Predictive Motor Control Based on a Generative Adversarial Network

Movitz Lenninger[1,2◑], Whan-Hyuk Choi[3,4◑], Hansol Choi[1,4*],

1 Bernstein Center Freiburg, BioIII, University of Freiburg, Freiburg, Germany
2 Division of Information Science and Engineering, KTH Royal Institure of Technology
3 Department of Mathematics, Kangwon National University, 1 Gangwondaehakgil, Chuncheon, 24341, Republic of Korea
4 Department of Biomedical Engineering, College of Information-Bio Convergence Engineering, Ulsan National Institute of Science and Technology (UNIST), Ulsan 44919, Republic of Korea

◑These authors contributed equally to this work.
* capine@gmail.com

Abstract

Predictive processing models suggest that the brain decides actions through inference. While most predictive models have been formalized using explicit representations of the probability distributions, using explicit structures and parameters, we explore an alternative representation using an implicit model known as generative adversarial networks, which has been widely explored in recent days in machine learning studies as they can learn the distribution directly from data. We demonstrate how a generative adversarial network can be trained to learn an implicit generative model of motor dynamics applicable to motor control. Furthermore, we show that such a model can perform approximate inference using the trained model, providing the necessary computations for both the forward and inverse model of motor control. We suggest that the cortical-basal ganglia circuit may behave as the generator and discriminator in the generative adversarial network model.

Introduction

Predictive processing models aim to explain action generation by the brain as an inference process [1–5]. In one example, planning-as-inference framework [3], the agent has an internal probabilistic state transition model about the state and action consequences. When making a decision, the agent uses its internal model to infer the action most likely to lead to a particular reward given its current state. Another influential example is the active inference framework, where the agent utilizes a hierarchical model of the state-action sequences. An agent decides its action by inference of a higher level latent cause (drinking coffee) from the current state (scene of a coffee cup and hand) and target states as biased prior (taking a coffee is good). The latent cause thus expands to predict the current action (moving the hand), which is realized by the body [5].

Generative models $P(x, z)$ are at the core of the predictive processing. They can be understood as "stochastic recipes for generating observed data" [6]. If x denotes observed data and z the latent variable, the system can generate new samples x by first drawing the latent state z from the prior $P(z)$ and then drawing the data from the conditional distribution $P(x|z)$. Furthermore, it is also desirable for predictive processing to be able to invert the process and infer the latent cause from observed data, i.e., given some data x infer the latent cause z from

$P(z|x)$. One important question on the generative models is how the structures of the distributions are learnt. Most of the predictive processing models focus on learning parametric models and the structure of the model is often provided by the experimenter *a priori*.

Another important point of the recent predictive processing models is that they are formalized based on explicit generative models [5–9], which require explicit parametric specification of the distribution of the observed data [10]. However, several questions need to be addressed when modelling brain functions using explicit models. For example, how the states and statistical structures among the states are learnt and represented. And also, how the entities such as probabilities and uncertainties are encoded? Although several hypotheses such as population coding explain the representation of the probabilities in brains, it is not an easy-to-answer question [11].

We suggest a predictive brain model based on an generative adversarial networks (GAN). It uses implicit generative model as an alternative formalization and does not suffer from the need to represent probabilities of samples. Instead, the target distribution $P(x|z)$ is implicitly represented by the distribution of the generated samples. Also importantly, [12–18] provide a successful framework for training such implicit generative models. Importantly, in contrast to explicit models, GANs do not require imposing any *a priori* structure of the function to train them. Thus, while explicit models require some degree of imposed structure, GANs learn the structure from the training data.

Here, we will demonstrate predictive processing with the GAN implicit generative model in a motor control problem. First, using a simple toy environment, we show how GAN can learn a generative model of a motor system dynamics. Second, we demonstrate how approximative inference using the trained GAN can behave as two of the most important models for motor control: the forward and inverse models. The forward model predicts the system's future state from the current state and action, and the inverse model decides the actions required to achieve a target state [1, 19–21]. Thus, our results show that the GAN framework can formalize the two core features of predictive processing: the structural and parametric learning of the generative model and using the learned model to perform inference. Interestingly, in the proposed framework, the architecture of the forward and inverse models is the same. We also speculate that the functional architecture of the Cortico-Basal Ganglial system can be interpreted in terms of the generator-discriminator system of our GAN predictive processing model.

Materials and methods

A Toy Environment

In the following, we consider a toy arm environment that is composed of two rods with two joints. (Fig 1**A**) Each joint has state entities of an angle and an angular velocity. The first joint connects the first rod to the ground. The second joint connects the first rod to the second rod. Thus, the state of each joint is determined by an angle and an angular velocity. A cursor is also located at the end of the second rod, indicating the end position in Cartesian coordinates. Note that although the cursor point is implicitly defined by the angles of the joints, this transformation is non-trivial. The agent can control its arm by changing the joints' angular acceleration, which we will consider as the action. The environment is deterministic and defined in discrete time. Thus, if the current states of the arm (i.e., the angles and angular velocities of the joints and their corresponding cursor position) and the current action are known, the future states of the arm at the following time point can be predicted.

