Framework and implications of virtual neurorobotics

Philip H. Goodman1,2,†, Quan Zou1 and Sergiu-Mihai Dascalu2

1 Department of Medicine and Program in Biomedical Engineering, University of Nevada, Reno, USA.
2 Department of Computer Science and Engineering, University of Nevada, Reno, USA.

Despite decades of societal investment in artificial learning systems, truly “intelligent” systems have yet to be realized. These traditional models are based on input-output pattern optimization and/or cognitive production rule modeling. One response has been social robotics, using the interaction of human and robot to capture important cognitive dynamics such as cooperation and emotion; to date, these systems still incorporate traditional learning algorithms. More recently, investigators are focusing on the core assumptions of the brain “algorithm” itself—trying to replicate uniquely “neuromorphic” dynamics such as action potential spiking and synaptic learning. Only now are large-scale neuromorphic models becoming feasible, due to the availability of powerful supercomputers and an expanding supply of parameters derived from research into the brain’s interdependent electrophysiological, metabolomic and genomic networks. Personal computer technology has also led to the acceptance of computer-generated humanoid images, or “avatars”, to represent intelligent actors in virtual realities. In a recent paper, we proposed a method of virtual neurorobotics (VNR) in which the approaches above (social-emotional robotics, neuromorphic brain architectures, and virtual reality projection) are hybridized to rapidly forward-engineer and develop increasingly complex, intrinsically intelligent systems. In this paper, we synthesize our research and related work in the field and provide a framework for VNR, with wider implications for research and practical applications.

Keywords: neurorobotics, human-robot interface, virtual reality, artificial intelligence, social robotics, epigenetic robotics, reinforcement

INTRODUCTION
An overarching societal goal is to understand animal and human intelligence and translate that knowledge into technology for prosthetic, assistive, and decision support applications. Traditional research in this field considers the brain to be a specially adapted information-processing system, which can be modeled using mathematical optimization or production rule artificial intelligence systems. Despite many decades of investment in such learning and classification systems, however, this approach has yet to yield truly “intelligent” systems. One proposed remedy comes from research in social robotics, which attempts to augment the understanding of intelligent behavior by capturing the important dynamics of cognition using robotic interaction with humans (Dautenhahn, 2007; Scheutz et al., 2007). However, almost all social robotics systems to date continue to incorporate some mixture of existing machine learning and production rule cognitive systems.

For this reason, investigators are now asking whether critical neural dynamics have indeed been left out of the traditional models. Fortunately, the past two decades of neuroscience research has yielded an abundance of quantitative parameters that characterize the brain’s interdependent electrophysiological (Markram et al., 1997; Schindler et al., 2006), genomic (Toledo-Rodriguez et al., 2004), proteomic (Toledo-Rodriguez et al., 2005),
metabolomic and anatomic (Wang et al., 2006) networks.

Researchers now have access to over a hundred neuroscience databases (Society for Neuroscience, 2007), including automated warehousing collections such as the Allen Brain Atlas (Allen Institute, 2007) and a new data-sharing website sponsored jointly by the U.S. National Science Foundation and National Institutes for Health, called the Collaborative Research in Computational Neuroscience (Teeters et al., 2008).

A previous limitation to the use of biologically realistic models has been the computational overhead. Fortunately, the past decade has witnessed an order of magnitude increase in computation power of individual computers and cluster configurations with a tremendous drop in cost for system components. A few groups have already reported simulations on the order of one million simplified neural elements (Izhikevich et al., 2004; Ripplinger et al., 2004) using supercomputer clusters.

Growth in computational technology has also encouraged nontechnical persons to participate across the Internet using “avatars” in complex virtual reality games and social networking “communities” (e.g., Second Life), which may include not only other human participants but also programmed robots. Thus, taken together, advances in computer technology and interactive 3-D software have set the stage not only to facilitate supercomputer modeling of realistic brains, but also to promote acceptance by humans that virtual reality projections may be capable of meaningful cognitive interaction.

