

Received May 5, 2020, accepted May 25, 2020, date of current version June 5, 2020.

Digital Object Identifier 10.1109/ACCESS.2020.2998396

Computational Modeling of Prefrontal Cortex for Meta-Cognition of a Humanoid Robot

EVREN DAGLARLI¹, (Member, IEEE)

Faculty of Computer and Informatics Engineering, Istanbul Technical University, 34469 Istanbul, Turkey

e-mail: evren.daglarli@itu.edu.tr

This work was supported by the NVIDIA Corporation through the GPU Grant Program.

ABSTRACT For robot intelligence and human-robot interaction (HRI), complex decision-making, interpretation, and adaptive planning processes are great challenges. These require recursive task processing and meta-cognitive reasoning mechanism. Naturally, the human brain realizes these cognitive skills by prefrontal cortex which is a part of the neocortex. Previous studies about neurocognitive robotics would not meet these requirements. Thus, it is aimed at developing a brain-inspired robot control architecture that performs spatial-temporal and emotional reasoning. In this study, we present a novel solution that covers a computational model of the prefrontal cortex for humanoid robots. Computational mechanisms are mainly placed on the bio-physical plausible neural structures embodied in different dynamics. The main components of the system are composed of several computational modules including dorsolateral, ventrolateral, anterior, and medial prefrontal regions. Also, it is responsible for organizing the working memory. A reinforcement meta-learning based explainable artificial intelligence (xAI) procedure is applied to the working memory regions of the computational prefrontal cortex model. Experimental evaluation and verification tests are processed by the developed software framework embodied in the humanoid robot platform. The humanoid robots' perceptual states and cognitive processes including emotion, attention, and intention-based reasoning skills can be observed and controlled via the developed software. Several interaction scenarios are implemented to monitor and evaluate the model's performance.

INDEX TERMS Artificial intelligence, brain modeling, cognitive robotics, human-robot interaction.

I. INTRODUCTION

Today, the humanoid robots are working together with humans as a personal assistant. However, they will take on more duties in the future. Due to its bipedal structure and physical capabilities, humanoid robots are very compatible to interact with human nature. Recent advancements in paradigms of artificial intelligence and cognitive neuroscience can contribute to the evolution of their technologies [1], [2]. In the next decades, it is expected that humanoid robots having cognitive skills will be widely used in social areas such as assistive, entertainment, and rehabilitation fields [3]–[6]. Robots using social areas require some human-like cognitive skills such as reasoning, decision making, problem-solving [4]. These embodied skills may organize very complex behavior patterns rather than perform deterministic or repetitive tasks [5]. Humanoid robots with enhanced embodied cognitive abilities can be used

The associate editor coordinating the review of this manuscript and approving it for publication was Wei Zhang.

to assist disabled individuals struggling to interact with their social environment by guiding their accessibility and communication [4], [5].

As future deliverables of this study, it is considered that humanoid robots competing with the intelligence of the human race can socially interact and collaborate with every social area as behaves like being a part of humanity so that they increase the living standards of its society. In order to realize that, the main purpose of this study intends to construct a suitable computational framework of a cognitive model for a humanoid robot. Because cognitive skills that are biologically existed in the human race can help to realize these goals [7]–[9]. Therefore, it is essential to investigate the biological nature of cognitive systems by the viewpoint of a humanoid robot.

In nature, these cognitive functions and skills could be biologically realized via some cortical and cerebral lobes in the human brain [10], [11]. The anatomical structure of the cerebral cortex includes two main cortical structures called frontal and posterior parts [12], [13].

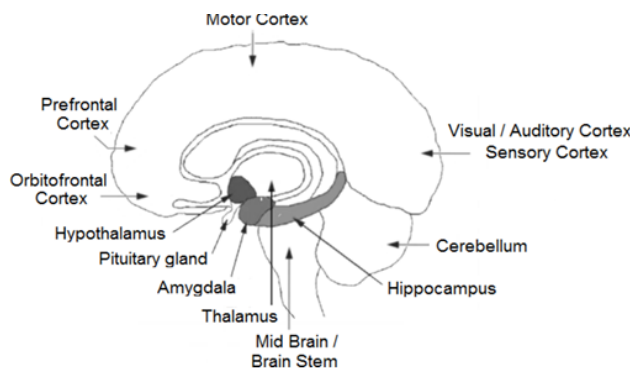


FIGURE 1. Human brain cortical zones [7].

Cognitive functions related to reasoning, planning, and working memory skills involve in the prefrontal cortex which is a part of the frontal lobe of the cerebral cortex (Figure 1). This part of that region is divided into four sub-cortical areas including dorsolateral, ventrolateral, anterior, and medial prefrontal regions [14]–[16]. Dorsolateral and ventrolateral prefrontal cortices realize spatial and temporal reasoning tasks respectively. The anterior prefrontal cortex (or Brodmann Area 10) perform emotional reasoning tasks. The medial prefrontal cortex including anterior cingulate cortex (ACC), dorsomedial PFC (dmPFC) and orbitofrontal cortex (OFC) regions, involves working memory and meta-cognition. Monitoring and association of these reasoning skills are realized by this computational unit.

As a big challenge, this study focuses on solving interaction games, logical puzzles to investigate planning problems, interaction failures during human-robot interaction (HRI) [17]–[27]. The content of interaction game or logical puzzle can include memory-based sorting or prediction tasks. Interaction games, logical puzzles can useful testbed for integrating emotion, intention, and attention-based reasoning skills into these problems [26], [27]. Therefore, the problem statement of the paper covers and aims whether a humanoid robot which is a personal assistant or companion can achieve these goals by establishing rational social interaction with humans. Since human-like social interaction of the robot with humans is a very hard issue and requires high-level mental skills including joint attention, the performance of spatial-temporal and emotional reasoning, meta-cognitive planning gain very critical importance in the computational mental model of a humanoid robot, during the interaction experiments between robot and human [17]–[20]. While the humanoid robot is struggling to interact with a human, they can establish a link including communication channels of emotion (feeling), skill, and physical resource sharing [21]–[23]. Thus, humanoid robot as a multi-modal mediator can be utilized in assistive case studies such as rehabilitation of the children suffering from attention deficit or autism spectrum disorders [24], [25]. Because the humanoid robots like these children may not achieve accurate meta-cognitive planning, focusing, sensory-motor association and may not execute right decisions, rational emotional

responses, precise motor commands which help to exhibit spatial behavior patterns (e.g., movement skills like imitation gestures, tracking the objects) [22]–[24].

Recently, brain-inspired computational cognitive architectures were developed for human-robot interaction (HRI) with humanoid robots to solve behavior planning, meta-cognitive reasoning problems [17]–[21]. Such projects are rapidly increased, and they are expected to rise as an emergent issue in the future. Some very good and remarkable examples related to a computational architecture based on cerebral/neocortex models of the brain were presented in recent years [1]–[6]. In these studies, various machine learning methods were employed for modeling of embodied cognition and adaptive behavior. Khamassi *et al.* developed a computational neurocognitive model of the prefrontal cortex for navigation of rat-like robots [28]–[34]. In their model, reinforcement learning was utilized to realize the cognitive functions of these cortical structures. Yildirim *et al.* [35] introduced an integrative cognitive architecture for object driven cortex. They dealt with a combination of multiple AI methods including causal generative models, hybrid symbolic-continuous planning algorithms with physics-based dynamics. In order to realize these, the object-driven cortex system benefits from the biological evidence of the specific regions in the human brain. Gómez-Martínez *et al.* [36] put forward a bio-inspired self-responding emotional behavior system for virtual creatures. This model enabling virtual creatures works as a concurrent and parallel distributed system to accord to their environment and generate more reliable behavior. A case study with the interaction scenario to be executed was implemented by the 3D simulation platform so that the results of the study were observed when the creature interacts with the environment. Hernández García *et al.* [37] utilized a biologically plausible spiking neural network to realize visual attention and object labeling for the iCub humanoid robot. According to this, the cortical neural network model imitating the human brain functions provides an associative memory that can learn words of the objects by visual attention. Experiments were conducted with both a simulated and a real iCub robot platform which is capable of associating a label to an object when visual and word stimuli are provided simultaneously in the experiment scene. The planning and decision-making based computational model considering effective information was developed by Cervantes *et al.* [38]. This computational model inspired by some brain areas organized as many sub-modules providing related cognitive functions such as emotional system, rational decision-making, episodic memory, and motivational system. The claims of their study were tested by a case study that experimental results included a remarkable comparison. Mizutani *et al.* [39] endeavored to realize whole-brain connectome architecture for achieving general purpose artificial intelligence. Their study aims to develop biologically meaningful AI architecture which includes neural circuit representations of the entire brain. Verschure recently reported a set of design principles underlying

the mind, brain, and body nexus (MBBN) [40], proposing a brain-inspired cognitive architecture composed of several computational columns and layers. An embodied biologically plausible model which has adaptive behavior skill studied by Maffei *et al.* [41]. In both models, the computational model involving the brain's ability to maintain strong interaction between an embodied agent and its environment through action is ensured via distributed adaptive control (DAC) architecture. The DAC-X architecture [41] is organized into two complementary structures of layers and columns, providing the computational neuro-cognitive architecture for use in robots. A set of layers, including reactive, adaptive, and contextual layers, describes the developmental stages of the architecture. The columnar structures, including the processing of states of the world, the self, and the generation of action, describe functional sub-modules of the architecture. A range of hypotheses related to the work has been verified utilizing humanoid and mobile robots in a range of application scenarios. Zeng *et al.* [42] struggled on imitating self-consciousness using brain-inspired modeling for a humanoid robot. In their study, it is aimed to build a computational approximate model of the primate mirror neuron system to realize cognitive functions such as self-recognition. In the first stage, the correlations between self-generated actions and visual feedbacks in motion are learned by the robot using a spike timing dependent plasticity (STDP) algorithm. In the second stage, the expected appearance of body gesture that the visual feedback is consistent with its motion is learned. During these tasks, the robot utilizes multi-sensor fusion to learn its own body in the real world and in the mirror. Vernon *et al.* [43] pursued a study for joint episodic-procedural memory in cognitive robotics. According to their idea, joint episodic-procedural memory enables the internal simulation to be conditioned by current context, semantic memory, and the agent's value system providing the motives, while context and semantics restrain the integrated explosion of potential perception-action associations and allow effective action selection to achieve the goals. Samsonovich proposed an emotional biologically inspired cognitive architecture (eBICA) based on mental schemas, producing a model involving schemas and mental states as the major components and moral schemas embodied into appraisals as attributes [44]. In this model, patterns of social emotions and semantic spaces are represented and controlled by these appraisals, and emotional modulation processes behavior and clusters emotional states in the arousal-valence domain. In an experiment involving human subjects and virtual agents, the proposed architecture was tested based on a simple paradigm involving a virtual world. The results revealed that eBICA with a moral schema was able to manipulate human behavior. Such a framework may be useful for enabling collaboration between virtual partners and humans, self-regulated learning of virtual agents, and the development of realistic emotional intelligence. In addition, Sengor *et al.* reported a robot model based on cortico-striato-thalamic circuits, developing a computational

model of the basal ganglia for use with a mobile robot [45]. The main purpose of this field of research is to investigate the potential use of robot models for implementing intelligent systems to inspire new approaches and techniques. A number of studies have examined reward-based computational mechanisms, shedding light on biological explanations of animal action selection processes. One study used a Khepera II mobile robot to study the implementation of goal-directed behavior. Moreover, computational neural models of cognitive processes have been utilized to describe higher-order processes like goal-directed behavior. A neuro-computational model of the auditory-cue fear acquisition was investigated by Navarro-Guerrero *et al.* [46], elucidating several principles of fear conditioning necessary for developing adaptive, self-protective systems, and suggesting that sensory-motor processing is an essential component of fear learning. This hybrid approach has been reported to be an effective tool for examining the temporal relationship between auditory sensory cues. A detailed study of computational mechanisms based on neural circuits in the brain has supported the development of safer robots and a better understanding of fear processing.