The variables of our toy arm environment are as below.

- $Q = (q_1, q_2) \in \mathbb{R}^2$: The angle of the first and second joint. The first joint connects the first rod to the ground. The angle is defined over the adjacent rods.
- $\dot{Q} = (\dot{q_1}, \dot{q_2}) \in \mathbb{R}^2$: the angular velocities of the joints.

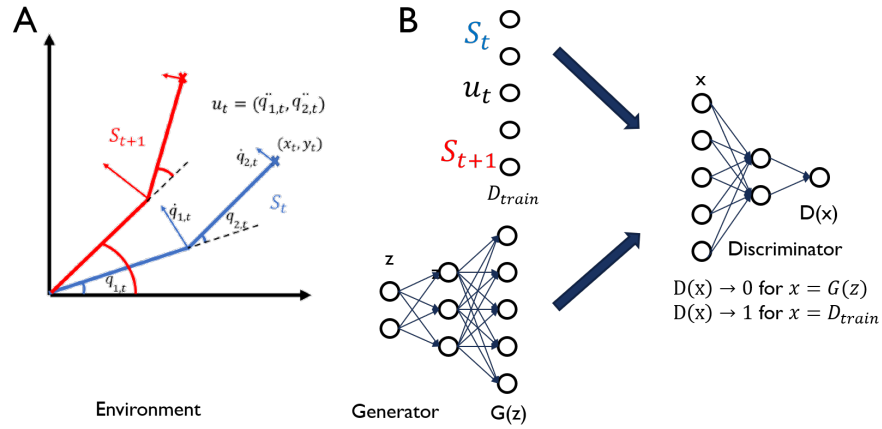

Fig 1. A toy environment and a GAN model. (**A**) A two-joint-arm system. A state S is composed of joint angles (q_1 and q_2), the angular velocities of the joints ($\dot{q}_1$ and $\dot{q}_2$, and the cursor position (x, y ; cross). The current state S_t (blue) and action U_t determines the future state S_{t+1}(red). (**B**) A generative adversarial network with a generator and discriminator. A generator is to produce samples $G(z)$ whose distribution is close to the real distribution D_{train}. This is done by co-training a discriminator, whose objective is to distinguish the real samples from the toy environment and the fake samples from the generator.

- $U = (u_1, u_2) \in \mathbb{R}^2$: u_i is the action given on the i-th joint.
- $(x, y) \in \mathbb{R}^2$: The position of the cursor at the end of the arm.
- $(Q, \dot{Q})$: The joint states, i.e., the angles and the angular velocities of the joints.
- $S = (Q, \dot{Q}, x, y)$: The state of the arm.
- S_t : The instant state of the arm at time t.
- U_t : Instant action of the arm at time t.

The dynamics of the system are as follows:

- $\dot{Q}_{t+1} = \dot{Q}_t + U_t$
- $Q_{t+1} = Q_t + \dot{Q}_{t+1}$

In the following, we will denote a State-Action-State motor sequence as

- $SUS = (q_{1,t}, q_{2,t}, \dot{q}_{1,t}, \dot{q}_{2,t}, x_{1,t}, y_{2,t}, u_{1,t}, u_{2,t}, q_{1,t+1}, q_{2,t+1}, \dot{q}_{1,t+1}, \dot{q}_{2,t+1}, x_{1,t+1}, y_{2,t+1})$.

In total, the toy environment has six degrees of freedom. If current joint states and actions are determined, all other parameters can be decided. We will indicate the relationship with the notation $a \underset{given}{|} b$ when a is fixed with given b. For example,

$$S_{t+1} \underset{given}{|} S_t, U_t$$

represents the future state, S_{t+1} of our toy environment when a set of current states and actions, S_t and U_t, are given.

Sampling training data

We uniformly sampled current joint states and the current actions. The joint angles were between $[0, \pi/2]$, angular velocities between $[-\pi/6, \pi/6]$ and the actions between $[-\pi/6, \pi/6]$. The other parameters of **SUS** were determined from the constraints imposed by the dynamics of the system. Lastly, each dimension was independently normalized between -1 and 1 by first subtracting the average of each dimension and then dividing by the maximum deviation from the average.

Training a generative model

Using samples obtained from the toy environment, $\mathbf{SUS}_{train}$, we trained a generative model using a well-established training protocol for GAN models [12–14]. In brief, two fully connected multilayer perceptrons with leaky ReLU activation functions are trained. One is the generator network **G**, and the other is the discriminator network **D** (Fig. 1**B**; 6-8-14 nodes for **G** and 14-8-1 nodes for **D**). **G** maps a random vector z, sampled from a uniform distribution $U[-1, 1]$, to the toy environment motor sequence. **D** maps an input motor sequence to an evaluation signal. The purpose of **G** is to generate realistic motor sequences, while **D** plays an adversarial role to discriminate between the motor sequence generated from **G** and the sequence sampled from the actual toy environment. **G** and **D** are trained by optimizing the loss function:

$$\mathcal{L} = max_D min_G V(G, D) = \mathbb{E}_z[log D(G(z)))] - \mathbb{E}_{x \in \mathbf{SUS}_{train}}[log D(x)] \tag{1}$$

where $\mathbf{SUS}_{train}$ is the actual sequences from the toy environment.

