Neural Networks 32 (2012) 130–137

Contents lists available at SciVerse ScienceDirect

Neural Networks

journal homepage: www.elsevier.com/locate/neunet

2012 Special Issue

Real-time human–robot interaction underlying neurorobotic trust and intent recognition

Laurence C. Jayet Bray [a,b,*], Sridhar R. Anumandla [b], Corey M. Thibeault [b,c], Roger V. Hoang [b], Philip H. Goodman [d], Sergiu M. Dascalu [b], Bobby D. Bryant [b], Frederick C. Harris Jr. [a,b]

[a] Brain Computation Lab, University of Nevada, Mail stop 456, Reno, NV 89557, USA
[b] Department of Computer Science and Engineering, University of Nevada, Reno, NV 89557, USA
[c] Department of Electrical and Biomedical Engineering, University of Nevada, Reno, NV 89557, USA
[d] Brain Computation Lab and Internal Medicine, University of Nevada, Reno, NV 89557, USA

ARTICLE INFO

This paper is dedicated to the late Dr. Philip H. Goodman, our mentor and friend who initiated work on brain modeling and virtual neurorobotics at The University of Nevada, Reno

Keywords:
Real-time computing
Virtual neurorobotics
Learning
Trust and intent recognition

ABSTRACT

In the past three decades, the interest in trust has grown significantly due to its important role in our modern society. Everyday social experience involves "confidence" among people, which can be interpreted at the neurological level of a human brain. Recent studies suggest that oxytocin is a centrally-acting neurotransmitter important in the development and alteration of trust. Its administration in humans seems to increase trust and reduce fear, in part by directly inhibiting the amygdala. However, the cerebral microcircuitry underlying this mechanism is still unknown. We propose the first biologically realistic model for trust, simulating spiking neurons in the cortex in a real-time human–robot interaction simulation. At the physiological level, oxytocin cells were modeled with triple apical dendrites characteristic of their structure in the paraventricular nucleus of the hypothalamus. As trust was established in the simulation, this architecture had a direct inhibitory effect on the amygdala tonic firing, which resulted in a willingness to exchange an object from the trustor (virtual neurorobot) to the trustee (human actor). Our software and hardware enhancements allowed the simulation of almost 100,000 neurons in real time and the incorporation of a sophisticated Gabor mechanism as a visual filter. Our brain was functional and our robotic system was robust in that it trusted or distrusted a human actor based on movement imitation.

© 2012 Elsevier Ltd. All rights reserved.

1. Introduction

Computational neuroscience is an interdisciplinary science, which combines diverse fields, especially neuroscience, and computer science and engineering. It uses computational approaches to investigate properties of the central and peripheral nervous systems at different levels of detail (De Schutter, 2008). It involves cycling back and forth between high-performance computing, experimental data recording and complex brain modeling to further construct real-time intelligent systems and understand neurological disorders.

There are a variety of neural simulators available to computational neuroscience researchers today. Although many share similar approaches and features, most have their unique qualities. The two well-known simulators that allow compartmental modeling are: NEURON, which is a simulation environment for creating and using empirically based models of biological neurons and neural circuits (Brette et al., 2007; Carnevale & Hines, 2006) ; GENESIS (the General Neural Simulation System), which is an extensive general simulation system for realistic modeling of neural and biological systems (Bower & Beeman, 1998; Brette et al., 2007). There is also a variety of spiking neural network simulators, such as BRIAN (Goodman & Brette, 2009), NEST (Gewaltig & Diesmann, 2007), PyNN (Davison et al., 2008). Even though these simulators have their own strengths, they do not have the capability for real-time simulations, which is very important for artificial intelligence (AI) and robotic interfaces. The latest spiking neural network was developed by Richert, Nageswaran, Dutt, and Krichmar (2011), and it generates large-scale computational models in real-time. However, it runs on single Graphic Processing Units (GPUs) (Richert et al., 2011).

* Corresponding author at: Brain Computation Lab, University of Nevada, Mail stop 456, Reno, NV 89557, USA. Tel.: +1 864 506 5489.
E-mail addresses: ljayet@gmail.com (L.C.J. Bray), sridhar.anumandla@gmail.com (S.R. Anumandla), corey@cmthibeault.com (C.M. Thibeault), rvhoang@gmail.com (R.V. Hoang), dascalus@cse.unr.edu (S.M. Dascalu), bdbryant@cse.unr.edu (B.D. Bryant), fred.harris@cse.unr.edu (F.C. Harris Jr.).

0893-6080/$ – see front matter © 2012 Elsevier Ltd. All rights reserved.
doi:10.1016/j.neunet.2012.02.029

Here we have utilized the updated version of our NeoCortical Simulator (NCS) (Brette et al., 2007) to compute integrate-and-fire, conductance-based synaptic neurons of brain regions involved in human emotions such as trust. The Virtual NeuroRobotic (VNR) system developed in our laboratory (Goodman, Buntha, Zou, & Dascalu, 2007; Goodman, Zou, & Dascalu, 2008) was also used to rapidly engineer a robotic system that can interact with humans.

The field of social robotics has been focused on better understanding the important dynamics of human emotion (Dautenhahn, 2007; Goodman et al., 2008; Scheutz, Schermerhor, Kramer, & Anderson, 2007). For many decades, intelligent systems have tried to replace a human mind in planning, learning new functions, and making decisions, but some traits such as social skills have been hard to replicate. Breazeal et al. have focused on developing creating socially intelligent robot partners that can interact with humans. Also, Thomaz et al. have been interested in machines learning new goals from people. These machine intelligent systems can effectively cooperate with people, but they are still socially limited (Thomaz & Breazeal, 2008). Therefore, it is tempting to try to incorporate realistic neuromorphic properties (hypothalamic oxytocin release) to demonstrate how an emotional behavior such as trust can be established.