In this study, we present a novel complete solution that covers a computational model of the prefrontal cortex which is composed of several sub-modules including dorsolateral, ventrolateral, anterior, and medial prefrontal regions for humanoid robots. The main contribution of this framework is to develop a computational approximation of the consciousness, awareness so that the humanoid robot can perform high level (meta-cognition) reasoning which associates spatial-temporal and emotional reasoning skills. In this architecture embodied into a humanoid robot, one of the major impacts is that the robot as a synthetic live form establishes human-like interaction with society and its environment. Using reinforcement learning algorithm, spatial-temporal and emotional reasoning skills realize several processes including attention, short-term memory, decision making (e.g. arithmetic and logical), planning, analysis of cause-effect relations, and problem-solving. These activities are monitored by the medial part of the computational prefrontal cortex model which is working as a reward mechanism. All cognitive transactions in the proposed prefrontal cortex model are organized into working memory. The working memory which is a type of associative memory is stored in the weights of an attractor network. According to forward and backward propagation in the network, imagination, recalling and prediction abilities can appear. For this framework, weights (synaptic strengths) of the network related to working memory are updated by various learning paradigms corresponding with the prefrontal cortex model. The unsupervised learning method is utilized for the clustering of the observation data constructed in the cognitive map (working memory). For attention, intention, emotion and meta-cognitive reasoning, the probabilistic model utilizing reinforcement learning can be considered for learning skills including inference, decision making and planning in the robot. In the proposed model, post

linguistic processing is realized by applying grammar rules, and acquired linguistic statements are located in the working memory.

The article follows with chapter 2 that explains the problem statement and description of related hypotheses. In the following section, chapter 3 gives background knowledge about materials and methods. In chapter 4, it is determined the design principles of the computational model of the prefrontal cortex realizing working memory, decision making, metacognitive planning, spatial-temporal reasoning, and inference on the humanoid robot. In chapter 5, the detailed information about the experimental platform and implementation of test scenarios are introduced. Then chapter 6 depicts experimental results and performance evaluation statistics. Finally, discussion, concluding remarks, and future works are presented in chapter 7.

II. PROBLEM STATEMENT AND EVALUATION CRITERIA OF THE GOAL

For decades, as a concept of cognitive control, functions of prefrontal cortical regions where behavioral procedures require to be modified or reorganized, responsible for the regulation of complex decision making, meta-cognitive planning, and reasoning, have been widely worked by neuro-cognitive and robotics researchers [47], [48]. These efforts contributed to the social interaction experiments with humanoid robots.

In this section, the major concern to be addressed is how to evaluate or measure the objectives of the problems in a rational social interaction between the humanoid robot and the human. In order to do this, an interaction game like a logical puzzle including memory-based sorting or prediction tasks is utilized to test hypotheses achieving the objectives of this study. Therefore, the proposed study is hypothesized not only to involve relations between subdivisions of the lateral PFC including dorsolateral, ventrolateral, and anterior regions but also to involve relations between the medial and lateral PFC for the field of computational neuro-cognitive robotics and its social interaction applications.

In shed light of these facts, I proposed a hypothesis (H1) to ensure that memory-based sorting or prediction tasks like a logical puzzle game are robustly solved by using spatiotemporal and emotional reasoning skills trained by reinforcement learning in a computational model of lateral PFC during rational social interaction. Spatial reasoning skills performed by dorsolateral PFC correspond to behavioral (movement based) intentions. Attention-based mental activities are associated with temporal reasoning skills realized by ventrolateral PFC. Emotional reasoning skills are performed by anterior PFC region. According to this, if all these skills including attention, intention, and emotion-based reasoning activities are simultaneously employing to contribute solution of the problem, it is expected that durations of achieving simple sorting or prediction tasks in the puzzle game are decreased while the accuracy of the tasks is increasing during the experiments.

Also, I put forward a hypothesis (H2) to ensure that spatial-temporal and emotional reasoning skills in lateral PFC are regulated by a computational model of medial PFC during human-like social interaction with a humanoid robot. A computational model of medial PFC composed of several sub-modules including anterior cingulate cortex (ACC), dorsomedial PFC, and orbitofrontal cortex (OFC) involves association and supervision procedures in order to monitoring and reorganizing of these mental transactions. In the medial PFC, the computational aspect of meta-cognition which is interpreted as a multi-modal reward mechanism provides support reinforcement learning activities of lateral PFC involving in spatiotemporal and emotional mental skills. The reward computation activities of these mental skills, based on the action-reward association are proceeded by ACC dynamics. And stimulus-reward association grounded reward computation activities are executed by OFC dynamics. For decades, by many researchers, concepts of consciousness and self-awareness have been associated with meta-cognition which is considered to take a major role in human-like cognition. For meta-cognition, if the meta-cognition mechanism in a computational model of medial PFC is integrating appropriate rewards to all mental functions of lateral PFC for regulating complex decision making, meta-cognitive planning and reasoning activities, convergence errors in reinforcement learning of lateral PFC related skills including attention, intention and emotion-based reasoning activities are decreased while success rates of the simple sorting or prediction tasks in the puzzle game are increasing during the experiments.

Also, another key issue is that dopaminergic gain released by ventral tegmental area (VTA) for tuning the parameters (e.g. learning rate, reward discount factor and exploration rate, etc.) affects not only reinforcement learning process phase but also reward computation phase in medial PFC. In order to realize synaptic neural plasticity which can provide a great advantage for adaptation, an optimization algorithm is employed by an unsupervised learning method. Furthermore, according to different conditions, synaptic neural plasticity can be parameterized via dopamine modulation for optimizing convergence speed. In shed light of this, as a neural plasticity mechanism, if an optimization algorithm contributes to obtaining a better dopaminergic gain parameter, learning processes will be faster and decision-making activities will be more robust. Also, this optimization mechanism with parametric modulation is a very important issue in which short-term memory in PFC interacts with episodic memory (e.g. historical data related to past events during the interaction) released by hippocampal regions.

III. BACKGROUND ABOUT MATERIALS AND METHODS

In order to achieve goals and examine proposed hypotheses, anatomical and functional features of PFC in a human brain should be analyzed by principles of mathematics, cognitive science, and neuroscience. According to cognitive-neuroscientific evidence and the viewpoint of computer science, the prefrontal cortex which is the key

point for developmental meta-cognition might be interpreted by combinations of different computational methods and machine learning paradigms. As an artificial intelligence methodology, this approach can be employed by neurocognitive robotics, human-robot interaction, and social robotics applications.

The neural models are very useful methods for representing and encoding cognitive activities. Naturally, computational modeling of cognitive and mental processes in the human brain, for application in humanoid robots, requires high-density neural structures [7], [8]. Among these models, spiking neural networks (SNN) [49]–[51] including spike response models, integrate-fire models, Izhikevich model, biologically plausible models (Hodgin-Huxley model, conductance or ion flow-based models) are the most similar to biological neural systems. Spiking neural cells can be represented as a nonlinear circuit with a capacitor like integrator. In addition to this, stochastic and chaotic behaviors like dynamic attractions are taken into consideration. The conductance-based neural model and its population activity can be generally expressed by

$$C \frac{dV_j}{dt} = \sum_i g_i \varphi_i(E_i - V_j) + I_{ext} \quad (1)$$

$$A_k(t) = \frac{1}{\Delta t} \frac{\sum_k \sum_j \delta(t; t + \Delta t)}{N_k} \quad (2)$$

where I_{ext} represents the external input current, including synaptic and interconnectional. Also, V_j depicts the j^{th} neuron's membrane potential, with a capacitance coefficient C . Potential E_i indicates reversal potentials related to their ion channels. Their conductance parameters g_i and reversal potentials E_i help to generate ion currents through gate variables φ_i , which are computed by ordinary differential equations. Ensemble or neural population activity is computed as $A_k(t)$, where δ , N_k are a number of spikes in the time interval (Δt), neuron numbers in k^{th} neural population respectively. The cloud of spiking neurons is driven by firing rate models of a neural population. Besides the main neuron pool, there can be some kind of sub-populations called excitatory and inhibitory populations. The general form of neural population density can be interpreted by Jansen-Rit's neural mass model equations [52]. There is a very close relationship between the signal relay from a population density of a neural mass and electroencephalography (EEG) activity [53]. The term $A_k(t)$ which is the mean ensemble firing rate of presynaptic excitatory input is passed into the computational model of the neural mass. Then population dynamics driven by neural mass equation provide information flow in the computational neural tissue.

A neural population activity stimulating the neural mass is generated by the contribution of mean firing rate coding which involves synaptic firing density, and spiking counts in a neural population of spiking neurons. Assembling small groups of neural populations produces larger complex structures (e.g., cortical regions or neural tissues) [54].

Their connection topologies allow brain activities with chaotic dynamic characteristics to be computationally reproduced for artificial brain frameworks.

As a large-scale neural activity, field dynamics that visualize cognitive behaviors might be associated with functional magnetic resonance imaging (fMRI) [55], [56]. Dynamic neural fields (DNF), a dynamic field (DF) branch, are driven by the Amari equations [57], [58]. The field activities which exhibit certain stochastic and nonlinear dynamic properties behave like wave packets traveling along the neural field.

$$\frac{d}{dt} U_i(x, t) = -U_i(x, t) + \int w_i(\ell) f(U_i(x - \ell, t - \tau)) d\ell + \int S_j(x, y) I_d(y, t) dy + h \quad (3)$$

where $\ell = x - x'$ is a spatial information distance from the mean of the cortical field. A weight matrix $w_i(\ell)$, which is a function of the spatial information distances, includes connection strengths of the synaptic activation inside the field $U_i(x, t)$. The function f is a sigmoid activation function. The parameter h corresponds to the bias. External effects can be realized in the exogenous connection matrix $S_j(x, y)$ and exogenous input field activity $I_d(y, t)$. As parallel and distributed computing blocks, the field dynamics of the computational neural tissue improve the model employing biophysical, meaningful spiking neuron populations and neural masses.

According to field activity $U_i(x, t)$, the dynamic associative memory related to cognitive skills can be shaped by multi-modal or strange attractors including chaotic behaviors, bifurcations, limit-cycles, and chattering. The attractive or repulsive forces are occurred by the gradient of the field potentials in n -dimensional $U_i(x, t)$. As a fragment of the short-term memory, the associative memory pattern related to a cognitive behavior is formed by the trajectory of the specific memory state expressed with $q_i(t)$. The attractor network behavior which is a nonlinear dynamic system can be expressed by following differential equation;

$$\frac{d}{dt} q_i(t) = -q_i(t) + \alpha(t) \nabla U_i(x, t) \quad (4)$$

where $\alpha(t)$ defines the magnitude of the attractive or repulsive forces $\nabla U_i(x, t)$ and the speed of memory state transition along with the trajectory. Also, stability features of the attractor dynamics can be evaluated by Lyapunov analysis which affects convergent/divergent characteristics, saddle-points, cyclic behaviors (oscillations), or point attractors (e.g. stable focus, a basin of attractor) [59].

In addition, nonlinear neural dynamics contributing realization of the cognitive skills provide a good opportunity for solving problems of the dynamic programming system [60], [61]. Thus activities of cognitive skills can be interpreted as the dynamic programming system like a typical finite state machine with discrete state transitions [60], [62]. The probabilistic representation of the discrete state transitions can be evaluated by Markov process which is a stochastic system [60]–[62]. According to whether sequential states

are observable, the Markov process which is similar to Bayes networks is divided into several types such as hidden Markov models (HMM), Markov decision process (MDP), partially observable MDP (POMDP) [63], [64]. The Bayesian networks driven by Bayes rule can be represented as graph topology of the state network including probabilistic causal relations a_{ij} between the nodes (states).

$$a_{ij} = P(x_t = s_j | x_{t-1} = s_i) = P(x_t | x_{t-1}, \dots, x_1) \quad (5)$$

$$P(x_1, x_2, \dots, x_{T-1}, x_T) = P(x_1) \prod_{t=1}^T P(x_t | x_{t-1}) P(y_t | x_t) \quad (6)$$

The causal relations or transition probabilities are stored in a matrix that defines the network topology related to Bayesian network-based models or Markovian models [65], [66]. Also, these models which are essential for reasoning and inference process can utilize different training algorithms including maximum likelihood estimation, mixture models (e.g. Gaussian mixture model (GMM)), Viterbi algorithm, forward-backward, Baum-Welch algorithms and Bayes learning rule [64]–[67]. Usually, the MDP based models that can be solved with value iteration algorithms are related to reinforcement learning methods employing Bellman learning rule [61].

