophrenia," *Trends Neurosci,* vol. 27, pp. 683-90, Nov 2004.

[35] A. Blanchard and L. Cañamero, "Modulation of Exploratory Behavior for Adaptation to the Context.," in *Proc. of the AISB'06 Symposium Biologically Inspired Robotics (Biro-net)*, University of Bristol, UK, 2006, pp. 131-137.

[36] M. Likhachev and R. Arkin, "Robotic comfort zones.," in *SPIE Sensor Fusion and Decentralized Control in Robotic Systems III*, 2000, pp. 27-41.

[37] A. McMahon, D. Scott, P. Baxter, and W. Browne, "An autonomous explore/exploit strategy. ," in *AISB'06 Symposium on Nature Inspired Systems*, 2006, pp. 192-201.

[38] M. A. Salichs and M. Malfaz, "Using emotions on autonomous agents. The role of happiness, sadness and fear.," in *Proceedings of the AISB'06 Symposium on Integrative Approaches to Machine Consciousness* Bristol, UK, 2006, pp. 157-164.

[39] W. H. Alexander and O. Sporns, "An Embodied Model of Learning, Plasticity, and Reward," *Adaptive Behavior,* vol. 10, pp. 143-159, 2002.

[40] J. L. Krichmar and G. M. Edelman, "Machine Psychology: Autonomous Behavior, Perceptual Categorization, and Conditioning in a Brain-Based Device," *Cerebral Cortex,* vol. 12, pp. 818-830, 2002.

[41] J. L. Krichmar, D. A. Nitz, J. A. Gally, and G. M. Edelman, "Characterizing functional hippocampal pathways in a brain-based device as it solves a spatial memory task," *Proc Natl Acad Sci U S A,* vol. 102, pp. 2111-6, Feb 8 2005.

[42] O. Sporns and W. H. Alexander, "Neuromodulation and plasticity in an autonomous robot," *Neural Netw,* vol. 15, pp. 761-74, Jun-Jul 2002.

[43] L. A. Briand, H. Gritton, W. M. Howe, D. A. Young, and M. Sarter, "Modulators in concert for cognition: Modulator interactions in the prefrontal cortex," *Prog Neurobiol,* vol. 83, pp. 69-91, Oct 2007.