Realistic brain simulation faces several remaining challenges, however. Developing tenable models to capture the essence of natural intelligence for real-time application requires that we discriminate features underlying information processing and intrinsic motivation from those reflecting biological constraints (such as maintaining structural integrity and transporting metabolic products). Furthermore, despite the large and increasing number of physiological parameters provided by experimental inquiry, most of the data relates either to the very small scale of individual or small groups of neurons (e.g., intracellular, 2-photon, or unit recordings at discrete recording sites), or at the other extreme, the joint effect of thousands or millions of neurons over millimeter (optical imaging) or centimeter fields (fMRI and PET). Thus the architecture and response patterns at the middle scale, or “mesocircuit”, remain largely uncharacterized, requiring that the brain modeler proposes and systematically tests plausible connection patterns and learning dynamics.

Another challenge in designing neuromorphic systems is that they must in some way be driven intrinsically by a motivational influence such that the dynamics that subserve information processing are themselves affected by a drive to accomplish the tasks (with neural learning that reinforces successful behavioral adaptation) (Oudeyer and Kaplan, 2007; Oudeyer et al., 2007; Samejima and Doya, 2007; Schweighofer et al., 2007). The motivational system must capture “the aboutness” of its own relationship to other behaving entities (and vice versa) in its environment (i.e., intentionality).

Considered together, physiological responsiveness to intrinsic motivation with intentionality should reflect behaviors consistent with emotional drive rather than by rules or objectives specified under the traditional information-processing paradigm. This suggests that “intelligence” has evolved most directly as a way to better serve emotional drive (rather than in spite of it).

We therefore hypothesize that the development of truly intelligent systems cannot occur outside the real-time, emotional interaction of humans with an intentionality-capable neuromorphic system. This does not exclude the possibility that intelligent systems, once refined, could ultimately be cloned at a point in development where they are ready to learn advanced tasks. Hence, to grow intelligent systems we must start with minimalist brain architectures that are capable of being driven by intrinsic motivation and intentionality in scenarios requiring intelligent behavior in a real-world context. One approach to growing human-like intelligence is to recapitulate the way in which children develop cognitive functions over the first several years of social experience.

**DEFINITION OF VNR**

In testing our hypothesis, it would be relevant not only to grow such intelligent systems but also to comprehend, at each step, the differential changes in architecture giving rise to novel and intelligent cognition. To address these objectives, in a recent publication, Goodman et al. (2007) proposed a hybridization of neuromorphic brain modeling validation using virtually projected robots interacting with human actors, which we call “virtual neorobotics” (VNR). Our proposed definition, open to future collaborative revision, is defined in Table 1. The definition expands upon the definition of neorobotics, which alone would imply a biologically representative robotic control system (criterion 3), and "virtual", which suggests interaction with a human. To test our hypothesis, we additionally require that the robotic system demonstrate sufficient physical and cognitive realism that the human accepts the robot as deserving of emotional reward (criteria 1 and 4) in a real-time interactive loop (criterion 2), with a cognitive architecture potentially extensible to larger cognitive scale (criterion 5).
Table 1 | Definition of virtual neurorobotics (VNR).

A computer-facilitated behavioral loop wherein a human interacts with a projected robot that meets the following five criteria:

1. The robot is sufficiently embodied for the human to tentatively accept the robot as a social, emotional partner
2. The human-robot interaction loop operates in real time, with no pre-specified parcellation into receptive and responsive time windows
3. The cognitive control is a neuromorphic brain emulation incorporating realistic neuronal dynamics with time constants that reflect synaptic activation and learning, established membrane and circuitry properties
4. The neuromorphic architecture can potentially provide circuitry underlying intrinsic motivation and intentionality, using "emotional" rather than rule-based learning & reinforcement
5. The neuromorphic architecture is expandable to progressively larger scale and complexity to support brain model development and validation

Table 2 | Specification of a virtual neurorobotic system.