A variety of external cues are involved in trust, but we mostly concentrate on imitation. This recurrent mechanism has been shown in human behavior from early ages to adulthood and is associated with cooperation and trust (Saavedra, Smith, & Reed-Tsochas, 2010). Observations of social interactions between two or more people have shown that humans unconsciously and unintentionally learn to imitate trustworthy behaviors. Several laboratories have worked on human–robot interaction, especially on mimicking behaviors, such as children and robots playing hide and seek (Trafton et al., 2006) or learning from demonstration (Koenig, Takayama, & Mataric, 2010). Recently, Paukner, Suomi, Visalberghi, and Ferrari (2009) showed that capuchin monkeys display affiliation toward humans who imitate and spend more time interacting with them (Paukner et al., 2009).

Intent recognition has been a wide research area in robotics since it is very important for developing social robots that can cooperate with humans in performing tasks (Kelley, 2009; Tahboub, 2005). Often, the intention of a person can be represented explicitly by visual or auditory stimuli. For social robots to operate in the human world, intent recognition plays a key role and the importance of it can be understood from real world examples such as: for a robot-controlled car to travel on roads, it needs to understand the intentions of other drivers and also let the other drivers know about its own intentions by giving appropriate signals to avoid accidents; during military operations such as anti-terrorist activities robots have to determine the intentions of the enemy in order to prevent them from performing illegal activities; if robots are supposed to guard the society from anti-social elements then there is a crucial need for them to sense the intentions of people in order to resolve whether anyone is going to create any damage to the society; and, for a robot to assist a human in industrial work it needs to know what he/she is expecting from it.

Oxytocin has been found to be responsible for developing human trust (Baumgartner, Heinrichs, Vonlanthen, Fischbacher & Fehr, 2008; Heinrichs, von Dawans, & Domes, 2009; Huber, Veinante, & Stoop, 2005; Kosfeld, Heinrichs, Zak, Fischbacher, & Fehr, 2005; Zak, Kurzban, & Matzner, 2005). It is a neuropeptide mammalian hormone that is known to primarily be produced by the hypothalamus and acts as a neurotransmitter in the brain. It is considered to be responsible for certain physiological functions in females such as oxytocin serves to stimulate and regulate continuous contractions of smooth muscle tissue of the uterus during labor, milk letdown for breastfeeding in lactating women, pair bonding, maternal care, and sexual behavior (Anderson-Hunt & Dennerstein, 1995; Carter, 2003; Ferguson, Young, & Insel, 2002; Lim & Young, 2006; Young & Wang, 2004). Apart from these well-known functions oxytocin is also considered to be responsible for developing trust in humans (Baumgartner et al., 2008; Heinrichs et al., 2009; Huber et al., 2005; Kosfeld et al., 2005; Zak et al., 2005).

We speculate that the role of hypothalamic oxytocin and amygdala is essential in simulating complex neuromorphic brain models of social cognition involving short term memory (Domes, Heinrichs, Michel, Berger, & Herpertz, 2007; Rimmele, Hediger, Heinrichs, & Klaver, 2009), trust (Baumgartner et al., 2008; Kosfeld et al., 2005; Zak et al., 2005) and suppression of fear (Kirsch et al., 2005; Petrovic, Kalisch, Singer, & Dolan, 2008). An increase in the level of oxytocin establishes trust by suppressing activity in the amygdala (Domes et al., 2007; Kirsch et al., 2005; Petrovic et al., 2008), which is present in the medial temporal lobe and is responsible for social cognition and fear in mammals (Choleris et al., 2007; Ohman, 2005; Whalen et al., 2004).

Based on these observations, our goal was to model a neuromorphic brain and implement it on a virtual neurorobot, which could learn to trust a human if its actions were imitated by a human actor. To the best of our knowledge, we propose the first real-time simulated biological brain model (including visual, parietal, premotor and inferotemporal cortices, hypothalamus, and amygdala) for human–robot interaction.

2. Computational design

2.1. Software platforms

This model included leaky integrate-and-fire neurons with conductance-based synapses. The simulations required adequate computer power, and therefore were performed using our NeoCortical Simulator (NCS) (Brette et al., 2007; Drewes, Zou, & Goodman, 2009; Wilson, Goodman, & Harris, 2001). Each integrate-and-fire neuron is characterized by a membrane time constant of 20 ms, a membrane resistance of 100 MΩ, and a resting membrane potential of −60 mV. Whenever the membrane potential crosses the spiking threshold of −50 mV, an action potential is generated, and the membrane potential is reset to the resting potential, where it remains clamped for a 5 ms refractory period. Membrane voltages are updated at each time step (1 ms) as follows:

$$C_N \frac{dV}{dt} - I_{AHP} - I_{syn} - I_{ext} + I_{leak} = 0.$$

With small ionic currents (I_{AHP}), which contribute to the membrane voltage by controlling spike-frequency adaptation; synaptic currents (I_{syn}); external input currents (I_{ext}); and a voltage-independent leakage current (I_{leak}).

Spike-timing-dependent plasticity (STDP) was defined as:

$$W(\triangle t) = \begin{cases} A_+ \exp\left(\dfrac{\triangle t}{\tau_+}\right) & \text{if } \triangle t < 0 \\ -A_- \exp\left(\dfrac{-\triangle t}{\tau_-}\right) & \text{if } \triangle t \geq 0. \end{cases}$$

With the maximum amounts of synaptic modification (A); the positive and negative windows (Δt); and positive and negative decay constants (τ) (Caporale & Dan, 2008; Dan & Poo, 2004; Markram, Gerstner, and Sjöström, 2011; Song, Miller, & Abbott, 2000; Zhang, Tao, Holt, Harris, and Poo, 1998).

