ARTICLE IN PRESS

Biologically Inspired Cognitive Architectures xxx (xxxx) xxx–xxx

Contents lists available at ScienceDirect

Biologically Inspired Cognitive Architectures

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

Research article

A computational model of cognitive development for the motor skill learning from curiosity

Hongge Ren, Chu Liu*, Tao Shi

North China University of Science and Technology, China

ARTICLE INFO

Keywords:
Curiosity
Sustained dopamine modulation
Motor skill learning
Actor-critic
Cognitive development

ABSTRACT

Curiosity is a key enabler of autonomous learning for cognitive development of humans. It allows them to learn motor skills through autonomous exploration of the environment. This paper proposes a new neurobiologically-inspired cognitive computational model: C-DCCM (Curiosity-Driven Cognitive Computing Model). The objective of this model is to implement the effect of curiosity that can be applied to the balance learning problem of the two-wheeled self-balancing robot. As basis of the modeling we use and extend the Actor-Critic model. Verification is conducted via simulation on a computing system of sustained dopamine modulation mechanism and its effects on the cerebral cortex. In order to illustrate the effect of the curiosity on the autonomous exploration, some comparison experiments about the balance learning problem have been done.

Introduction

Cognitive development is the process that humans form and develop autonomously knowledge and skills during the interaction between humans and environment. This process of humans is gradually developed and improved for their whole life. The first stage of humans' cognitive development is the acquisition of sensory motor skills Piaget and Cook (1976). However, traditional machine learning approaches to motor learning often fail to meet the requirements of autonomy and real-time. Learning and simulating the neural activities and self-regulating mechanisms of humans and animals and embedding them to robots are important research projects for cognitive robots.

In the opinion of neurophysiology, the motor skills of humans and animals are gradually formed and developed during the continuous interaction between their related organs and the environment. The studies of neurophysiology have shown that motor skills learning of humans and animals are related to the cerebral cortex and basal ganglia Doyon, Penhune, and Ungerleider (2003). The cerebral cortex is the most advanced center that regulates or controls the movement of body. There are some studies about the cerebral cortex for the motor learning of the robots in recent years. Gupta, Vig, and Noelle (2012) presented a biologically plausible computational model about prefrontal cortex for the motor skill automaticity. Srinivasa and Chelian (2012) presented a computational model of prefrontal cortex and accompanying brain systems to develop and refine reward-eliciting behaviors. Zeno (2017) presented a new neurophysiologic model about the dynamics of a rodent's navigation spatial awareness cells found in the hippocampus and entorhinal cortex. That model was used for the navigation of robots. However, they did not involve the basal ganglia.

The main function of the basal ganglia is the control of autonomous movements, the integration of fine-tuned awareness activities and motor responses. The computational model of basal ganglia mainly includes some conceptual models on the system level Joel, Niv, and Ruppin (2002) and the GPR model on the cell level Gurney, Prescott, Wickens, and Redgrave (2004). The actor-critic model is the typical representatives of the system-level model. It is widely used for the motor skill learning. Our model uses this model as the basis of modeling.

The structure of Basal Ganglia is similar to the actor-critic model. The corresponding relationship can be seen in Fig. 1. The striatum matrix corresponds to the actor part. And the striatum patch corresponds to the critic part. The dopamine signal originating from the substantia nigra pars compacta (SNc) corresponds to temporal differences (TD) prediction error signal Gillies and Arbuthnott (2015).

The role of dopamine in the basal ganglia is very important; it imparts normal reward stimuli to the motive attribute, attracting the attention and desire of humans and animals. In recent studies, if there is decay/forgetting of learned values, DA can sustain regulate. This can facilitate fast goal-reaching Kato and Morita (2016). We introduce this mechanism into the model to improve the model's learning ability. Besides, it can adjust humans' emotions. There are some models which are worth mentioning in the monoamine neurotransmitters of emotions affect. Lövheim (2012) presented the neuropsychological picture of Emotions. Where emotions are expressed in a three dimensional model

* Corresponding author.
E-mail address: liuc2240@163.com (C. Liu).

https://doi.org/10.1016/j.bica.2018.05.001
Received 4 April 2018; Received in revised form 15 May 2018; Accepted 30 May 2018

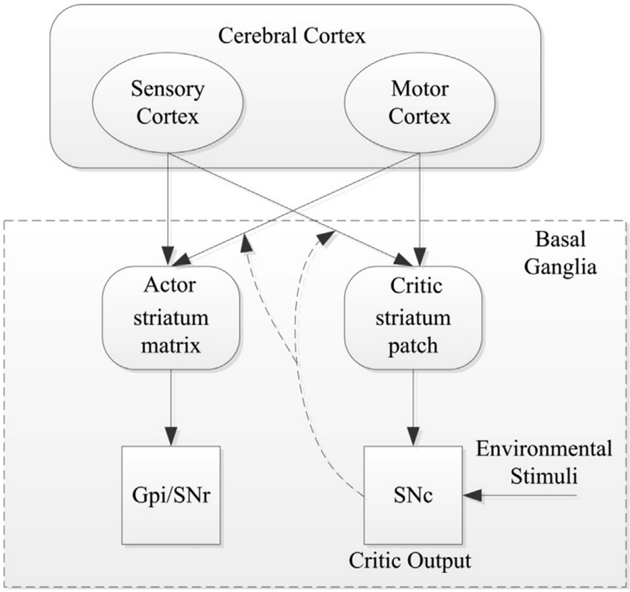

Fig. 1. The actor-critic model with possible neurobiological implementation of the basal ganglia.

of monoamine neurotransmitters: serotonin, dopamine, noradrenaline. Vallverdú et al. (2016) proposed a new neurobiologically-inspired affective cognitive architecture to implement the dopamine neuromodulation of the fear affect in the computing system. However, in these models, a clear and complete computational account of how humans generate the emotional feelings remains elusive.