The representation of every sequential node (states) in a network of Bayesian or Markovian domains can attribute extra computations including some decision-making processes [60], [66]. These processes related to reasoning and inference operations might involve not only deterministic or rule-based approaches but also probabilistic or stochastic approaches. The rule-based decision making methods (e.g. decision trees, random forests, fuzzy logic tools like fuzzy cognitive maps) are used to make specific rule-based decision operations that can provide “if-then” like cause-effect reasoning activities [68]–[72]. Decision tree approaches that constitute stacked rules employ hierarchically structured form including several nodes and branches [68], [70]. As a bundle composed of multiple decision trees, random forest which is an ensemble learning method is aggregated via a stochastic approach for classification, regression, and other complex decision-making tasks with multi-modal goals [69], [70]. Several supervised or unsupervised learning methods can be applied to these models. In addition, probabilistic or fuzzy uncertainties can be attributed to decision-making tasks. Different rules related to these tasks can be created via several fuzzy variables as uncertainty which is a result of membership function [71]. The fuzzy cognitive map (FCM) introduced by Bart Kosko is defined as a network topology of causal relationships between the nodes (e.g., concepts, events) [72], [73]. For rule-based FCM, these nodes can be replaced by the rules with fuzzy variables [74], [75]. As a typical variation of FCM, the causal relationships can have probabilistic functions. Hebbian learning or genetic algorithm based training methods have been mostly utilized in FCM [76].

Because of the fact that very huge amount of data processing is required for the realization of the cognitive skills

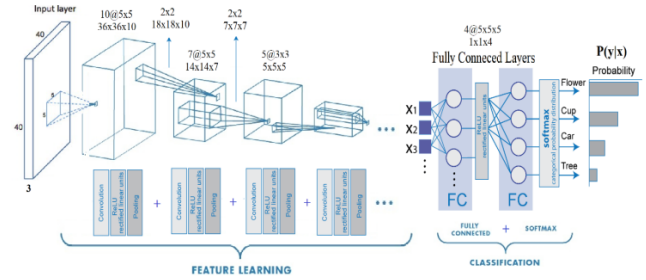


FIGURE 2. Convolutional neural network (CNN) architecture.

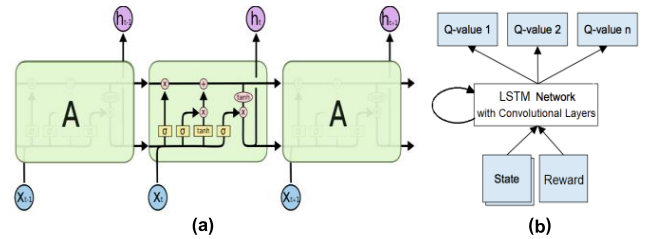


FIGURE 3. a) LSTM, b) LSTM based DQN.

and executing mental activities in PFC, different deep learning paradigms can be considered such as a convolutional neural network (CNN), recurrent neural network structures like long-short term memory (LSTM) and reinforcement deep learning model as deep Q-network (DQN). Especially, according to conditions, they might be combined as a hybrid machine learning approach.

For the classical structure of a convolutional neural network which has strong feature learning capacity resembling high-level abstraction processes in the human brain, the sequence of the cascaded layers is stacked as [input (x) – convolution layer - ReLU – max pool layer] (Figure 2). The convolution operator requires convolution filters (weights) as a rectangular prismatic tensor. Rectified linear unit (ReLU) is an activation function. This neural network is trained by stochastic gradient descent (SGD) with momentum so that features to be encoded are hierarchically extracted from the pattern of large scale neural transactions [77]–[79]. As a recurrent neural network (RNN), the LSTM (Figure 3.a) resembling memory activations in the human brain can be constructed by convolutional layers. It is capable of storing short-term memory contexts for a long-term period in the episodic memory pattern [77], [78], [80]. A conventional component of the LSTM is composed of a memory cell including input, output, and forget gates. For the training session, the backpropagation through time algorithm can be preferred [80], [81].

Deep Q network (DQN) combining the Q-learning algorithm with the deep neural networks can be re-organized by a LSTM network embodied with convolutional layers [82]–[84] (Figure 3.b). As a reinforcement learning method, DQN based models which allow sequential decision-making tasks can be associated with meta-cognitive activities in medial PFC in the human brain. In addition, instead

of a conventional reinforcement learning approach, these methods with deep neural networks provide a more flexible function approximation feature to reach human-level performance on rational social interaction involving complex decision-making skills. For optimization in the training phase, the loss function of the network is computed as follows,

$$L(\theta_i) = E_{s,a} \left[(r + \gamma \max_{a'} Q(s', a' | \theta_{i-1}) - Q(s, a | \theta_i))^2 \right] \quad (7)$$

where q is a learning parameter set. The experience memory replay (D) is a set of tuples composed of (s, a, r, s') . r and $Q(s, a)$ are a reinforcement reward and Q value including state (s) and action (a) pairs respectively. SGD algorithm updates the experience memory replay, while it is struggling to minimize the loss function. For the computational PFC model of the humanoid robot, additional model parameters (e.g. dopaminergic gain, motivation factor, learning rate, etc.) can be optimized by an adaptation algorithm. The optimization mechanism which employs an adaptation algorithm is mostly chosen by unsupervised learning methods (e.g., self-organizing map (SOM)) or metaheuristic methods (e.g., genetic algorithm (GA)). According to the chosen fitness criteria, the parameters to be adapted can be searched by a genetic algorithm. During this process, some conventional operations including mutation and crossover are utilized to obtain optimal values of the parameters [85], [86]. The self-organized map is preferred for tasks such as clustering, dimension reduction on the parameters to be adapted [87], [88].

IV. COMPUTATIONAL PREFRONTAL CORTEX MODEL

The computational model of the prefrontal cortex (Figure 4) which is crucial for many mental abilities such as meta-cognition, consciousness, self-awareness is built on large scale ensemble bio-inspired spiking neural structure-based dynamic domain such as neural field, attractor model. By this approach, computational modeling of the mental activities including cognitive skills, behaviors, emotions and rational thinking (planning or sequential decision-making tasks) which are represented as a sequence of statements, stacked rules, can be dynamically encoded by the temporal synaptic activations (e.g. spiking, ensemble and field dynamics) in the high-density complex network topologies (weights). The computational neural structure of PFC which is suitable to neuro-morphological equivalent (e.g., human brain) is composed of several subcortical regions (or modules) such as dorsolateral (dlPFC), ventrolateral (vlPFC), anterior (aPFC) and medial (mPFC) prefrontal cortex areas. Two major data streams coming from parietal and temporal regions (lobe) of the sensory cortex are propagated into dlPFC and vlPFC modules of the computational model of lateral PFC by dorsal (where) and ventral (what) pathways respectively. In addition, the prefrontal cortex is supported by connections of limbic system components including amygdala, thalamus,

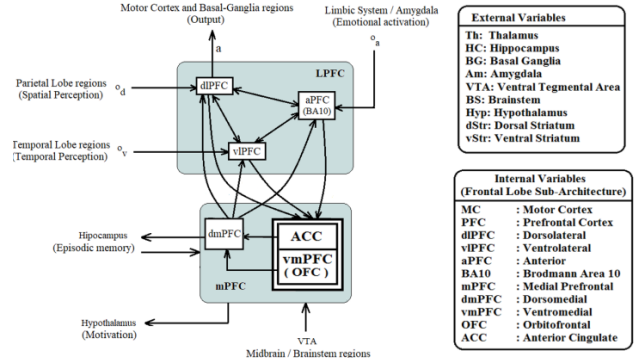


FIGURE 4. Computational architecture of prefrontal cortex (PFC).

hippocampus, basal ganglia, hypothalamus so that it regulates cognitive skills like decision making and planning abilities.

Extensive parallel and distributed neurocognitive transactions employing the stochastic recursive mathematical models and dynamic nonlinear programming methodologies occur in the computational PFC model including hybrid machine learning algorithms like a recurrent and reinforcement deep learning algorithm with partially observable variables. These partially observable variables or internal states are represented by a sequence of statements, stacked rules in the proposed huge neural network structure. In order to achieve adaptation, some tunable hyper-parameter set as a dopaminergic gain vector, learning, and exploration rates can be optimized by metaheuristic learning methods (e.g., genetic or evolutionary algorithm) so that learning processes and decision-making tasks are performed faster and more robust.

A. LATERAL PREFRONTAL CORTEX (LPFC)

The computational model of lateral PFC including dlPFC, vlPFC, aPFC allows that memory-based sorting or prediction tasks like a logical puzzle game are robustly solved by using spatiotemporal and emotional reasoning skills during rational social interaction. The observations o^t are fetched as system inputs $[o^d(t), o^v(t), o^a(t)]^T$. The states s^t are represented and computed by $[dl(t), vl(t), e(t)]^T$ from dlPFC, vlPFC and aPFC components of the lateral PFC model respectively. The computational model of lateral PFC can be described as transitions $(o^t, s^t, s^{t+1}, a^t, R^t)$ which are stored in replay memory D . The reward R^t is generated by the medial PFC model. The system outputs a^t which are derived in dlPFC are relayed to basal-ganglia and motor cortex regions. The state transition model is defined as $T(s, a^t, s') = P(s' | s, a)$ and the observation model is represented as $O(s^t, a^t, o^t) = P(o | s^t, a)$.

$$\begin{aligned}
 & T(s, a_t, s') \\
 &= P(s' | s, a) \\
 &= \begin{cases} P^{dlPFC}(dl(t+1)|e(t), vl(t), a(t)) & \sim N(\mu^{dl}, \sigma^{dl}) \\ P^{vlPFC}(vl(t+1)|e(t), dl(t)) & \sim N(\mu^{vl}, \sigma^{vl}) \\ P^{aPFC}(e(t+1)|dl(t), vl(t)) & \sim N(\mu^e, \sigma^e) \end{cases} \quad (8)
 \end{aligned}$$

$$\begin{aligned}
 &O(s', a_t, o^t) \\
 &= P(o|s', a) \\
 &= \begin{cases} P^{dlPFC}(od(t)|e(t+1), vl(t+1), a(t)) & \sim M(\mu^{dl}, \sigma^{dl}) \\ P^{vlPFC}(ov(t)|e(t+1), dl(t+1)) & \sim M(\mu^{vl}, \sigma^{vl}) \\ P^{aPFC}(oa(t)|dl(t+1), vl(t+1)) & \sim M(\mu^e, \sigma^e) \end{cases} \quad (9)
 \end{aligned}$$

As the exogenous stimuli, dlPFC accepts and process spatial data stream (dorsal or where pathway). In the computational model of lateral PFC, dorsolateral prefrontal cortex (dlPFC) enables to realize abilities of spatial cognitive skills including intentions, spatial planning, and decision-making tasks. These tasks are related to objects or events of spatial properties including coordinates, orientations, distances, dimensions and movements. Moreover, dlPFC involves in developing working memory related to spatial thinking. It is linked to ventrolateral, anterior (Brodmann area 10), and medial prefrontal cortex regions. As an output data stream, the connections between basal ganglia and dlPFC are associated with action selection on sequences of motor commands and procedural memory development. In addition, the connections between dlPFC and motor regions (or motor cortex) are responsible for the generation of motor commands (or actions). The dorsolateral part of the state transition $P^{dlPFC}(dl(t+I)| e(t), vl(t), a(t))$ and the observation models $P^{dlPFC}(od(t)| e(t+I), vl(t+I), a(t))$ are derived by a computational dynamic memory model including high level neural transactions such as field activations and spiking behaviors which are generated by the certain number of spiking neurons.