NCS has focused on increasing simulation speed and improving the functionality of our brain models. This optimization work has produced better than sevenfold sequential speedup over previous versions and decreased the memory footprint by almost 90%, while shrinking our code base by more than 25%. Currently a

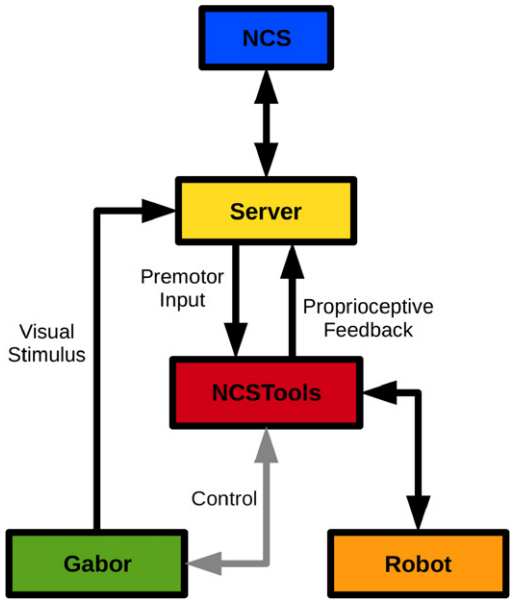

Fig. 1. Robotic system configuration. Closed loop robotic system with its five major components.

single compute node can run a simulation with 35,000 cells and approximately 6.1 million synapses using 72% of a 4 GB memory. Memory use per node is approximately halved as the number of nodes is doubled. Each node requires only a few tens of KB overhead for each other node, so the practical upper bound on the size of a brain we can create is determined by the available hardware.

2.2. Hardware platforms

We now use two clusters: one consisting of 208 Opteron cores, 416 GB of RAM, and more than a Terabyte of disk storage; and the other one consisting of 4 SUN 4600 machines (16-processors each) connected via Infiniband with 128 GB Ram per machine, and 24 Terabytes of disk storage. This has allowed us to look at large scale NCS models, and begin explore optimizations for shared-memory parallelism.

2.3. Robotic system

The virtual robotic system for this project was designed around a number of components unique to NCS and the VNR paradigm (Whalen et al., 2004). The neural simulation was executed on a remote computing cluster and was networked to the other system components using our Brain Communication Server (BCS) (Choleris et al., 2007), a publish–subscribe server developed specifically for integration with NCS. The closed-loop robotic system is shown schematically in Fig. 1, and the other major components are described below.

NCSTools - is a software system written in C++ that provides a number of mechanisms for communicating in real-time with a running NCS simulation (Goodman et al., 2007). NCSTools accepts plain text strings from clients connected through a built-in socket server. Through a custom configuration file, users can then assign these strings to define the input stimulus or simulation controls for a running NCS instance. Similarly, NCSTools can be configured to process simulation reports in a number of different ways, the results of which can be sent to connected clients through the server interface. With this mechanism designers of remote tools can interface with neural simulations in a way that abstracts them from the details of the model, ultimately allowing reuse of code without modification for different models (only the NCSTools configuration needs to be changed).

The NCSTools server monitors the robotic avatar and creates the appropriate stimulus for proprioceptive feedback and premotor movement to replicate the role of a biological brainstem. Similarly, the NCSTools software receives spiking information from the premotor region of the neural simulation. This activity is monitored, and when a configured threshold is reached the appropriate command is sent to the robotic avatar, initiating the appropriate motion.

Virtual robotic interface - was constructed using Webots 5 (Cyberotics, Lausanne, Switzerland). Motions were programmed in C++ using the provided interfaces and the communication was accomplished using the NCSTools C++ client.

Gabor filter -is considered one of the better representations of mammalian visual receptive fields (Jones & Palmer, 1987). Often used for edge detection, the Gabor filter can spatially filter an image by frequency and orientation (Mehrotra, Namuduri, & Ranganathan, 1992). Although Gabor filters provide a good approximation of human visual processing, the computational cost is often too high for real-time applications. To reduce the processing time, a GPU-based Gabor filter application was developed. The Gabor processing application is designed around NVidia's CUDA GPU programming environment. CUDA provides mechanisms for GPU algorithm development with an emphasis on high-performance applications. Processing for our simulation begins with the capture of a 320×240 pixel image at a frequency that is determined by the user during configuration. A 128×128 pixel area of interest concentrated on the image is then selected and grayscale information is extracted. The area of interest is differenced with the previous frame. This differencing provides an immediate representation of the motion between the two frames. The differenced image is then padded with zeros to 256×256 pixels and a fast Fourier transform (FFT) is computed. The frequency space/domain image is processed with a pre-computed complex Gabor kernel and the inverse FFT is computed. The original 128×128 pixel area of interest is extracted and segmented into a user-defined number of regions. Each resulting values, especially the ones that are not in the range of [0, 1] are normalized using the convolution theorem. These values are averaged for each region and sent directly to the running NCS brain simulation as a stimulus using the network interface.

3. Behavioral scenario

In our application, a virtual humanoid neurorobot stood behind a table holding a yellow rod, and this virtual environment was projected onto a large screen. The human participant sat in front of a camera placed on a table and held a rod similar to the one present in the virtual environment. The screen and table were arranged so that the two subjects faced one another. A dumbbell was placed on the table, and a similar object was loaded into the virtual environment at a later phase of the experiment.

The experiment consisted of two phases, a learning phase and a challenge phase. These were integrated into a single interaction between the human participant and the virtual neurorobot. During the learning phase the neurorobot was configured to perform sequences of vertical or horizontal motions with the rod for 5 s at a time. During this time the motion of the interacting human was captured by the camera and processed by the system described in Section 2.3. If the human matched the neurorobot motion (concordant motion) the brain model learned to trust the participant. However, if the human performed the opposite motion (discordant motion) then the brain model caused the