I. Temporal sequencing (TIMING)
   1. Computation and communication provide nearly real-time ROBOT response, to maintain cognitive and emotional linkage
   2. Time is not segmented a priori for ROBOT or ACTOR receptiveness or reaction

II. Environment (SCENE)
   1. Realistic contents, including sights, sounds, and objects
   2. May be affected by the actions of ROBOT or ACTOR
   3. May include other ROBOTs or multiple ACTORs

III. Live participant (ACTOR)
   1. Human, child or adult, depending on type of target intelligence
   2. Willingness to accept SCENE as realistic situation
   3. Willingness to assume ROBOT has ability to perceive and respond meaningfully
   4. Willingness to attribute intentionality to ROBOT

IV. Neuromorphic system (ROBOT)
   1. Central nervous subsystem (BRAIN) may include neocortex, hippocampus, basal ganglia, and/or other limbic regions relating to attention, reward, and fear
   2. Computational architecture must incorporate biologically plausible learning algorithms and support for expansion to progressively larger scale and complexity
   3. Repertoire of sensors, expressions, and behaviors are commensurate with its physical and brain complexity

The components of the real-time loop are further delineated in Table 2. Here, we emphasize that the interaction between human and virtual robot be unscripted, of a spontaneous, action-reaction nature. That is, there is no segmentation of behavioral time into periods wherein the robot is receptive, waiting, analyzing, and/or taking action. The action-reaction requirement implies also that the system operate nearly in real time. In our experiences, human actors readily accept delays of up to 3 or 4 s without becoming frustrated about unrealistic robotic response and losing cognitive and emotional linkage. Of course, the range of behaviors is indirectly constrained by the sensory and motor capabilities of the robot, the types of behaviors exhibited by the human, and the context (e.g., background activity).

RESEARCH

To-date there is a paucity of literature meeting our criteria (see Related Work, below). Thus we focus here on our own research (Goodman et al., 2007) as an example of the VNR principles. In that work, we chose an instinctual “friend vs. foe” response wherein a resting dog responds to movement in its visual field with either (1) a cautious growl while remaining in a lying position, (2) threatening bark while sitting up, or (3) happy breathing and tail-wagging while fully standing. A human actor was told that he/she is visiting a home with a dog unknown to him/her. As shown in Figure 1, a robotic dog was projected in pseudo-3D onto the forward screen, with external sensors that enable its simulated brain to “see” and respond to the actor’s movements, in the context of a background scene projected onto the rear screen (for
In this demonstration, we used a static image of a suburban neighborhood. The robot's eyes (a tracking pan-tilt-zoom camera) and ears (monaural or spaced stereo microphones) capture the actor's movements and voice in the context of the background scene, which is projected independently (and may contain moving elements, including other animals or actors). The BRAINSTEM is a supercomputer running threads that synchronously (1) capture and preprocess video images, sound, and touch, (2) convert preprocessed sensory images into probabilities of spiking for each primary neocortical region, (3) upload spike probability vectors to the BRAIN simulator, (4) then from the BRAIN simulator accept motor neuron region output spike density vectors and trigger corresponding dominant motor sequences (e.g., for the virtual dog robot: sitting, lying, barking, walking) via the robotic simulator program (Webots/URBI), which makes the corresponding changes in behavior of the projected robot (and incorporates internal sensation such as proprioception and balance). The BRAIN simulator is a neuromorphic modeling program running on a supercomputer, executing a pre-specified spiking brain architecture, which can adapt as a result of learning (using reward stimuli offered by the ACTOR's voice or stroking of the touch pad).

Based on successful performance, researchers iteratively "plug in" alternative or more complex brain architectures. A proposed enhancement would be to couple live in vivo or in vitro neural tissue (BRAIN SLICE) to the brain simulation using multielectrode arrays and optical imaging, in order to continuously calibrate and constrain synthetic brain dynamics.

The simple neuromorphic brain consisted of 64 single-compartment neurons divided into four columns representing pre-motor regions (precursors to coordinated behavioral sequences), each connected to one of the visual field preferences based on Gabor filter configurations. According to the probability vector received from BRAINSTEM, NCS injected short (1 ms) step current (3 nA) pulses sufficient to reach the threshold of −50 mV and generate a single spike. Membrane voltages updated at a frequency of 1 kHz. The ACTOR in this scenario was told in advance that moving vertically-oriented objects (including body parts) will pose a threat to the robot, whereas moving horizontally-oriented objects will be perceived as friendly gestures; the actor was free to choose any sequence of movements in response to the perceived intent of the ROBOT. ROBOT behavioral sequences are triggered when the neuromorphic BRAIN output to BRAINSTEM has 50 ms of consistent spiking in one pre-motor region compared with another. Periods without domination of one pre-motor region over another trigger the ROBOT to lie down and growl. In cell rasters, each row represents the timing of action potentials (spikes) of a single neuron; darker gray markers indicate clustered bursts of spikes. Figure 2 shows pre-motor action potential spike rasters from a typical 10-s VNR interaction with a human actor. A video is available online at http://brain.unr.edu/VNR.