In the computational model of lateral PFC, temporal data stream (ventral or what pathway) which is an exogenous stimulus is processed by ventrolateral PFC (vlPFC). It allows realizing abilities of temporal cognitive skills including attention, temporal planning, and decision-making tasks. These tasks are related to objects or events of temporal properties including descriptions, shapes, colors. Furthermore, vlPFC is effective in developing the temporal part of working memory related to non-spatial thinking. It is connected to dorsolateral, anterior (Brodmann area 10), and medial prefrontal cortex regions. The neocortex inspired computational dynamic memory model including high level neural transactions such as field activations and spiking behaviors, is employed for derivation of $P^{vlPFC}(vl(t+I)| e(t), dl(t))$ and $P^{vlPFC}(ov(t)| e(t+I), dl(t+I))$ which are the ventrolateral PFC part of the state transition and the observation models respectively.

Emotion related cognitive skills like emotional memory processes, emotional reasoning, and inference tasks are performed by anterior prefrontal cortex (aPFC) module. The aPFC accepts exogenous input data stream including emotional responses released from the amygdala as a part of the limbic system. In addition, aPFC, providing excitatory or inhibitory control on dlPFC and vlPFC modules, decide how emotional information (states) reacts to behavior sequences. A computational dynamic memory model driven by the

neocortex like high level neural transactions such as field activations and spiking behaviors is utilized to obtain the anterior prefrontal cortex (aPFC) part of the state transition model $P^{aPFC}(e(t+I)| dl(t), vl(t))$ and the observation model $P^{aPFC}(oa(t)| dl(t+I), vl(t+I))$. The belief propagation is calculated by the following equation;

$$\begin{aligned}
 b_t(s') &= \tau(b_{t-1}, a_t, o^t) \\
 &= \eta O(s', a_t, o^t) \sum_{s \in S} T(s, a_t, s') b_{t-1}(s) \quad (10)
 \end{aligned}$$

The reward function $\rho(b, a)$ depending on the belief states is computed using the reward function based on the world states.

$$\rho(b, a) = \sum_{s \in S} R(s, a_t) b_t(s) \quad (11)$$

The Q values of the system are learned according to Bellman equation;

$$\begin{aligned}
 Q_t(b, a; \theta_i) &= \rho(b, a) + \gamma \sum_{b'} \tau(b_{t-1}, a_t, o^t) V(b') \quad (12) \\
 V_t(b) &= \max_a Q_t(b, a) \quad (13)
 \end{aligned}$$

The optimal value $V^t(b)$ is obtained by maximizing the Q values with respect to actions. The computational model of lateral PFC is optimized by minimizing the loss function. The loss function $L(b, a; \theta_i)$ to be minimized is written;

$$\begin{aligned}
 L(b, a; \theta_i) &= E_{b,a} \left[(Q_{target}(b, a; \theta_i) - Q_t(b, a; \theta_i))^2 \right] \quad (14) \\
 \theta_{i+1} &= \theta_i + \alpha \nabla_{\theta} L(b, a; \theta_i) \quad (15)
 \end{aligned}$$

A computational model of lateral PFC (Figure 5) consisting of a deep reinforcement learning algorithm with partially observable dynamics is constructed by a LSTM network embodied with convolutional layers. Where α is a parameter related to training speed. The network weights θ_i which are updated by the gradient descent algorithm provide to be obtained the optimal Q values via decreasing the difference between approximate $Q_t(b, a; \theta_i)$ and actual (target) Q values $Q_{target}(b, a; \theta_i)$ which are computed by Bellman equation. Obtained network output h^t is associated with approximate Q values. The observation information and the computed reward coming from mPFC are concatenated and represented with x^t which is fed into the LSTM. c^t indicates the cell memory.

Due to contribution of the cognitive skills including attention, intention, and emotion-based reasoning activities for solution of the rational social interaction problem with a humanoid robot, it is expected that durations of achieving simple sorting or prediction tasks in the puzzle game are decreased while the accuracy of the tasks is increasing during the experiments.

B. MEDIAL PREFRONTAL CORTEX (MPFC)

As one of the most critical regions, the medial PFC module which is composed of several subcortical areas including dmPFC, vmPFC (OFC), and ACC, has supervisor or regulator functionality on the lateral PFC. During a rational social interaction between human and humanoid robot, the states of the spatial-temporal and emotional reasoning skills are

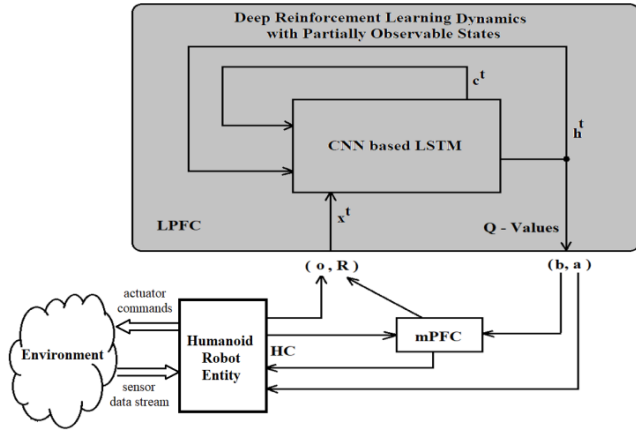


FIGURE 5. Machine learning model of lateral prefrontal cortex (LPFC).

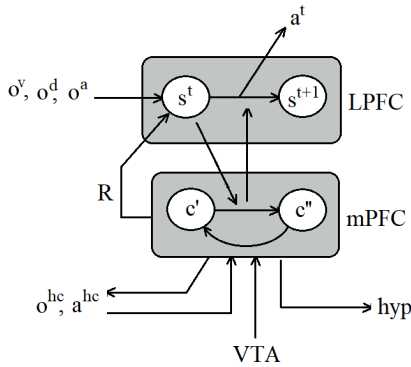


FIGURE 6. State transitions between LPFC and MPFC.

monitored by a computational model of mPFC (Figure 6) which is responsible for some cognitive and mental progress such as monitoring and reorganizing of complex decision making (e.g. planning or problem-solving), attention, meta-reasoning, associative working memory, adaptive learning, behavior execution, emotion regulation, multi-modal integration of action-reward and/or stimulus-reward association, language processes and expectation (prediction) tasks.

An arousal signal produced by mPFC is released to the hypothalamic module so that motivational gain (e.g., hypothalamic responses) modifies perceptual parameters for optimizing cognitive perception skills in the sensory cortex regions. In addition, past information related to long-term or episodic memory, coming from the hippocampus is observed for reward computation processes. Besides of that, distinctive information related to short-term or working memory, processed by modules of the PFC is broadcasted to the hippocampus model. In order to optimize the parameters of the reward computation process, the dopaminergic gain released by the ventral tegmental area (VTA) is employed by sub-modules in the computational medial PFC. As a property of meta-cognition on a rational social interaction scenario between the human and the humanoid robot, the reinforcement learning activities on the lateral PFC involving in spatiotemporal and emotional mental skills are

controlled via multi-modal rewards computed by sub-modules of mPFC. This approach serves like inverse reinforcement learning methodology realizing reward estimation function. At the first stage, the inference of the internal sub-states (c) extracted from $b(s^t)$ and hippocampal information (o_{hc}^t, a_{hc}^t) are estimated by the random forest model.

$$R^{acc}(s^t, a^t) = P(c|b, a_{hc}^t) \quad (16)$$

$$R^{ofc}(s^t, a^t) = P(c|b, o_{hc}^t) \quad (17)$$

As a part of multi-modal reward computation activities, the action-reward association functions $R^{acc}(s^t, a^t)$ are performed by the ACC module which is sensitive to past actions. A sequence of past actions shapes these activities. Also, the reward computation processes related to the stimulus-reward association functions $R^{ofc}(s^t, a^t)$, which are realized by the OFC module, are modified via a sequence of past observations (stimuli). The fusion of these associative functions (the stimulus-reward and the action-reward) is supervised by dmPFC. In addition, the dmPFC module undertakes the failure detection events (or conflict monitoring tasks).

In order to realize reasoning tasks which can create cause-effect relationships through these internal sub-states (c), a deep belief network (DBN) consisting of convolutional layers is utilized to train the rule-based fuzzy cognitive map, which executes the rule base with if-then like statements, which yields to consecutive rational (cause-effect) relations such as $(c' \xrightarrow{\varphi_i} c'')$, where φ_i defines the weight matrix of the network.

$$L(b, a; \varphi_i) = E_{b,a} \left[(R_{target}(b, a; \varphi_i) - R_t(b, a; \varphi_i))^2 \right] \quad (18)$$

This network is tuned by an evolutionary algorithm with respect to network cost function (fitness) so that its performance is guaranteed to increased or maintained. Its training procedure is based on a gradient descent algorithm.

$$\varphi_{i+1} = \varphi_i + \alpha \nabla_{\varphi} L(b, a; \varphi_i) \quad (19)$$

The reinforcement learning progress with computed reward released from mPFC is performed on the spatiotemporal and emotional reasoning skills which are hosted by modules of the lateral PFC. As a result of that, when the success rates of the memory-based simple sorting or prediction tasks in a rational social interaction scenario including a logical puzzle game are increasing, meta-cognitive planning and reasoning activities, convergence errors in reinforcement learning of lateral PFC related skills including attention, intention and emotion-based reasoning activities are decreased during the experiments.

V. EXPERIMENTAL SETUP AND IMPLEMENTATION

In this study, the main experiment platform is a Bioloid humanoid robot of Robotis (Figure 7) [89]. Bioloid is composed of 24 smart servo actuators (AX-12A), several peripheral body sensors (e.g. SparkFun 9 DoF Razor IMU M0, IR transmitters, 2 axis gyro, proximity sensor for distance measurement) and the main controller CM530 containing 32 bit Arm Cortex based microcontroller, external

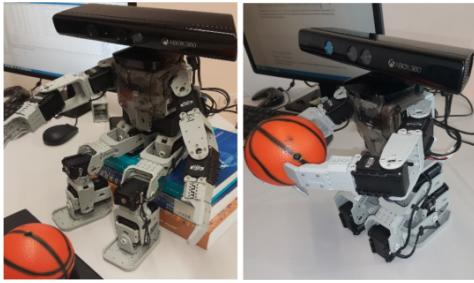


FIGURE 7. Humanoid robot partner: Bioid with Kinect mounted head.

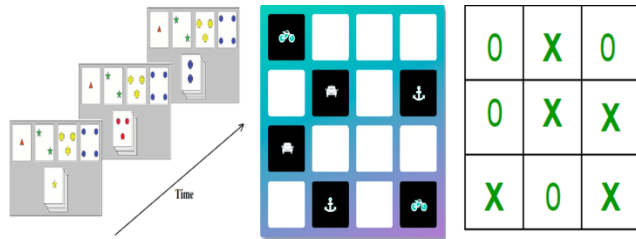


FIGURE 8. Wisconsin card sorting test (WCST), memory matching and Tic-Tac-Toe game.

I/O ports [89], [90]. In order to perform visual perception processes, the humanoid robot is supported with a ZED 3D stereo camera with 2K video, 6 DoF positional tracking, 20m depth spatial mapping [91]. As an audio environment, the additional equipment including a 2×2 microphone array, the stereo speaker is mounted on the robot's head with 2 DoF pan-tilt neck.

The computational workload of system architecture is hosted in a Nvidia Jetson TX2 embedded platform [92]. Brief specifications of the embedded system include 64 bit quad-core Arm Cortex A57 processor @2,1 GHz with 2 MB L2 cache and dual-core Denver processor @2,0 GHz with 2 MB L2 cache, 8 GB 128-bit DDR4 ram @59,9 GB/s, Nvidia Pascal architecture GPU with 256 CUDA cores and 32 GB eMMC, SATA storage. In order to control, connections to the humanoid robot can be realized by Bluetooth or USB 3,0 ports. We mostly preferred USB connection in the experiments.

As an implementation experiment (Figure 8), the interaction scenario is based on game like a logical puzzle including memory-based sorting and prediction tasks (e.g. Tic-Tac-Toe game, memory matching game and Wisconsin card sorting test (WCST)) which are used for evaluation of the cognitive functions and diagnosing the neurological disorders such as schizophrenia, dementia, and the other prefrontal cortex lesions [93]–[96]. In the WCST, people must categorize cards with respect to four different criteria such as color, shape, number, and symbol via the only feedback whether the classification is correct [97], [98]. The classification rule changing in every 10 cards has been expected to solve by the participant who has probability making one or more mistakes when the rule changes. The task evaluates how well participants can

adapt to the changing rules [99], [100]. First of all, the cards are randomly mixed up and placed as to face down on a $n \times n$ grid board in the memory matching game. In each move, the player who turns over any two cards should remember what was on each card and where it was. If the two cards match, keep them otherwise, turn them back over. Tic-tac-toe is a turn-taking logical game that the participant place an “X” card in a square while the opponent is placing an “O” card in the other square. In this game, the player who makes all “X” or “O” cards placed in a straight line (e.g. row, column, or crosswise), wins the match on a $n \times n$ grid board.