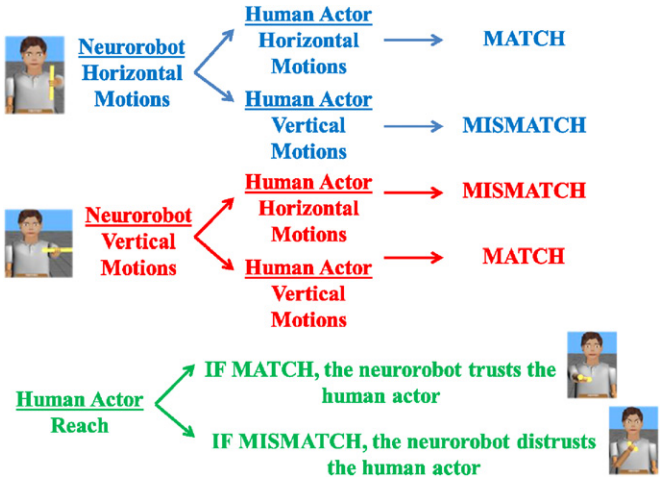

Fig. 2. Behavioral paradigm of the instinctual trust-the-intent scenario. During the learning phase, the human matched or mismatched the virtual neurorobot's horizontal or vertical motions. During the challenge phase, the human reached for an object to test the trust instinct of the neurorobot.

neurorobot to build distrust for the human participant. A transition to the challenge phase could be triggered either manually by someone monitoring the experiment or automatically by a signal from the brain simulator. This transition would redirect the visual information being captured by the camera to a different region of the brain model designed to control a reaching movement. During this phase the human reaches for the dumbbell, and if the neural model has learned to trust the participant the robot hands over the corresponding object in the virtual environment. If the neural model has not learned to trust the human participant, the robot grabs the object and pulls it away from the participant. This behavioral scenario is illustrated in Fig. 2.

4. Neuromorphic brain architecture

The computational neuromorphic brain architecture consisted of small regions of the visual (VC), parietal (PR), inferotemporal (IT), and pre-motor (PM) cortices and the hypothalamus (HYP) and amygdala (AMY) limbic systems, as shown in Fig. 3. These regions were represented by 12,000, 16,000, 8000, 8000, 26,000, and 13,000 neurons, respectively. This an order of magnitude larger, which improves the functionality of the brain compared to our previous model for this task (Anumandla et al., 2011). We used a realistic four-compartment triple apical dendrite model for the hypothalamic oxytocin cells, where the three dendrites received stimuli from columns in the visual, parietal, or inferotemporal cortex, one for each. The three dendrites together stimulated the hypothalamus, and thus the synapses connecting the inferotemporal cortex to the hypothalamus learned to trust a human.

The visual cortex was composed of three columns (reach, horizontal, and vertical) with 4000 neurons each (2000 excitatory and 2000 inhibitory). It received input from the external camera, which captured the human participant's actions, and differentiated horizontal, vertical and reach motions from one another.

The parietal cortex was composed of four columns (horizontal, vertical, trust, and distrust) with 4000 neurons each (2000 excitatory and 2000 inhibitory). It received input from the neurorobot, based on its motions, and also made the distinction between trust and distrust.

The inferotemporal cortex was composed of two columns (E and I pools) with 8000 neurons (4000 excitatory and 4000 inhibitory). It had direct projections from and to the hypothalamus, and reinforced the learning synapses for concordant motions.

The hypothalamus was composed of two columns (horizontal, and vertical) with 13,000 neurons each (10,000 excitatory and 3000 inhibitory). After being activated by VC, PR, and IT the hypothalamus fully allowed (or not) the inhibition of the amygdala.

The amygdala was composed of two columns (E and I pools) with 13,000 neurons (10,000 excitatory and 3000 inhibitory). It was stimulated by a monotonic current to replicate biological sensory stimuli and by the hypothalamus, and sent information to the parietal cortex for either trusting or distrusting the human participant.

The premotor cortex was composed of two columns (trust, and distrust) with 4000 neurons each (2000 excitatory and 2000 inhibitory). It received input from the parietal cortex, and triggered the action for keeping (distrust) or giving (trust) the object to the human participant.

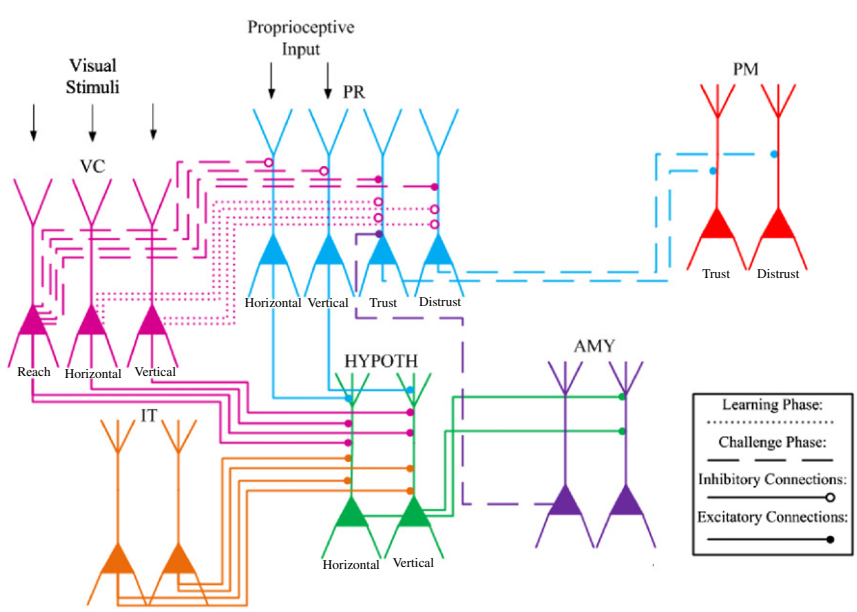

Fig. 3. Computational neuromorphic brain architecture of the instinctual trust-the-intent scenario. During the learning phase, the dominant pathway depends on vertical or horizontal motions (dotted line) while during the challenge phase, it depends on the reaching action (dashed line). (VC: Visual Cortex; PR: Parietal Cortex; PM: Premotor; IT: Inferotemporal; HYPOTH: Hypothalamus; AMY: Amygdala).

Table 1
Description and analysis of the computational neuromorphic brain.