RELATED WORK

Social embeddedness is a key characteristic of the proposed VNR approach, with an emphasis similar to that received initially in the stepwise, ontological development of robotic cognition (Breazeal and Scassellati, 2000; Brooks et al., 1998) and more recently in epigenetic robotics research focused on the interaction between cognitive and perceptual brain systems (Lungarella and Berthouze, 2002; Schlesinger, 2003). In order to map behavior to robotic cognition, almost all of these models rely on combinations of psychological production rules, fitness functions, and machine learning algorithms. Notably, this includes models aimed at capturing neuronal epiphenomena such as mirror neuronal activity (Triesch et al., 2007). Our proposed approach is different in several ways. First, we focus on understanding brain physiology at the "mesocircuit" level, relying on social-emotional robotics to reduce the multitude of potential architectures that could bridge the measurements at the cellular level (e.g., patch clamp and unit recordings) with those at the scales of millions of cells (e.g., optical and fMR imaging). Second, because the stipulation of neuromorphic architecture excludes the use of production rules or hierarchical algorithms as
psychological models, any assumptions on motivation, intentionality and behavioral triggering must emerge from the tissue models themselves, and learning from behavioral reinforcement must manifest as synaptic change. Third, since realistic social interaction requires temporal coherence between the simulated robotic brain and that of the human actor, the simulation must incorporate the actual distribution of physiological time constants that characterize membranes, channels, and synapses.

Due to the distinguishing characteristics of the proposed VNR approach, there is limited research work similar to ours. Some groups have reported success in navigational tasks using neuromorphic architectures (Banquet et al. 2005; Cuperlier et al., 2005; Krichmar et al., 2005; Ogata et al., 2004; Wiener and Arleo, 2003). Notable endeavors that share similarities with our work include identification of challenges and opportunities in robot-mediated neurorehabilitation (Harwin et al., 2006), development of prototypes that combine robotics and virtual reality to assist in the rehabilitation of brain-injured patients and support motor control research (Patton et al., 2006), and generation of artificial brains for virtual robots using a new paradigm based on the epigenetic approach (Pasquier, 2004, 2005).

DISCUSSION AND WIDER IMPLICATIONS

The rationale for the virtual paradigm in VNR is rooted fundamentally on engineering and human-computer interface considerations, and is similar to that put forward by Krichmar and Edelman (2005) for robotic instantiation of brain-based devices. Certainly, a closed-loop system could incorporate either real or virtual robots. In our VNR framework, however, we emphasize virtuality for the following reasons: (1) the human actor must find the robotic behavior believable; it is our impression that refined neurorobotic avatars are more readily accepted (perhaps due to the popularity of online virtual reality networking) than clumsy, unreliable physical robotic prototypes; (2) as investigators design and grow more complex neuromorphic brains, robotic behaviors will require additional sensory, motoric, and emotional sophistication, which in turn may entail major changes in the robot's body parts and dimensions, and degrees of freedom of joints and face—all of which can be accelerated using software (often in just hours) without the delays and costs of added hardware and its engineering; and, (3) at stages of neuromorphic development at which it would be important to demonstrate the functionality of a physical robot, the software API can be compiled and transferred to a prototype of the hardware robotic system (provided that VNR simulator used a realistic control API).

The use of human actors in the VNR approach might be seen as an obstacle in terms of time, resources, and variability. However, there is no other “gold standard” for realistic, spontaneous, emotionally intelligent interaction. Moreover, it is human-level cognition that we explore and seek to elucidate in our modeling and applications. In addition, the parameters for neuronal membranes, channels, and synapses are given as time constants on the order of milliseconds to seconds, as co-optimized by evolution. This means that, for example, a system that emulates connected neurons but operates at the temporal scale of microseconds cannot interact with the slower responses of humans. Therefore, both the joint distribution of known biological time constants and the need for emotionally intelligent responses require the use of a closed-loop interaction of the brain prototype with an actor. As an alternative, one might consider using animals in place of humans; however, animals rely on many subtle biological sensory cues such as smell, so will readily accept neither embodied nor virtual robots as socially interactive partners.