Achieving these tasks may involve some cognitive functions such as intention, attention, and emotion simultaneously [101]–[103]. The intention which involves requests of the action sequence is related to spatial skills originated by the dlPFC module. The attention which requires temporal focus and contextual attributes of an object (or event), is related to temporal skills hosted by the vlPFC module. As a control mechanism allowing the regulation on the behaviors, the emotional reasoning skill driven by the aPFC module associates the events (e.g. experiences, memory fragment during the interaction) with the artificial feelings (emotional states) such as satisfaction (e.g. well-being), excitation and inhibition (e.g. frustration). Then, a logical puzzle game including memory-based sorting or prediction tasks shaped according to this emotional modulation (control). Meta-cognition property of the computational model of the mPFC module allows supervision over spatial-temporal and emotional reasoning skills which are ensured by lateral PFC.

The performances of the spatial-temporal and emotional reasoning skills which are ensured by the functions of the computational lateral PFC are tested in the first experiment. While prediction works are testing the non-spatial or temporal decision-making skills which are related to attentional abilities, the spatial planning skills associated with intention-based functions are evaluated by the sorting tasks in some kind of logical puzzle like memory-based interaction game. During the solving of these tasks, emotional states of the humanoid robot are monitored so that the emotional reasoning functions can provide decisional bias (inference) to the mentioned skills of the computational lateral PFC. Durations and accuracies are measured through the experiment. At the end of the experiment, it is expected to verify the hypothesis (H1) via the contribution of the cognitive functions (e.g., attention, intention, and emotion-based reasoning) in the lateral PFC model for the solution of the rational social interaction problem with a robot.

In the second experiment, the efficiency of the meta-cognition property of the computational model of mPFC is evaluated via reward-based monitoring activities on the functions of the computational lateral PFC model. The stimulus associated reward activities originated in the OFC module are examined using human feedback events perceived by the humanoid robot in the interaction game. The action related reward activities derived from the ACC module are dealt with the humanoid robot's social clues through the human.

Then the performance of managing both functionalities handled by the dmPFC module is evaluated so that self-awareness and consciousness are investigated. During the experiment, success rates and convergence errors are observed periodically. It is promised to justify the hypothesis (H2) via the contribution of the reward computation based meta-cognition functions which are hosted by the mPFC model at the end of the experiment.

For the last experiment, the performances of all mental operations such as associative learning, complex decision-making and meta-cognitive planning are improved by an optimization algorithm that will ensure neural plasticity. In addition, the parameters including learning rate, reward discount factor, and exploration rate are adaptively tuned via dopaminergic gain. Moreover, depending upon different conditions, the dopamine modulation is parametrically driven by synaptic neural plasticity as tuning convergence speed. Training durations, convergence speeds, and learning performances are monitored during the experiment.

At the end of the experiments, the robot reports learning and interaction statistics. Under the supervision of the operator, levels of rehabilitation, and contribution of the neuro-cognitive architecture embodied in the humanoid robot are elaborately discussed.

VI. EXPERIMENTAL RESULTS

The computational prefrontal cortex system framework for a humanoid robot was modeled, simulated, and tested in the robot operating system (ROS) middleware with Kinetic distro run over Ubuntu 16.04 LTS operating system. This framework which is a ROS package is composed of many nodes written by python (rospy) and C++ (roscpp) client libraries of ROS [104]. For image processing and computer vision tasks, OpenCV and PointCloud libraries were preferred [105], [106]. In addition, TensorFlow library which is a machine learning framework was utilized to handle neural networks and deep learning applications [107]. Also, there exists a software development kit (SDK/API) for a bio-robot platform. For communication between the hardware of the humanoid robot and SDK, the firmware of bio-robot platform is updated before the experiments. The developed system is visualized by several 2D/3D graphical user interface (GUI) tools of ROS (e.g. Gazebo simulator, rviz visualization environment, rqt_graph, rqt_plot, rqt_bags)

The snapshots of the experiments implementing human-robot interaction scenarios are shown in figure 9. In the conducted experiments, the experiment scene (Figure 10) that the humanoid robot and human are presented as face to face is considered in where several game cards with different features (e.g. size, shapes, colors, number, location, and symbol, etc.) are located on the spatial workspace (e.g. small platform or desk) which resides between the human and the robot. In these experiment scenes, Wisconsin card sorting test (WCST), memory matching and Tic-Tac-Toe game scenarios were executed so that the memory-based logical puzzle game

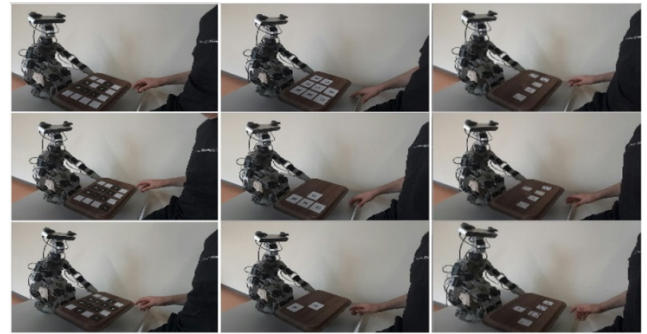


FIGURE 9. The scenes of the human-robot interaction experiments.

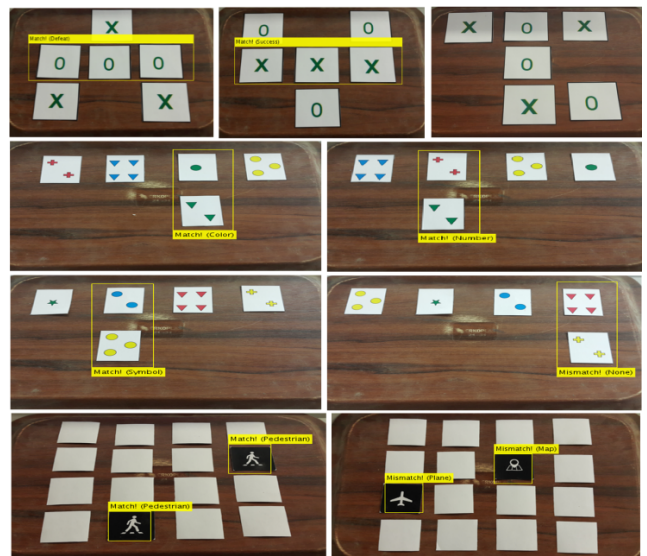


FIGURE 10. The preliminary operations and basic detections.

involving sorting and prediction tasks were implemented during the experiments.

Audio, visual, and set of joint information are received by the humanoid robot platform. While visual inputs are regarding as saliences (e.g. ID, color, location), gaze direction, and skeletal information of human for hand pointing for object or gesture recognition, audio input is evaluated as sound localization and speech recognition. The internal stimuli of the robot's body are joint angles, gyro, and proximity information. The robot platform can execute some gestural behaviors like a hand pointing as output actions (or motor commands). In addition to that, the system enables the robot to realize speech communication with a human.

In the experiments, turn-taking interaction tasks including WCST, memory game, and tic-tac-toe game were spontaneously performed between a humanoid robot and a human participant. To verify the proposed hypotheses, the computational prefrontal cortex model was realized and evaluated for a humanoid robot. The data presented in the table 1 provide indicators of performance evaluation, including ability scores for a performed activity and average response times.

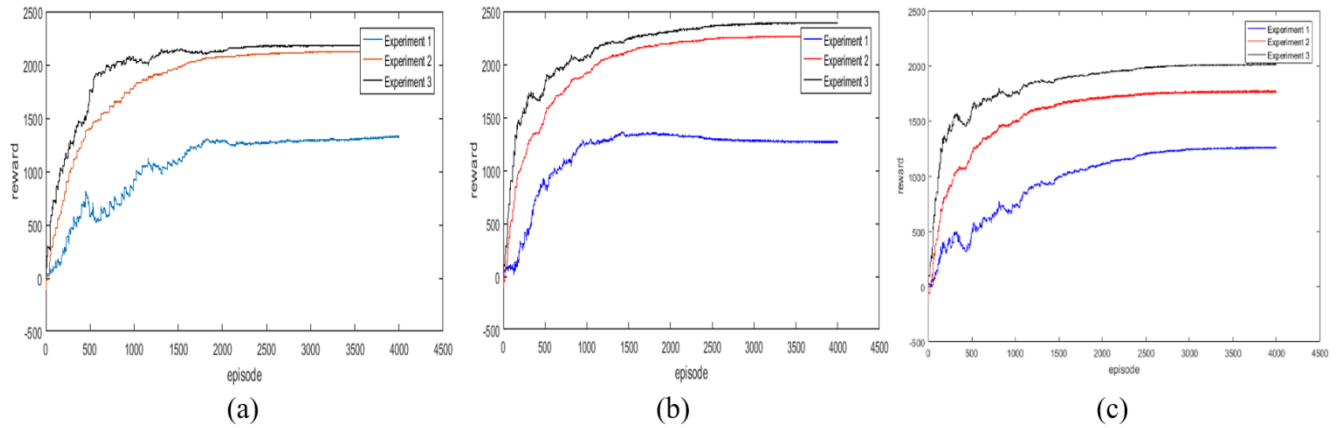


FIGURE 11. The reward trends of the experiments for (a) memory game, (b) WCST and (c) Tic-Tac-Toe.

TABLE 1. Accuracy and response time of interaction tasks.

Emotional State	Spatial-Temporal Events	Memory game		WCST		Tic-Tac-Toe	
		μ	τ	μ	τ	μ	τ
Excited	Content Matching	81	12	86	13	88	24
	Direction Prediction	77	9	79	6	84	15
Relaxed	Content Matching	75	14	72	16	79	31
	Direction Prediction	71	11	74	9	83	17
Bored	Content Matching	64	39	67	38	62	35
	Direction Prediction	59	34	58	29	54	32
Stressed	Content Matching	35	27	33	26	36	17
	Direction Prediction	22	31	24	22	19	13

Where μ is accuracy (percent) and τ is average Response Time (sec).

In order to investigate H1, the performance of the LPFC model functions including working memory, attention, intention, and emotional reasoning was evaluated using the average of the response times and the accuracy scores.

The accuracy scores were computed by “(achieved_activity_count) \times 100 / (total_activity)” in the experiments. Besides of that, the success rates were appraised by “1/(error_rate)” for a task (e.g. memory game, WCST and tic-tac-toe). The state-based events such as content matching evaluate the performance of attention skills associated with vIPFC. The performance of the intention skills related to dIPFC is dealt with the movement-based events (direction prediction). The emotions including “excited”, “relaxed”, “bored” and “stressed” are triggered by the responses of spatiotemporal events. Accordingly, the average response times of achieving basic sorting or prediction tasks decreased, meanwhile the average accuracy scores of the tasks increased during the first experiment, since all of these skills, including attention, intention and emotion-based reasoning behaviors, are simultaneously utilized to help solve the problem involving LPFC based planning tasks.

The performance of the meta-cognition mechanism in a computational model of medial PFC referred to in H2, was investigated by observing success rates (reward)

in figure11, and convergence errors (cost) in figure 12 for the experiment scenarios including memory game (figure 11.a), WCST (figure 11.b) and tic-tac-toe (figure 11.c). When the results are examined, the rewards are slightly better in the case of experiment 3 with respect to ones in case of experiment 2, while the rewards in experiment 1 are remaining behind of both experiments for all scenarios. The reward trend in the WCST task seems to have slightly less fluctuation according to the other tasks. The most oscillations occurred in the memory game task. The convergence errors in reinforcement learning of lateral PFC related skills including attention, intention and emotion-based reasoning activities decreased while success rates of the simple sorting or prediction tasks in the puzzle game are increasing during the experiments, because of that the meta-cognition mechanism in a computational model of medial PFC is integrated appropriate rewards to all mental functions of lateral PFC for regulating complex decision making, meta-cognitive planning and reasoning activities.