A	Model summary		
Populations	Inferotemporal cortex (IT), visual cortex (VC), parietal cortex (PR), premotor cortex (PM), hypothalamus (HYP), and amygdala (AMY)		
Neuron model	Leaky integrate-and-fire, fixed threshold, refractory time		
Plasticity	STDP		
Synapse model	Conductance-based		
Measurements	Membrane potential		
B	Extrinsic connectivity		
Type	Description	Probability (%)	Conductance (mS)
VC/PR/IT–HYP	Visual, parietal, and inferotemporal input to the hypothalamus dendrites	10	0.001–0.005
VC–PR	After the visual cortex receives input from the activity is projected to the parietal cortex	10	0.08
HYP_E–AMY_I	The hypothalamus activates interneurons to suppresses or not the amygdala	30	0.003
HYP–IT	Feedback loop from the hypothalamus to the inferotemporal cortex	10	0.005
AMY_E–PR_I	If the amygdala is not suppressed, information is sent to the parietal cortex	10	0.3
PR–PM	The parietal cortex is directly connected to the premotor cortex	10	0.01
C	Intrinsic connectivity		
VC	Visual cortex internal connections	10	0.09
PR_I–PR_E	Parietal cortex internal connections	10	0.03
AMY_I–AMY_E	Interneurons suppressed the amygdala activity based on the hypothalamus firing	10	0.008
D	External inputs		
Visual stimuli	The visual cortex receives input from the external camera through the Gabor filter which captures human motions	10	Amp = 1 mA
Proprioceptive	The parietal cortex receives input from the neurorobot based on its motions	10	Amp = 1 mA
Monotonic	The amygdala receives monotonic current to emphasize sensory stimuli	50	Amp = 1 mA

The primary mechanism was that the hypothalamus stimulated interneurons (AMY_I), which suppressed the amygdala (AMY_E). If AMY_E was completely suppressed, interneurons (PR_I) between AMY and PR were never activated, and thus never suppressed "trust" or PR_E (in the case of concordant motions). However, if AMY_E was not suppressed because of the lack of activity (and learning) in the hypothalamus, then "trust" or PR_E was suppressed (in the case of discordant motions). Depending on whether the parietal (PR_E) trust or distrust column was activated, a corresponding pre-motor column (PM) was triggered. Consequently, a motor command was sent to the Brainstem subsystem (Peng, 2006), which directed it in turn to the neurorobot.

STDP (Please see equations in Section 2.1) was used in the hypothalamic synapses to reinforce learning as the hypothalamus received a combination of signals from the visual and the parietal cortices at the same time during concordant motions, and at different times during discordant motions. This reinforcement was also accentuated by a feedback loop with the inferotemporal cortex.

For a better understanding of the network, the description and the analysis of the neuromorphic brain architecture are given in Table 1.

5. Results

After the stimulation (described in Section 4) was injected, the activity in the visual and parietal cortices corresponding to the human and robot motions along with the hypothalamic firing increased, as shown in Fig. 4. Significant firing rates (up to 80 Hz above baseline) occurred in the hypothalamus when the activities of the visual and parietal cortices overlapped in time (Fig. 4, horizontal and vertical concordant). However, insufficient consistent firing occurred in the hypothalamus because of discordant activities in the visual and parietal cortices, even though there was the same stimulus from the inferotemporal cortex (Fig. 4, horizontal and vertical discordant). The feedback loop between the hypothalamus and the inferotemporal cortex, in addition to STDP, reinforced these phenomena for learning purposes.

Since the visual, parietal, and inferotemporal cortices connected to the independent dendrites of the hypothalamus, firing occurred only when all three dendrites integrated and fired together. Spiking in the independent dendrites of the hypothalamus occurred in the case of concordant motions as the neurorobot learned to trust the human. However, in the case of discordant motions dendritic firing within the hypothalamus decreased.

During the reaching movement, a relatively significant firing rate (up to 30 Hz above baseline) occurred in only one of the parietal decision-making cortical columns, which depended on the types of motion (concordant or discordant) during the learning phase. Fig. 5 (left, concordant) shows higher firing rates in the parietal trust column when compared to the distrust column during concordant robot–human motions. In this figure, the firing started between 20 and 25 s after the end of the learning phase since a reach for an object was performed at that time. In Fig. 5 (right, discordant), the parietal distrust column fired more when compared to the trust column after discordant robot–human motion in the learning phase.

Additionally, the synaptic weight distributions of the vertical and horizontal columns for both concordant and discordant motions over both phases differed greatly. Whether the motion was vertical or horizontal, there was a significant increase in the synaptic weight during concordant motions, as shown in Fig. 6. It was twice as high as it was for discordant motions due to STDP present in the hypothalamus synapses. Also, in the case of concordant robot–human motions, after significant learning has occurred, the firing in the hypothalamus was strong enough to inhibit the activity in the amygdala, which thus never shut down trust in the parietal cortex. However, in the case of discordant robot–human motions, because there was insufficient or no firing in the hypothalamus, the activity in amygdala remained active, and trust never occurred.

By modeling 91,000 neurons, our neurorobot brain has become much more functional compared to our previous model (Anumandla et al., 2011) in terms of consistently trusting or