The VNR approach opens exciting and promising avenues of future research and application. For example, within the Webots/URBI environment we are currently developing a social-emotional humanoid robot with functional capabilities motivated by the MDS (mobile, dexterous, social) robot under development by the Personal Robotics Group of the MIT Media Lab (http://robotic.media.mit.edu). Our robot will incorporate language understanding and production using corresponding neocortical models based on praise and curiosity. We are also working on a related model of childhood autism.
We also plan to calibrate and constrain synthetic brain dynamics by coupling live in vivo (acute slice or sustained culture) or in vitro neural recordings to the brain simulation using multi-electrode arrays and optical stimulation and imaging.

The envisioned impact of the proposed approach is wide-reaching. First, neuroscience research is expected to directly benefit from VNR in terms of development and validation of new, expandable brain models and architectures as well as study and exploration of various brain disorders and injuries, including strokes and genetic disorders. Second, faster progress in a variety of medical applications areas will likely be enabled by VNR-based research, primarily in terms of advancements in neuroprosthetics and new solutions for brain-related assistive technologies. Third, a diversity of other application areas traditionally propelled by developments in artificial intelligence could take advantage of VNR method and tools. These include, but are not limited to, decision-making support driven by human-like behavior and motivation, enhanced robotics-centered navigation and security, and better understanding in the fields of neural development, neurophysiology, and neuropathology.

**CONFLICT OF INTEREST STATEMENT**

The authors declare that this research was conducted without any commercial or financial relationships that could be construed as a potential conflict of interest.

**ACKNOWLEDGMENTS**

This work was supported by grants from the U.S. Office of Naval Research (grants N000140010420 and N000140510525).

**REFERENCES**

Allen Institute (2007). Allen Brain Atlas, URL: http://www.brain-map.org.

Banquet, J. P., Ghausser, P., Quay, M., Revel, A., and Burmod, Y. (2005). A hierarchy of associations in hippocampo-cortical systems: cognitive maps and navigation strategies. *Neural Comput.* 17, 1339–1384.

Breaalal, C., and Sassenelati, B. (2000). Infant-like social interactions between a robot and a human caretaker. *Adapt. Behav.* 8, 49–73.

Brooks, R., Breaalal, C., Marjanovic, M., Scassellati, B., and Williamson, M. (1998). The Cog project: building a humanoid robot. In *Computation for Metaphors, Analogy, and Agents*, C. Nehaniv, ed (New York: Springer), pp. 52–87.

Cuperlier, N., Quay, M., Laroque, P., and Gausser, P. (2005). Transition cells and neural fields for navigation and planning. *Comput. Sci.* 5561, 346–355.

Dauthenhahn, K. (2007). Socially intelligent robots: dimensions of human-robot interaction. *Phil. Trans. R. Soc. B.* 362, 679–704.

Goodman, P. H., Buntha, S., Zou, Q., and Dascalu, S. (2004). Spike-timing dynamics of neuronal groups. *Cereb. Cortex* 14, 933–944.

Krichmar, J. L., and Edelman, G. M. (2005). Brain-based devices for the study of nervous systems and the development of intelligent machines. *Artif. Life* 11, 63–77.

Krichmar, J. L., Seth, A. K., Nit, D. A., Fleischer, J. G., and Edelman, G. M. (2005). Spatial navigation and causal analysis in a brain-based device modeling cortical-hippocampal interactions. *Neuroinformatics* 3, 197–221.

Lungarella, M., and Berthouze, L. (2002). On the interplay between morphological, neural, and environmental dynamics: a robotic case study. *Adapt. Behav.* 10, 223–241.

Markham, H., Lubke, J., Frotscher, M., and Sakmann, B. (1997). Regulation of synaptic efficacy by coincidence of postsynaptic APs and EPSPs. *Science* 275, 213–225.

Ogata, T., Sugano, S., and Tani, J. (2004). Intrinsic motivation systems for autonomous mental development. Innovations in applied artificial intelligence. *Lect. Notes Comput. Sci.* 3029, 435–444.