The learning performances involving training steps (episode), convergence speeds (cost) and are monitored during the experiments (Figure 12). The results show that costs in the reinforcement learning process of mental functions such as attention, intention, and emotion-based reasoning activities are decreased more efficiently by the meta-cognition mechanism integrating optimization mechanism for optimizing learning performances of the cognitive skills in experiment 3. The effectiveness of this mechanism is better for tic-tac-toe (figure 12.c) with respect to the other both scenarios. Despite the fact that there is a small difference between memory game (figure 12.a) and WCST (figure 12.b), the cost in the memory game task is slightly worse than the cost of the WCST task. Using this optimization algorithm, it is shown that the learning process is faster and the decision-making activities are more robust.

Size of convolution filter (weight tensor) is chosen as a stack of 4-time frames with a dimension of $84 \times 110 \times 2$ representing the image pattern of width, height, and depth

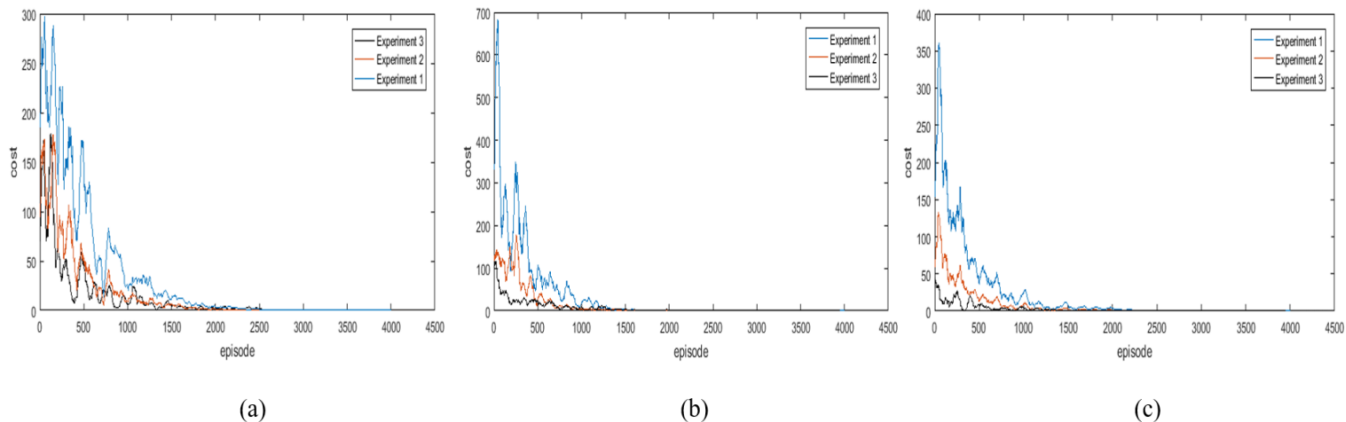


FIGURE 12. The costs: (a) memory game, (b) WCST and (c) Tic-Tac-Toe.

TABLE 2. LPFC model performances of interaction tasks.

No	Models for spatial-temporal planning tasks	Memory game		WCST		Tic-Tac-Toe	
		Tr	Ts	Tr	Ts	Tr	Ts
1	DQN	58	53	56	51	57	55
2	DDQN	74	72	75	68	76	64
3	DQN w/LSTM	76	75	77	69	79	68
4	Deep POMDP w/LSTM	86	77	80	73	87	75
5	Model 4 + Dynamic field with spiking neural behaviors	88	78	82	74	91	77

Where Tr represents train (percent) and Ts is test (percent).

respectively. Alpha coefficient known as the learning rate is 0,00025. The exploration probability which is constrained its minimum by 0,01, is 1,0 at the start and its decay rate is 0,00001. The reward discounting rate (gamma) is 0,9.

In table 2, different models related to LPFC are evaluated with respect to training performances. In order to achieve intention, attention and emotional reasoning skills, deep reinforcement learning centered models such as double deep Q networks (DDQN), DQN with LSTM, partially observable deep reinforcement learning with LSTM, and spiking neural dynamics are compared to show supremacy between them. According to results, although the results were close to each other, it is seen that there exist three bands when examined in detail. At first band, model 1 consist of DQN recorded the worst result with average accuracy %56-58 for train session and %51-55 for the test session. As the second band, the model 2-3 involving DDQN and DQN with LSTM respectively, acquire better scores than model 1. In the last band, deep POMDP with LSTM (model 4) and model 4 with spiking neural dynamics as model 5 gained better results from the other models. The best performance was slightly achieved via model 5 with average accuracy %82-91 for the train session and %74-78 for the test session.

The advantage of the models used in mPFC is presented by comparison with different models in table 3. As the inverse

TABLE 3. Reward generation model performances in mPFC.

No	Reward generation model (meta-cognition)	Memory game		WCST		Tic-Tac-Toe	
		Tr	Ts	Tr	Ts	Tr	Ts
1	DBN w/Random Forest	73	64	65	59	67	62
2	DBN w/ Rule-Based FCM	67	63	61	58	64	60
3	DBN w/ Random Forest Based FCM	81	70	77	66	79	65
4	Model 3 + GA optimization	87	72	79	67	84	71

Where Tr means train (percent) and Ts stands for test (percent).

reinforcement learning technique, various types of reward generator models composed of deep belief networks (DBN) with a random forest model, rule-based fuzzy cognitive map (FCM), and an optimization mechanism (genetic algorithm) are evaluated so that meta-cognition has been achieved. When the results are investigated, it is easily seen that the third model with GA optimization with average accuracy %79-87 for train session and %67-71 for test session has a remarkable advantage with respect to the other models. While DBN with random forest model gave a better result than the model with rule-based FCM, the scores of DBN with random forest-based FCM model prevailed against model 1 and model 2.

VII. CONCLUSION

The brain-inspired computational modeling approaches are expected to make huge impacts on human-robot interaction studies since as a synthetic life form, the socially aware robots equipped with neuro-cognitive architecture will be widely used in social areas such as assistive, entertainment and rehabilitation fields. Therefore, all these studies to enable human like cognition for the socially aware robots will completely improve the living standards of people [108]–[110]. As a result, this study is very important in that humanoid robots with enhanced embodied cognitive abilities can be used to assist disabled individuals struggling

to interact with their social environment by guiding their accessibility and communication. In addition, it is promised that this study will lead to a progressive impact on human-robot interaction researches achieving computational approximation of consciousness and human like cognition for social aware robots. These challenges which are mentioned require recursive task processing and meta-cognitive reasoning mechanism. Naturally, the human brain realizes these cognitive skills by prefrontal cortex which is a part of neocortex.

In this study, a new computational framework of the prefrontal cortex model which is composed of several sub-modules including dorsolateral, ventrolateral, anterior, and medial prefrontal regions was developed and tested for human-robot interaction (HRI). The new computational framework was embedded in a humanoid robot platform. The major novelty related to this framework is to develop a computational representation of the human brain enabling artificial consciousness and imitating self-awareness so that the humanoid robot achieved human level mental activities such as complex decision-making, goal-oriented behavior planning, and meta-cognitive reasoning until it attained the optimal goal state during the experiments involving human-robot interaction scenarios. In order to realize spatio-temporal and emotional reasoning skills involving attention, intention, short-term memory, decision making (e.g. arithmetic and logical), planning, analysis of cause-effect relations and problem solving, the computational prefrontal cortex framework based on the neuromorphic foundations of human's mental activities utilizes deep reinforcement learning algorithm with partially observable state dynamics and LSTM for LPFC model. As a reward computing mechanism, the activities of LPFC are monitored and regulated by the mPFC model of the computational prefrontal cortex framework, which employs deep belief network, rule-based fuzzy cognitive map (alternatively fuzzy random forest model) and genetic algorithm to achieve meta-cognitive reasoning. In addition, the mental activities in the computational framework constitute a working memory stored in the weights of an attractor network, as a sort of associative memory. The network weights in this working memory are updated by machine learning algorithms corresponding with the prefrontal cortex model. Mental functions such as emotional responses and spatiotemporal behavioral planning activities are observed during the experiments. The developed computational framework executing interaction scenarios worked on the software environment including related libraries (SDK/API) and ROS middleware hosted by Ubuntu operating system.

Three experiments were performed to verify the proposed architecture and related hypotheses H1 and H2. The experiments contain the interaction game scenario based on a logical puzzle including memory-based sorting and prediction tasks (e.g. Tic-Tac-Toe game, memory matching game, and Wisconsin card sorting test (WCST)). The first experiment deals with the performances of the spatio-temporal and emotional reasoning skills which are ensured by the functions

of the computational lateral PFC. During experiment1, it is seen that durations of achieving simple sorting or prediction tasks in the puzzle game are decreased while the accuracy of the tasks is increasing during the experiments. Thus, the hypothesis (H1) is verified by the contribution of the cognitive functions (e.g., attention, intention, and emotion-based reasoning) in the lateral PFC model for the solution of the rational social interaction problem with a humanoid robot. The reward generation based meta-cognition activities on the functions of the computational lateral PFC model evaluates the efficiency of the monitoring property of the computational model of mPFC in the second experiment. At the end of the experiment2, the fact that meta-cognitive planning and reasoning activities, convergence errors in reinforcement learning of lateral PFC related skills including attention, intention and emotion-based reasoning activities are decreased is observed while success rates of the simple sorting or prediction tasks in the puzzle game are increasing during the experiments. Therefore, the hypothesis (H2) is validated by the contribution of the reward computation based meta-cognition functions which are hosted by the mPFC model. In addition, the set of hyperparameters as a dopaminergic gain vector including learning rate, reward discount factor and exploration rate are adaptively tuned by an optimization algorithm. This mechanism ensuring neural plasticity makes significant enhancement on the performances of all mental operations such as associative learning, complex decision-making and meta-cognitive planning activities in the last experiment. Training durations, convergence errors (costs) and learning performances (rewards) are monitored during the experiment. At the end of the experiments, the humanoid robot reports learning and interaction statistics.

As a reverse engineering perspective of artificial intelligence, this approach may be extended to computational modeling of cognitive and mental functions related to various cortical regions such as limbic system (components including basal ganglia, amygdala, hippocampus, and hypothalamus), motor cortex, cerebellum, brain stem, parietal, temporal and occipital (visual cortex) lobes so that whole brain model is realized for a humanoid robot. In future studies, I will have pursued to investigate the computational realization of the other neuro-cognitive functions for meta-cognition, which are the key issues for modeling consciousness and self-awareness during a rational social interaction scenario between human and humanoid robot as beyond deliverables of this research. In addition, computational methods for adaptation of additional tuning in hyperparameters may provide a more detailed understanding of the processes involved. Beside of that, this proposed architecture could be further improved in future research, by integrating the personality model to brain-inspired neuro-cognitive architecture for the humanoid robot. Also, it may be considered that these efforts will build artificial life organisms (systems) and lead to achieve artificial general intelligence in the long future.

ACKNOWLEDGMENT

The Jetson TX2 embedded development platform used for this research was donated by NVIDIA Corporation.