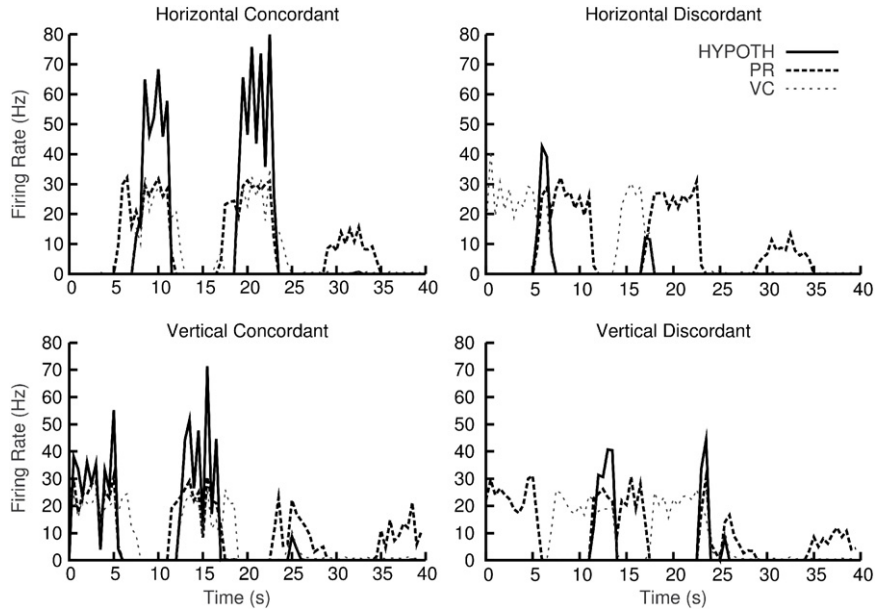

Fig. 4. Cortical and hypothalamic activities during concordant and discordant motions. Firing rates of three regions (VC: Visual Cortex; PR: Parietal Cortex; HYP: Hypothalamus) are represented during horizontal concordant (top left), horizontal discordant (top right), vertical concordant (bottom left), and vertical discordant (bottom right) motions.

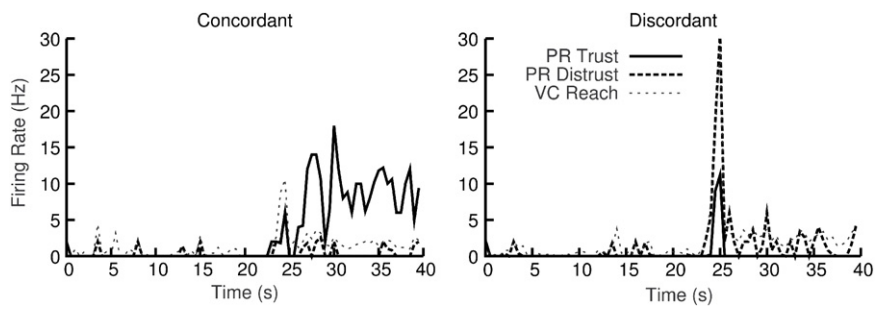

Fig. 5. Activity in cortical columns during reaching. Firing rates of three regions (PR trust: parietal cortex trust column; PR distrust: parietal cortex distrust column; VC reach: visual cortex during reaching) are represented during reaching and after concordant (left) or discordant (right) motions occurred during the learning phase.

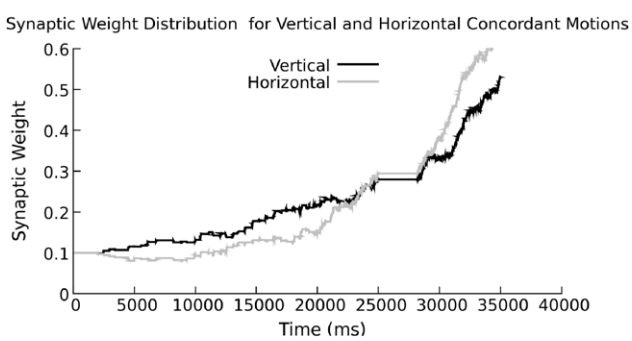

Fig. 6. Hypothalamic synaptic weights during concordant motions. The increase in the weight of hypothalamus synapses are represented during vertical (left) and horizontal (right) motions during the learning phase.

distrusting the human actor. Consequently, the neurorobot always learns discordant from concordant motions after a minimum of four trials. After improving the Gabor filtering mechanism as well as the brain architecture, the whole system has become more robust. This allowed higher flexibility and greater margins of error.

As we build and simulate larger and larger brains, we had to make design decisions regarding hardware and software in order to maintain the real-time performance we require. We have begun design and implementation of a new GPU based version of NCS. Preliminary results show that we are able to increase the number of neurons we can simulate by an order of magnitude or more. When this is coupled with multi-spatial simulations (that is the ability of having different simulator engines for different portions of the brain all working together), the capabilities our neurorobots will further extend in the future.

6. Conclusions

We modeled unique biological aspects of a mammalian brain related to the establishment of trust, including firing in the hypothalamus based on dendritic potentials, and the roles of oxytocin and the amygdala. The hypothalamic oxytocin cell model was developed as a realistic four compartment triple apical dendrite model where each dendrite receives stimulus from one of the three cortical columns: visual, parietal, and inferotemporal cortices. Our objective of increasing cooperativity between a virtual neurorobot and a human was achieved, in a model where the release of oxytocin into the hypothalamus (Bergquist & Ludwig, 2008; Komendantov, Trayanova, & Tasker, 2007; Neumann, 2007; Sabatier, Rowe, & Leng, 2007) is strengthened by reinforcement learning (Hurlemann et al., 2010) for concordant motions. This activity in turn inhibited the amygdala, which was responsible for cutting down cortical decision-making circuits. Thus the trust levels of the virtual neurorobot went up, whereas in the case of discordant behavior during learning there was no trust reaction due to the lack of inhibition in the amygdala. Our findings replicated the experimental results of Paukner et al. (2009), which demonstrated similar behavior in capuchin monkeys.

The trust established in our neurorobot is more consistent over several experiments than in our earlier report due to the improved Gabor mechanism and the higher number of neurons in every region of the brain. We conclude that our system is functional, robust, and biologically plausible.

The visual stimulus plays a dominant role in the VNR paradigm presented here, and the results demonstrated that role is sufficient in the trust building and decision making processes. This visual dependence has been illustrated in other studies, but is not a requirement in establishing trust (Tosoni, Galati, Romani, & Corbetta, 2008). As more advanced architectures are constructed based on these results, additional external stimuli such as facial recognition, auditory learning, or mirror neurons (Fitch, Huber, & Bugnyar, 2010) can complement or replace the visual cues.

The initial enhancements done on our hardware and software platforms allowed the experiments to run real-time simulations. The adequate processing power can handle a larger and more complex brain, especially the forward and reverse conductance between soma and dendrites of the hypothalamus.

We are now using this model as a basis for building complex neuromorphic brain architectures related to memory, trustworthiness, and social interactions. These were implemented on virtual robots in order to develop intelligent social agents using the VNR paradigm. Furthermore, insights into the functions of oxytocin may ultimately provide therapeutic benefits in neurological disorders such as autism (Andari et al., 2010; Hollander et al., 2003).