The Papez dual circuit theory of the functional topology of emotion in the brain Papez (1937) suggested that the cingulate cortex generate the emotion feelings. Psychologists believe that curiosity has the characteristics of cognitive and emotion. Curiosity that acts as a positive emotion plays an important role in the process of humans' cognitive development. It can promote humans to explore the unknown environment autonomously. Making robots curious is an important research subject for cognitive robots in recent years. Oudeyer and Kaplan (2004) proposed the intelligent adaptive curiosity IAC algorithm and realized the low-level motion selection in the robot's high-dimensional perceptual motion space. After that, The team improved IAC and proposed SAGG-RIAC algorithm, the algorithm has higher exploration capabilities Baranes and Oudeyer (2010). Hester and Stone (2017) proposed a variance and novelty-intrinsic motivation algorithm that enables agents to gradually explore unknown environments and learn more complex skills in a developmental and curious manner. In the above study, the calculation of curiosity is based on the error of model. However, the exploration activities of humans belong to model-free. Zhang, Ruan, Xiao, and Zhu (2016) proposed a curiosity algorithm based on the exploration number of states. However, it is only applicable for discrete actions, not for continuous actions.

The main contribution of this paper is twofold: (i) extending the Actor-Critic model to implement the influence of curiosity on the continuous and autonomous exploration of the environment; (ii) implementing neuromodulation mechanisms of sustained dopamine for the curiosity.

The rest of this paper is organized as follows: Section *Related Work* briefly introduces the two-wheeled self-balancing robot and its previous cognitive computational model. Section *Computational model of cognitive development: C-DCCM* introduces our computational model: C-DCCM. Section *Two-wheeled robot mathematical model* introduces the mathematical model of the two-wheeled self-balancing robot. The simulation results are provided in Section *Experimental design and results analysis*, followed by the Conclusions and future work.

Related work

The two-wheeled self-balancing robot has the characteristics of non linear and under-actuated, it has attracted great attention from many researchers in recent years Chan, Stol, and Halkyard (2013). The core of the control lies in the problem of motor balance. It can also serve as an experimental platform for many control algorithms and models. The researches mostly focus on the traditional control methods such as PID control Hu, Ieee, Xu, and Zhang (2015), linear quadratic regulator (LQR) control Alarfaj and Kantor (2011) and fuzzy control Li et al. (2011). There are some cognitive models of the two-wheeled self-balancing robot in recent studies.

Chen, Xiao-Gang, and Dai (2012) proposed a behavior cognition computational model involving the coordination of cerebellum and basal ganglia. He used the RBF neural network as the actor and critic function approximation. The simulation results showed that the learning speed is increased as well as the failure times are reduced by the proposed method than by the Actor-Critic method with the only Basal Ganglia mechanism. Ruan, Chen, and Yu (2012) proposed a new tropism-based ADHDP (action-dependent heuristic dynamic programming) learning mechanism involving the cortico-basal ganglia and cerebellar circuitry and the thalamic function. The comparison experiment showed that the model implement the learning of optimal control law without knowledge of the system model, which demonstrates the model-free learning method in the brain. Zhang et al. (2016) proposed a Two-wheeled robot sensor motor system based on learning automaton, and the concepts of curiosity and orientation are introduced, The results of experiment showed that the model improves the robot's ability of self-learning and self-organizing, but also avoids the small probability event effectively, which helps keep the robot with high stability.

Computational model of cognitive development: C-DCCM

In this section, we introduce our new cognitive computational model: C-DCCM. The idea is to let the sustained dopamine neuromodulators of curiosity apply to robots to learn motor skills through the autonomous exploration of the environment. The main differences between our model and previous Actor-Critic models are twofold:(i) Dopamine neurons are not only correspond to TD prediction error signals, but also correspond to curiosity produced by the cingulate cortex; (ii) The output of striatum patch will decay over time. Electrophysiological Schultz, Dayan, and Montague (1997) and fast scan cyclic voltmeter (FSCV) Day, Roitman, and Carelli (2007) studies found that DA represents reward-prediction error. Morita and Kato (2014) have shown that reward-prediction-error can actually sustain after training if decay/forgetting of learned values.

The computational details of curiosity

For the self-balanced learning of two-wheeled robots, it belongs to the SMDP (stationary Markov decision process); In other words, the state transformation and reward function do not vary with time Ruan et al. (2012). The reward function can be regarded as the stimulus of the environment to the robot, as in formula (1).

$$r = \begin{cases} 0, \ |\theta| \leqslant \pi/10 \\ -1, \ |\theta| > \pi/10 \end{cases} \tag{1}$$

where, θ is the tilt angular of the two-wheeled robot; r is the reward value.The robot judges the current status and generates the critic value according to the stimuli given by the environment, as in formula (2).

$$J(\mathrm{t}) = r(\mathrm{t}+1) + \gamma r(\mathrm{t}+2) + \gamma^2 r(\mathrm{t}+3) + \ldots, \tag{2}$$

The critic value at time t − 1