REFERENCES

- [1] T. Chaminade and G. Cheng, "Social cognitive neuroscience and humanoid robotics," *J. Physiol.-Paris*, vol. 103, nos. 3–5, pp. 286–295, May 2009.
- [2] M. Asada, K. F. MacDorman, H. Ishiguro, and Y. Kuniyoshi, "Cognitive developmental robotics as a new paradigm for the design of humanoid robots," *Robot. Auto. Syst.*, vol. 37, nos. 2–3, pp. 185–193, Nov. 2001.
- [3] J. A. Reggia, G. E. Katz, and G. P. Davis, "Humanoid cognitive robots that learn by imitating: Implications for consciousness studies," *Frontiers Robot. AI*, vol. 5, p. 1, Jan. 2018.
- [4] A. Wykowska, T. Chaminade, and G. Cheng, "Embodied artificial agents for understanding human social cognition," *Philos. Trans. Roy. Soc. B, Biol. Sci.*, vol. 371, no. 1693, May 2016, Art. no. 20150375.
- [5] A. Wykowska, G. Metta, C. Becchio, R. Hortensius, and E. Cross, "Cognitive and social neuroscience methods for HRI," in *Proc. Companion ACM/IEEE Int. Conf. Hum.-Robot Interact.*, Mar. 2018, pp. 405–406.
- [6] J. D. Kralik, D. B. C. Muldrew, D. Gunasekaran, and R. D. Lange, "Cognitive and action control for goal-directed reaching in a humanoid robot," in *Proc. IEEE Int. Conf. Robot. Biomimetics (ROBIO)*, Dec. 2017, pp. 1868–1873.
- [7] J. A. Weyhenmeyer and E. A. Gallman, *Rapid Review of Neuroscience*. Amsterdam, The Netherlands: Elsevier, 2007, p. 102.
- [8] M. D. Lieberman, "Social cognitive neuroscience: A review of core processes," *Annu. Rev. Psychol.*, vol. 58, no. 1, pp. 259–289, Jan. 2007.
- [9] E. E. Smith, S. M. Kosslyn, and L. W. Barsalou, *Cognitive Psychology: Mind and Brain*, vol. 6. Upper Saddle River, NJ, USA: Prentice-Hall, 2007.
- [10] R. S. Frackowiak, *Human Brain Function*. Amsterdam, The Netherlands: Elsevier, 2004.
- [11] B. Kolb and I. Q. Whishaw, *An Introduction to Brain and Behavior*. New York, NY, USA: Worth Publishers, 2001.
- [12] G. E. Dawson and K. W. Fischer, *Human Behavior and the Developing Brain*. New York, NY, USA: Guilford Press, 1994.
- [13] A. Konar, *Artificial Intelligence and Soft Computing: Behavioral and Cognitive Modeling of the Human Brain*. Boca Raton, FL, USA: CRC Press, 1999.
- [14] J. Fuster, *The Prefrontal Cortex*. New York, NY, USA: Academic, 2015.
- [15] E. K. Miller and J. D. Cohen, "An integrative theory of prefrontal cortex function," *Annu. Rev. Neurosci.*, vol. 24, no. 1, pp. 167–202, Mar. 2001.
- [16] J. N. Wood and J. Grafman, "Human prefrontal cortex: Processing and representational perspectives," *Nature Rev. Neurosci.*, vol. 4, no. 2, pp. 139–147, Feb. 2003.
- [17] M. A. Goodrich, "Using models of cognition in HRI evaluation and design," in *Proc. AAAI Fall Symp. Ser., Intersection Cogn. Sci. Robot., Interfaces Intell.*, Oct. 2004, pp. 21–24.
- [18] P. Langley, J. E. Laird, and S. Rogers, "Cognitive architectures: Research issues and challenges," *Cognit. Syst. Res.*, vol. 10, no. 2, pp. 141–160, Jun. 2009.
- [19] M. A. Goodrich and A. C. Schultz, "Human-robot interaction: A survey," *Found. Trends Hum.-Comput. Interact.*, vol. 1, no. 3, pp. 203–275, 2007.
- [20] A. Tapus, M. Maja, and B. Scassellatti, "The grand challenges in socially assistive robotics," *IEEE Robot. Automat. Mag.*, vol. 14, no. 1, 2007.
- [21] K. Dautenhahn, "Socially intelligent robots: Dimensions of human-robot interaction," *Philos. Trans. Roy. Soc. B, Biol. Sci.*, vol. 362, no. 1480, pp. 679–704, 2007.
- [22] B. Robins and K. Dautenhahn, "Developing play scenarios for tactile interaction with a humanoid robot: A case study exploration with children with autism," in *Proc. Int. Conf. Social Robot.* Berlin, Germany: Springer, Nov. 2010, pp. 243–252.
- [23] M. Häring, J. Eichberg, and E. André, "Studies on grounding with gaze and pointing gestures in human-robot-interaction," in *Proc. Int. Conf. Social Robot.* Berlin, Germany: Springer, Oct. 2012, pp. 378–387.
- [24] E. Mower, D. J. Feil-Seifer, M. J. Mataric, and S. Narayanan, "Investigating implicit cues for user state estimation in human-robot interaction using physiological measurements," in *Proc. 16th IEEE Int. Symp. Robot Human Interact. Commun. (RO-MAN)*, 2007, pp. 1125–1130.
- [25] C. Clabaugh and M. Mataric, "Escaping oz: Autonomy in socially assistive robotics," *Annu. Rev. Control, Robot., Auto. Syst.*, vol. 2, no. 1, pp. 33–61, May 2019.
- [26] M. Patacchiola, "A developmental model of trust in humanoid robots," Doctoral dissertation, Univ. Plymouth, Plymouth, U.K., 2018.
- [27] M. Asada, "Artificial empathy," in *Proc. 31st Int. Congr. Psychol. Diversity Harmony-Insights Psychol.* Chichester, U.K.: Wiley, Oct. 2018, pp. 19–41.
- [28] M. Khamassi, R. Quilodran, P. Enel, P. F. Dominey, and E. Procyk, "Behavioral regulation and the modulation of information coding in the lateral prefrontal and cingulate cortex," *Cerebral Cortex*, vol. 25, no. 9, pp. 3197–3218, Sep. 2015.
- [29] M. Khamassi, P. Enel, P. F. Dominey, and E. Procyk, "Medial prefrontal cortex and the adaptive regulation of reinforcement learning parameters," in *Progress in Brain Research*, vol. 202. Amsterdam, The Netherlands: Elsevier, 2013, pp. 441–464.
- [30] M. Khamassi, C. Wilson, R. Rothé, R. Quilodran, P. F. Dominey, and E. Procyk, "Meta-learning, cognitive control, and physiological interactions between medial and lateral prefrontal cortex," in *Proc. Neural Basis Motivational Cogn. Control*, 2011, pp. 351–370.
- [31] M. Khamassi, S. Lallée, P. Enel, E. Procyk, and P. F. Dominey, "Robot cognitive control with a neurophysiologically inspired reinforcement learning model," *Frontiers Neurobotics*, vol. 5, p. 1, Jul. 2011.
- [32] M. Khamassi, R. Quilodran, P. Enel, E. Procyk, and P. F. Dominey, "A computational model of integration between reinforcement learning and task monitoring in the prefrontal cortex," in *Proc. Int. Conf. Simulation Adapt. Behav.* Berlin, Germany: Springer, Aug. 2010, pp. 424–434.
- [33] M. Khamassi, R. Quilodran, P. Enel, P. Dominey, and E. Procyk, "Role of the frontal cortex in solving the exploration-exploitation trade-off," in *Proc. Cinquième conférence plénière française de Neurosciences Computationnelles (Neurocomp)*, Aug. 2010.
- [34] M. Khamassi, "Complementary roles of the rat prefrontal cortex and striatum in reward-based learning and shifting navigation strategies," Doctoral dissertation, Université Pierre et Marie Curie-Paris VI, 2007.
- [35] I. Yildirim, J. Wu, N. Kanwisher, and J. Tenenbaum, "An integrative computational architecture for object-driven cortex," *Current Opinion Neurobiol.*, vol. 55, pp. 73–81, Apr. 2019.
- [36] D. G. Gómez-Martínez, J.-H. Rosales, V. Muñoz-Jiménez, F. Ramos, and M. Ramos, "A bio-inspired self-responding emotional behavior system for virtual creatures," *Biologically Inspired Cognit. Archit.*, vol. 26, pp. 26–40, Oct. 2018.
- [37] D. Hernández García, S. Adams, A. Rast, T. Wennekers, S. Furber, and A. Cangelosi, "Visual attention and object naming in humanoid robots using a bio-inspired spiking neural network," *Robot. Auto. Syst.*, vol. 104, pp. 56–71, Jun. 2018.
- [38] J.-A. Cervantes, J.-H. Rosales, S. López, F. Ramos, and M. Ramos, "Integrating a cognitive computational model of planning and decision-making considering affective information," *Cognit. Syst. Res.*, vol. 44, pp. 10–39, Aug. 2017.
- [39] H. Mizutani, M. Ueno, N. Arakawa, and H. Yamakawa, "Whole brain connectomic architecture to develop general artificial intelligence," *Procedia Comput. Sci.*, vol. 123, pp. 308–313, Jan. 2018.
- [40] P. F. M. J. Verschure, "Distributed adaptive control: A theory of the mind, brain, body nexus," *Biologically Inspired Cognit. Archit.*, vol. 1, pp. 55–72, Jul. 2012.
- [41] G. Maffei, D. Santos-Pata, E. Marcos, M. Sánchez-Fibla, and P. F. M. J. Verschure, "An embodied biologically constrained model of foraging: From classical and operant conditioning to adaptive real-world behavior in DAC-X," *Neural Netw.*, vol. 72, pp. 88–108, Dec. 2015.
- [42] Y. Zeng, Y. Zhao, J. Bai, and B. Xu, "Toward robot self-consciousness (II): Brain-inspired robot bodily self model for self-recognition," *Cogn. Comput.*, vol. 10, no. 2, pp. 307–320, 2018.
- [43] D. Vernon, M. Beetz, and G. Sandini, "Prospection in cognition: The case for joint episodic-procedural memory in cognitive robotics," *Frontiers Robot. AI*, vol. 2, p. 19, Jul. 2015.
- [44] A. V. Samsonovich, "Emotional biologically inspired cognitive architecture," *Biologically Inspired Cognit. Archit.*, vol. 6, pp. 109–125, Oct. 2013.
- [45] B. Denizdurduran and N. S. Sengor, "A realization of goal-directed behavior-implementing a robot model based on cortico-striato-thalamic circuits," in *Proc. ICAART*, 2012, pp. 289–294.