Acknowledgments

The authors would like to thank Gareth Ferneyhough, Nick Ceglia and Emily Barker for their help.

References

Andari, E., Duhamel, J. R., Zalla, T., Herbrecht, E., Leboyer, M., & Sirigu, A. (2010). Promoting social behavior with oxytocin in high-functioning autism spectrum disorders. *Proceedings of the National Academy of Sciences of the United States of America*, *107*(9), 4389–4394.

Anderson-Hunt, M., & Dennerstein, L. (1995). Oxytocin and female sexuality. *Gynecologic and Obstetric Investigation*, *40*(4), 217–221.

Anumandla, S.R., Jayet Bray, L.C., Thibeault, C.M., Hoang, R.V., Dascalu, S.M., & Harris, J.F.C. et al. (2011). Modeling oxytocin induced neurorobotic trust and intent recognition in human robot interaction. In *Paper presented at the proceedings of the international joint conference on neural networks, IJCNN*. San Jose, CA.

Baumgartner, T., Heinrichs, M., Vonlanthen, A., Fischbacher, U., & Fehr, E. (2008). Oxytocin shapes the neural circuitry of trust and trust adaptation in humans. *Neuron*, *58*(4), 639–650.

Bergquist, F., & Ludwig, M. (2008). Dendritic transmitter release: a comparison of two model systems. *Journal of Neuroendocrinology*, *20*(6), 677–686.

Bower, J. M., & Beeman, D. (1998). *The book of GENESIS: exploring realistic neural models with the general neural simulation system* (2nd ed.). New York, NY: Springer.

Brette, R., Rudolph, M., Carnevale, T., Hines, M., Beeman, D., Bower, J. M., et al. (2007). Simulation of networks of spiking neurons: a review of tools and strategies. *Journal of Computational Neuroscience*, *23*(3), 349–398.

Caporale, N., & Dan, Y. (2008). Spike timing-dependent plasticity: a Hebbian learning rule. *Annual Review of Neuroscience*, *31*, 25–46.

Carnevale, N. T., & Hines, M. L. (2006). *The NEURON book*. Cambridge, UK: Cambridge University Press.

Carter, C. S. (2003). Developmental consequences of oxytocin. *Physiology and Behavior*, *79*, 383–397.

Choleris, E., Little, S. R., Mong, J. A., Puram, S. V., Langer, R., & Pfaff, D. W. (2007). Microparticle-based delivery of oxytocin receptor antisense DNA in the medial amygdala blocks social recognition in female mice. *Proceedings of the National Academy of Sciences of the United States of America*, *104*(11), 4670–4675.

Dan, Y., & Poo, M. M. (2004). Spike timing-dependent plasticity of neural circuits. *Neuron*, *44*(1), 23–30.

Dautenhahn, K. (2007). Socially intelligent robots: dimensions of human–robot interaction. *Philosophical Transactions of the Royal Society of London. Series B. Biology Science*, *362*, 679–704.

Davison, A. P., Brüderle, D., Eppler, J., Kremkow, J., Muller, E., Pecevski, D., et al. (2008). PyNN: a common interface for neuronal network simulators. *Front Neuroinform*, *2*, 11.

De Schutter, E. (2008). Why are computational neuroscience and systems biology so separate? *Plos Computational Biology*, *4*(5).

Domes, G., Heinrichs, M., Glascher, J., Buchel, C., Braus, D. F., & Herpertz, S. C. (2007). Oxytocin attenuates amygdala responses to emotional faces regardless of valence. *Biological Psychiatry*, *62*, 1187–1190.

Domes, G., Heinrichs, M., Michel, A., Berger, C., & Herpertz, S. C. (2007). Oxytocin improves "mind-reading" in humans. *Biological Psychiatry*, *61*(6), 731–733.

Drewes, R., Zou, Q., & Goodman, P. H. (2009). A python toolkit to aid in the design, simulation, and analysis of spiking neural networks with the neocortical simulator. *Front Neuroinformatics*, *3*, 16.

Ferguson, J. N., Young, L. J., & Insel, T. R. (2002). The neuroendocrine basis of social recognition. *Frontiers in Neuroendocrinology*, *23*, 200–224.

Fitch, W. T., Huber, L., & Bugnyar, T. (2010). Social cognition and the evolution of language: constructing cognitive phylogenies. *Neuron*, *65*(6), 795–814.

Gewaltig, M., & Diesmann, M. (2007). NEST (Neural Simulation Tool). *Scholarpedia*, *2*(4), 1430.

Goodman, D. F., & Brette, R. (2009). The brian simulator. *Frontiers in Neuroscience*, *3*(2), 192–197.

Goodman, P. H., Buntha, S., Zou, Q., & Dascalu, S. M. (2007). Virtual neurorobotics (VNR) to accelerate development of plausible neuromorphic brain architectures. *Frontiers in Neurorobotics*, *1*, 1–7.

Goodman, P. H., Zou, Q., & Dascalu, S. M. (2008). Framework and implications of virtual neurorobotics. *Frontiers in Neuroscience*, *2*, 123–128.

Heinrichs, M., von Dawans, B., & Domes, G. (2009). Oxytocin, vasopressin, and human social behavior. *Frontiers in Neuroendocrinology*, *30*(4), 548–557.

Hollander, E., Novotny, S., Hanratty, M., Yaffe, R., DeCaria, C. M., Aronowitz, B. R., et al. (2003). Oxytocin infusion reduces repetitive behaviors in adults with autistic and Asperger's disorders. *Neuropsychopharmacology*, *28*(1), 193–198.

Huber, D., Veinante, P., & Stoop, R. (2005). Vasopressin and oxytocin excite distinct neuronal populations in the central amygdala. *Science*, *308*(5719), 245–248.