Oudeyer, P. Y., and Kaplan, F. (2007). What is intrinsic motivation? *A typology of computational approaches. Neuro Robotics* 1, 6. doi: 10.3389/neuro.12.006.2007.

Oudeyer, P. Y., Kaplan, F., and Hafner, V. V. (2007). Intrinsic motivation systems for autonomous mental development. *IEEE Trans. Evol. Comput.* 2, 265–286.

Pasquier, J. L. (2004). *NAC*, an artificial brain for an epigenetic robot. *Proc. World Automation Cong.* 15, 535–540.

Pasquier, J. L. (2005). Applying the GEM prospecitive paradigm to the autonomous and adaptable control of a virtual robot. *MICAl 2005: advances in artificial intelligence. Lect. Notes Comput. Sci.* 3789, 959–969.

Patton, J., Dawe, G., Scharver, C., Massa-Ivaldi, F., and Kenyon, R. (2006). Robotics and virtual reality: a perfect marriage for motor control research and rehabilitation. *Assist. Technol.* 18(2), 181–193.

Ripplinger, M. C., Wilson, C. J., King, J. G., Frye, J., Drewes, R., Harris, F. C., and Goodman, P. H. (2004). Computational models of interacting brain networks. *J. Investig. Med.* 52, 155.

Samejima, K., and Doya, K. (2007). Multiple representations of belief states and action values in cortico basal ganglia loops. *Ann. NY Acad. Sci.* 1104, 213–228.

Schneiherger, M., Schermerhor, P., Kramer, J., and Anderson, D. (2007). First steps toward natural human-like HRI. *Auton. Rob.* 22, 411–423.

Schindler, K. A., Goodman, P. H., Wieser, H. G., and Douglas, R. J. (2006). Fast oscillations trigger bursts of action potentials in neocortical neurons in vitro: a quasi-white-noise analysis study. *Brain Res.** 1110, 201–210.

Schlesinger, M. (2003). A lesson from robotics: modeling infants as autonomous agents. *Adapt. Behav.* 11, 97–107.

Schweighofer, N., Tanaka, S. C., and Doya, K. (2007). Serotonin and the evaluation of future rewards-theory, experiments, and possible neural mechanisms. *Ann. NY Acad. Sci.* 1104, 289–300.

Society for Neuroscience (2007). Databases Having Experimental Data, URL: http://ndg.sfn.org.

Teeters, J. L., Harris, K. D., Millman, K. J., Ohlhausen, B. A., and Sommer, F. T. (2008). Data Sharing for Computational Neuroscience. *Neural Informatics* 6, 47–55. doi: 10.1007/s12021-008-9009-y. URL: http://www.crcns.org.

Toledo-Rodriguez, M., Blumenfeld, B., Wu, C. Z., Luo, J. Y., Attali, B., Goodman, P., and Markham, H. (2004). Correlation maps allow neuronal electrical properties to be predicted from single-cell gene expression profiles in rat neocortex. *Cereb. Cortex* 14, 1310–1327.

Toledo-Rodriguez, M., Goodman, P., Illsc, M., Wu, C. Z., and Markham, H. (2005). Neuropeptide and calcium-binding protein gene expression profiles predict neuronal anatomical type in the juvenile rat. *Physiol. (Lond.*) 567, 401–413.

Triesh, J., Jasso, H., and Deak, G. O. (2007). Emergence of mirror neurons in a model of gaze following. *Adapt. Behav.* 15, 149–165.

Wang, Y., Markham, H., Goodman, P. H., Berger, T. K., Ma, J. Y., and Goodman-Rakic, P. S. (2006). Heterogeneity in the pyramidal network of the medial prefrontal cortex. *Nat. Neurosci.* 9, 534–542.

Wiener, S. I., and Arleo, A. (2003). Persistent activity in limbic system neurons: neurophysiological and modeling perspectives. *J. Physiol. (Paris)* 97, 547–555.

Received: 10 April 2008; accepted: 04 June 2008.

Citation: Front. Neurosci. (2008) 2, 1; 123–128, doi: 10.3389/neuro.01.007.2008

Copyright © 2008 Goodman, Zou and Dascalu. This is an open-access article subject to an exclusive license agreement between the authors and the Frontiers Research Foundation, which permits unrestricted use, distribution, and reproduction in any medium, provided the original authors and source are credited.