- [46] N. Navarro-Guerrero, R. Lowe, and S. Wermter, "A neurocomputational amygdala model of auditory fear conditioning: A hybrid system approach," in *Proc. Int. Joint Conf. Neural Netw. (IJCNN)*, Jun. 2012, pp. 1–8.
- [47] A. Lieto, M. Bhatt, A. Oltramari, and D. Vernon, "The role of cognitive architectures in general artificial intelligence," *Cognit. Syst. Res.*, vol. 48, pp. 1–3, May 2018.
- [48] A. Marchetti, F. Manzi, S. Itakura, and D. Massaro, "Theory of mind and humanoid robots from a lifespan perspective," *Zeitschrift für Psychologie*, vol. 226, no. 2, pp. 98–109, Apr. 2018.
- [49] M. Van Gerven and S. Bohte, Eds., "Artificial neural networks as models of neural information processing," *Frontiers Comput. Neurosci.*, vol. 11, p. 114, Dec. 2017.
- [50] Z. Bing, C. Meschede, F. Röhrbein, K. Huang, and A. C. Knoll, "A survey of robotics control based on learning-inspired spiking neural networks," *Frontiers Neurobotics*, vol. 12, p. 35, Jul. 2018.
- [51] M. G. Johnson and S. Chartier, "Spike neural models part II: Abstract neural models," *Quant. Methods Psychol.*, vol. 14, no. 1, pp. 1–16, 2018.
- [52] S. Ahmadizadeh, P. J. Karoly, D. Nešić, D. B. Grayden, M. J. Cook, D. Soudry, and D. R. Freestone, "Bifurcation analysis of two coupled jansen-rit neural mass models," *PLoS ONE*, vol. 13, no. 3, Mar. 2018, Art. no. e0192842.
- [53] J. P. Koren, J. Herta, F. Fürbass, S. Pirker, V. Reiner-Deitemyer, F. Riederer, and C. Baumgartner, "Automated long-term EEG review: Fast and precise analysis in critical care patients," *Frontiers Neurol.*, vol. 9, p. 454, Jun. 2018.
- [54] E. Balaguer-Ballester, R. Moreno-Bote, G. Deco, and D. Durstewitz, "Metastable dynamics of neural ensembles," *Frontiers Syst. Neurosci.*, vol. 11, p. 99, Jan. 2018.
- [55] R. T. Thibault, A. MacPherson, M. Lifshitz, R. R. Roth, and A. Raz, "Neurofeedback with fMRI: A critical systematic review," *NeuroImage*, vol. 172, pp. 786–807, May 2018.
- [56] L. Bréchet, D. Brunet, G. Birot, R. Gruetter, C. M. Michel, and J. Jorge, "Capturing the spatiotemporal dynamics of self-generated, task-initiated thoughts with EEG and fMRI," *NeuroImage*, vol. 194, pp. 82–92, Jul. 2019.
- [57] G. Schönner and J. P. Spencer, *Dynamic Thinking: A Primer on Dynamic Field Theory*. Oxford, U.K.: Oxford Univ. Press, 2016.
- [58] M. A. Giese, *Dynamic Neural Field Theory for Motion Perception*, vol. 469. Cham, Switzerland: Springer, 2012.
- [59] R. S. Zemel and M. C. Mozer, "Localist attractor networks," *Neural Comput.*, vol. 13, no. 5, pp. 1045–1064, May 2001.
- [60] M. L. Puterman, *Markov Decision Processes: Discrete Stochastic Dynamic Programming*. Hoboken, NJ, USA: Wiley, 2014.
- [61] R. E. Bellman and S. E. Dreyfus, *Applied Dynamic Programming*, vol. 2050. Princeton, NJ, USA: Princeton Univ. Press, 2015.
- [62] S. M. Ross, *Introduction to Stochastic Dynamic Programming*. New York, NY, USA: Academic, 2014.
- [63] C. C. White, *Markov Decision Processes*. New York, NY, USA: Springer, 2001, pp. 484–486.
- [64] L. R. Rabiner, "A tutorial on hidden Markov models and selected applications in speech recognition," *Proc. IEEE*, vol. 77, no. 2, pp. 257–286, Feb. 1989.
- [65] A. Gelman, H. S. Stern, J. B. Carlin, D. B. Dunson, A. Vehtari, and D. B. Rubin, *Bayesian Data Analysis*. London, U.K.: Chapman & Hall, 2013.
- [66] F. V. Jensen, *An Introduction to Bayesian Networks*, vol. 210. London, U.K.: UCL Press, 1996, pp. 1–178.
- [67] D.-S. Lee, J. J. Hull, and B. Erol, "A Bayesian framework for Gaussian mixture background modeling," in *Proc. Int. Conf. Image Process.*, vol. 3, Sep. 2003, p. III-973.
- [68] S. R. Safavian and D. Landgrebe, "A survey of decision tree classifier methodology," *IEEE Trans. Syst., Man, Cybern.*, vol. 21, no. 3, pp. 660–674, May 1991.
- [69] L. Breiman, "Random forests," *Mach. Learn.*, vol. 45, no. 1, pp. 5–32, 2001.
- [70] J. Ali, R. Khan, N. Ahmad, and I. Maqsood, "Random forests and decision trees," *Int. J. Comput. Sci. Issues*, vol. 9, no. 5, p. 272, 2012.
- [71] G. J. Klir and B. Yuan, *Fuzzy Sets and Fuzzy Logic: Theory and Applications*. Upper Saddle River, NJ, USA, 1995, p. 563.
- [72] B. Kosko, "Fuzzy cognitive maps," *Int. J. Man-Mach. Stud.*, vol. 24, no. 1, pp. 65–75, 1986.
- [73] H.-J. Song, Z.-Q. Shen, C.-Y. Miao, Z.-Q. Liu, and Y. Miao, "Probabilistic fuzzy cognitive map," in *Proc. IEEE Int. Conf. Fuzzy Syst.*, 2006, pp. 1221–1228.
- [74] J. P. Carvalho and J. A. B. Tomé, "Rule-based fuzzy cognitive maps-fuzzy causal relations," in *Computational Intelligence for Modelling, Control, and Automation*, M. Mohammadian, Ed. 1999.
- [75] J. P. Carvalho and J. A. B. Tomé, "Rule based fuzzy cognitive maps-qualitative systems dynamics," in *Proc. 19th Int. Conf. North Amer. Fuzzy Inf. Process. Soc. (NAFIPS)*, Jul. 2000, pp. 407–411.
- [76] E. I. Papageorgiou, C. D. Stylios, and P. P. Groumpos, "Active hebbian learning algorithm to train fuzzy cognitive maps," *Int. J. Approx. Reasoning*, vol. 37, no. 3, pp. 219–249, Nov. 2004.
- [77] Y. LeCun, Y. Bengio, and G. Hinton, "Deep learning," *Nature*, vol. 521, no. 7553, p. 436, 2015.
- [78] I. B. Y. Goodfellow and A. Courville, *Deep Learning*. Cambridge, MA, USA: MIT Press, 2016.
- [79] Y. LeCun and Y. Bengio, "Convolutional networks for images, speech, and time series," *Handbook Brain Theory Neural Netw.*, vol. 3361, no. 10, pp. 1–14, 1995.
- [80] S. Hochreiter and J. Schmidhuber, "Long short-term memory," *Neural Comput.*, vol. 9, no. 8, pp. 1735–1780, 1997.
- [81] T. N. Sainath, O. Vinyals, A. Senior, and H. Sak, "Convolutional, long short-term memory, fully connected deep neural networks," in *Proc. IEEE Int. Conf. Acoust., Speech Signal Process. (ICASSP)*, Apr. 2015, pp. 4580–4584.
- [82] V. Mnih, K. Kavukcuoglu, D. Silver, A. A. Rusu, J. Veness, M. G. Bellemare, A. Graves, M. Riedmiller, A. K. Fidjeland, G. Ostrovski, S. Petersen, C. Beattie, A. Sadik, I. Antonoglou, H. King, D. Kumaran, D. Wierstra, S. Legg, and D. Hassabis, "Human-level control through deep reinforcement learning," *Nature*, vol. 518, no. 7540, pp. 529–533, Feb. 2015.
- [83] V. Mnih, K. Kavukcuoglu, D. Silver, A. Graves, I. Antonoglou, D. Wierstra, and M. Riedmiller, "Playing atari with deep reinforcement learning," 2013, *arXiv:1312.5602*. [Online]. Available: <http://arxiv.org/abs/1312.5602>
- [84] M. Hausknecht and P. Stone, "Deep recurrent Q-learning for partially observable MDPs," in *Proc. AAAI Fall Symp. Ser.*, Sep. 2015, pp. 1–9.
- [85] L. Davis, "Handbook of genetic algorithms," Tech. Rep., 1991.
- [86] D. Whitley, "A genetic algorithm tutorial," *Statist. Comput.*, vol. 4, no. 2, pp. 65–85, Jun. 1994.
- [87] T. Kohonen, "The self-organizing map," *Proc. IEEE*, vol. 78, no. 9, pp. 1464–1480, Sep. 1990.
- [88] J. Vesanto and E. Alhoniemi, "Clustering of the self-organizing map," *IEEE Trans. Neural Netw.*, vol. 11, no. 3, pp. 586–600, May 2000.
- [89] (May 2017). [Online]. Available: <https://www.robotis.com>
- [90] (Jun. 2018). [Online]. Available: <https://learn.sparkfun.com/tutorials/9dof-razor-imu-m0-hookup-guide/all/>
- [91] (Apr. 2018). [Online]. Available: <https://www.stereolabs.com/>
- [92] (Dec. 2017). [Online]. Available: <https://www.nvidia.com/tr-tr/autonomous-machines/embedded-systems/jetson-tx2/>
- [93] B. Schneider and M. Koenigs, "Human lesion studies of ventromedial prefrontal cortex," *Neuropsychologia*, vol. 107, pp. 84–93, Dec. 2017.
- [94] S. Agarwal, R. Ram, and H. Iida, "Measuring force of game information in the brain: Linking game refinement theory and neuroscience," in *Proc. 8th Int. Conf. Knowl. Smart Technol. (KST)*, Feb. 2016, pp. 291–294.
- [95] N. Bardikoff and M. Sabbagh, "The differentiation of executive functioning across development: Insights from developmental cognitive neuroscience," in *New Perspectives on Human Development*. 2017, pp. 47–66.
- [96] G. P. Band, C. Basak, H. A. Slagter, and M. W. Voss, "Effects of the game and game-like training on neurocognitive plasticity," *Frontiers Hum. Neurosci.*, vol. 10, p. 123, Mar. 2016.
- [97] S. W. Anderson, H. Damasio, R. D. Jones, and D. Tranel, "Wisconsin card sorting test performance as a measure of frontal lobe damage," *J. Clin. Experim. Neuropsychol.*, vol. 13, no. 6, pp. 909–922, Nov. 1991.
- [98] S. Teubner-Rhodes, K. I. Vaden, J. R. Dubno, and M. A. Eckert, "Cognitive persistence: Development and validation of a novel measure from the wisconsin card sorting test," *Neuropsychologia*, vol. 102, pp. 95–108, Jul. 2017.
- [99] G. Díaz-Blancat, J. García-Prieto, F. Maestú, and F. Barceló, "Fast neural dynamics of proactive cognitive control in a task-switching analogue of the wisconsin card sorting test," *Brain Topography*, vol. 31, no. 3, pp. 407–418, May 2018.

- [100] K. Tanaka, M. J. Buckley, and F. A. Mansouri, "Functional division among prefrontal cortical areas in an analog of wisconsin card sorting test," in *The Prefrontal Cortex as an Executive, Emotional, and Social Brain*. Tokyo, Japan: Springer, 2017, pp. 17–38.
- [101] J. van Benthem, *Logical Dynamics of Information and Interaction*. Cambridge, U.K.: Cambridge Univ. Press, 2011.
- [102] J. Van Benthem, E. Pacuit, and O. Roy, "Toward a theory of play: A logical perspective on games and interaction," *Games*, vol. 2, no. 1, pp. 52–86, Feb. 2011.
- [103] C. L. Nehaniv, F. Forster, J. Saunders, F. Broz, E. Antonova, H. Kose, C. Lyon, H. Lehmann, Y. Sato, and K. Dautenhahn, "Interaction and experience in enactive intelligence and humanoid robotics," in *Proc. IEEE Symp. Artif. Life (ALife)*, Apr. 2013, pp. 148–155.
- [104] (May 2017). [Online]. Available: <http://wiki.ros.org/ROS/Tutorials>
- [105] (May 2017). [Online]. Available: https://docs.opencv.org/master/d9/df8/tutorial_root.html
- [106] (May 2017). [Online]. Available: <http://www.pointclouds.org/documentation/>
- [107] (Sep. 2017). [Online]. Available: <https://www.tensorflow.org/tutorials>
- [108] Y. Mohammad and T. Nishida, *Data Mining for Social Robotics*. Cham, Switzerland: Springer, 2015.
- [109] X. Huang, W. Wu, and H. Qiao, "Computational modeling of emotion-motivated decisions for continuous control of mobile robots," *IEEE Trans. Cognit. Develop. Syst.*, early access, Jan. 1, 2020, doi: [10.1109/TCDS.2019.2963545](https://doi.org/10.1109/TCDS.2019.2963545).
- [110] R. Ramirez-Pedraza, N. Vargas, C. Sandoval, J. L. del Valle-Padilla, and F. Ramos, "A bio-inspired model of behavior considering decision-making and planning, spatial attention and basic motor commands processes," *Cognit. Syst. Res.*, vol. 59, pp. 293–303, Jan. 2020.



EVREN DAGLARLI (Member, IEEE) received the B.Sc. degree in electrical and electronics engineering from Marmara University, the M.Sc. degree in mechatronics engineering with a focus on intelligent systems and robotics from Istanbul Technical University (ITU), and the Ph.D. degree in control and automation engineering with a focus on computational cognitive-neuroscience, human–robot interaction from ITU. He has worked as a Research Assistant with the Department of Electrical and Electronics Engineering, Atilim University. He has many publications that are published in international journals, conferences/symposiums related to mechatronics, intelligent control systems, and robotic areas. As a Researcher, he took some duties and responsibilities in several national/international projects. He has worked as a Project Engineer and the Department Manager at a private sector technology and engineering company. He is currently working as a Faculty Member and a Lecturer/Instructor with the Computer Engineering Department, Faculty of Computer and Informatics Engineering, Istanbul Technical University. He is a member of ACM and ASME.

• • •