Hurlemann, R., Patin, A., Onur, O. A., Cohen, M. X., Baumgartner, T., Metzler, S., et al. (2010). Oxytocin enhances amygdala-dependent, socially reinforced learning and emotional empathy in humans. *Journal of Neuroscience*, *30*(14), 4999–5007.

Jones, J. P., & Palmer, L. A. (1987). An evaluation of the two-dimensional gabor filter model of simple receptive-fields in cat striate cortex. *Journal of Neurophysiology*, *58*(6), 1233–1258.

Kelley, R. (2009). *Mind reading for social robots: stochastic models of intent recognition*. University of Nevada, Reno, NV.

Kirsch, P., Esslinger, C., Chen, Q., Mier, D., Lis, S., Siddhanti, S., et al. (2005). Oxytocin modulates neural circuitry for social cognition and fear in humans. *Journal of Neuroscience*, *25*(49), 11489–11493.

Koenig, N., Takayama, L., & Mataric, M. (2010). Communication and knowledge sharing in human–robot interaction and learning from demonstration. *Neural Networks*, *23*(8–9), 1104–1112.

Komendantov, A. O., Trayanova, N. A., & Tasker, J. G. (2007). Somato-dendritic mechanisms underlying the electrophysiological properties of hypothalamic magnocellular neuroendocrine cells: a multicompartmental model study. *Journal of Computational Neuroscience*, *23*, 143–168.

Kosfeld, M., Heinrichs, M., Zak, P. J., Fischbacher, U., & Fehr, E. (2005). Oxytocin increases trust in humans. *Nature*, *435*(7042), 673–676.

Lim, M. M., & Young, L. J. (2006). Neuropeptidergic regulation of affiliative behavior and social bonding in animals. *Hormones and Behavior*, *50*, 506–517.

Markram, H., Gerstner, W., & Sjöström, P. J. (2011). A history of spike-timing-dependent plasticity. *Front. Syn. Neurosci.*, *3*, 4.

Mehrotra, R., Namuduri, K. R., & Ranganathan, N. (1992). Gabor filter-based edge-detection. *Pattern Recognition*, *25*(12), 1479–1494.

Neumann, I. D. (2007). Stimuli and consequences of dendritic release of oxytocin within the brain. *Biochemical Society Transactions*, *35*, 1252–1257.

Ohman, A. (2005). The rote of the amygdala in human fear: automatic detection of threat. *Psychoneuroendocrinology*, *30*(10), 953–958.

Paukner, A., Suomi, S. J., Visalberghi, E., & Ferrari, P. F. (2009). Capuchin monkeys display affiliation toward humans who imitate them. *Science*, *325*(5942), 880–883.

Peng, Q. (2006). *Brainstem: a neocortical simulator interface for robotic studies*. University of Nevada, Reno, NV.

Petrovic, P., Kalisch, R., Singer, T., & Dolan, R. J. (2008). Oxytocin attenuates affective evaluations of conditioned faces and amygdala activity. *Journal of Neuroscience*, *28*(26), 6607–6615.

Richert, M., Nageswaran, J. M., Dutt, N., & Krichmar, J. L. (2011). An efficient simulation environment for modeling large-scale cortical processing. *Front Neuroinform*, *5*, 19.

Rimmele, U., Hediger, K., Heinrichs, M., & Klaver, P. (2009). Oxytocin makes a face in memory familiar. *Journal of Neuroscience*, *29*(1), 38–42.

Saavedra, S., Smith, D., & Reed-Tsochas, F. (2010). Cooperation under indirect reciprocity and imitative trust. *PLoS One*, *5*(10), Article No.: e13475.

Sabatier, N., Rowe, I., & Leng, G. (2007). Central release of oxytocin and the ventromedial hypothalamus. *Biochemical Society Transactions*, *35*, 1247–1251.

Scheutz, M., Schermerhor, P., Kramer, J., & Anderson, D. (2007). First steps toward natural human-like HRI. *Autonomous Robotics*, *22*, 411–423.

Song, S., Miller, K. D., & Abbott, L. F. (2000). Competitive Hebbian learning through spike-timing-dependent synaptic plasticity. *Nat Neurosci*, *36*, 919–926.

Tahboub, K.A. (2005). Compliant human–robot cooperation based on intention recognition. In *Paper presented at the intelligent control. Proceedings of the 2005 IEEE international symposium on, mediterrean conference on control and automation*, Limassol.

Thomaz, A. L., & Breazeal, C. (2008). Teachable robots: understanding human teaching behavior to build more effective robot learners. *Journal of Artificial Intelligence*, *172*(6–7), 716–737.

Tosoni, A., Galati, G., Romani, G. L., & Corbetta, M. (2008). Sensory-motor mechanisms in human parietal cortex underlie arbitrary visual decisions. *Nature Neuroscience*, *11*(12), 1446–1453.

Trafton, J.G., Schultz, A.C., Perzanowski, D., Bugajska, M.D., Adams, W., & Cassimatis, N.L. et al. (2006). Children and robots learning to play hide and seek. In *Paper presented at the human–robot interaction*.

Whalen, P. J., Kagan, J., Cook, R. G., Davis, F. C., Kim, H., Polis, S., et al. (2004). Human amygdala responsivity to masked fearful eye whites. *Science*, *306*, 2061.

Wilson, C. E., Goodman, P. H., & Harris, F. C. Jr. (2001). Implementation of a biologically realistic parallel neocortical-neural network simulator. In *Paper presented at the proceedings of the 10th SIAM conference on Parallel Process. for Sci. Comp.* Portsmouth, VA.

Young, L. J., & Wang, Z. (2004). The neurobiology of pair bonding. *Nature Neuroscience*, *7*, 1048–1054.

Zak, P. J., Kurzban, R., & Matzner, W. T. (2005). Oxytocin is associated with human trustworthiness. *Hormones and Behavior*, *48*(5), 522–527.

Zhang, L. I., Tao, H. W., Holt, C. E., Harris, W. A., & Poo, M.-M. (1998). A critical window for cooperation and competition among developing retinotectal synapses. *Nature*, *395*, 37–44.