We validate our generative model by testing whether the generated samples reflect the physical constraints of the environment. To do so, we randomly sample 100 inputs z_{test} from the same uniform distribution. The generated samples $G(z_test)$ can be divided into three parts: the current arm state (S_t^g), the current action (u_t^g), and the future arm state (S_{t+1}^g)(black arm in Fig. 2). Ideally, the generated future state S_{t+1}^g should be close to the actual future state from the environment, given the generated current state and action. In other words, S_{t+1}^g should be close to $S_{t+1} \underset{given}{|} S_t^g, U_t^g$. Thus, we compare the actual future state of the joints and the cursor position against the generated future states (X_{t+1}^g and $Q_{t+1}^g, \dot{Q}_{t+1}^g$) in Euclidean distances.

The forward and inverse model as approximate inference

While, at first glance, the structure of GAN generative models does not allow for the inference necessary for the forward and inverse motor control problems, we here show how these problems can be treated as omitted data completion problems. The forward problem is to compute the future state of the environment given the current state S_t^f and action U_t^f:

$$S_{t+1} \underset{given}{|} S_t^f, U_t^f \tag{2}$$

where S_t^f, U_t^f are current states and actions. Our forward model approximately solves the problem as a sample completion problem. Thus, the task is to complete the **SUS**-tuple where the future state of the arm is missing (black squares below represent the partially omitted sample):

$$SUS^f = (q_{1,t}, q_{2,t}, \dot{q_{1,t}}, \dot{q_{2,t}}, x_{1,t}, y_{2,t}, u_{1,t}, u_{2,t}, \blacksquare, \blacksquare, \blacksquare, \blacksquare, \blacksquare, \blacksquare) \tag{3}$$

For the inverse problem, on the other hand, the task is to find an action, U^i, which brings the arm from the current state S_t to a desired state S_{t+1}. Reinterpreted as an omitted data competition problem, the task is to complete the following **SUS**-tuple:

$$SUS^i = (q_{1,t}, q_{2,t}, \dot{q_{1,t}}, \dot{q_{2,t}}, x_{1,t}, y_{2,t}, \blacksquare, \blacksquare, \blacksquare, \blacksquare, \blacksquare, \blacksquare, x_{1,t+1}, y_{2,t+1}) \tag{4}$$

Note that only the desired cursor position is given in the inverse problem but not the future joint states.

Yeh et al. ([14]) demonstrated how to infer the omitted part of data using GAN generative model in an image completion problem. The main idea is to find an optimal input $\hat{z}$, which minimizes a combination of perceptual loss, L_p, and contextual loss, L_c. In short, the idea is that the perceptual loss keeps the $G(\hat{z})$ evaluated as a real sample by the discriminator and that the contextual loss keeps the sample $G(\hat{z})$ close to the observable parts of the given data:

$$L_p(z) = D(G(z)) \tag{5}$$

$$L_c(z) = \|G(z) \odot M - \mathbf{SUS}_{given} \odot M\| \tag{6}$$

where $\mathbf{SUS}_{given}$ is the evidence given to the model and M is a vector that indicates the dimensions with given data. The mask for forward M_f and inverse models M_i are as below:

$$M_f = (1, 1, 1, 1, 1, 1, 1, 1, 0, 0, 0, 0, 0, 0) \tag{7}$$

and

$$M_i = (1, 1, 1, 1, 1, 1, 0, 0, 0, 0, 0, 0, 1, 1). \tag{8}$$

The total loss is a sum of the two losses, $L = L_p + \lambda L_c$ where λ is a weight parameter.

Wwen a forward or a inverse problem is givien, our system approximately searches the optimal input $\hat{z}$. The system begins from 256 different initial random inputs, z. For each sample, back-propagation is used to update the input z to minimize the loss, $L(z)$. This iteration is repeated 1,000 times, and the $\hat{z}$ with the smallest loss across the 256 updated inputs is chosen. Lastly, our system expands $\hat{z}$ to a sample $G(\hat{z})$, to find the future state for the forward model and the action for the inverse model.

We validate the forward model by testing whether the predicted future state is close to the actual future state by comparing the cursor position and the joint states (angular position and velocities) from our forward model and the ones determined by the actual toy environment. We validate the behavior of the inverse model by testing whether the action U_t^i from the inverse model directs the cursor to the target position. We computed the root mean square error between the target cursor position (x^i, y^i) and the position from the current state given the action from the inverse model $(x, y \underset{given}{|} S_t^i, U_t^i)$ and the target position (x^i, y^i).

Results

Training Generative Model for a Two-Arm System

We trained a GAN generative model on data generated from the toy environment. Fig. 2**A**-**B** shows representative samples from the naive and trained model, respectively. During the training process, we evaluated the current model by sampling 1000 random inputs $\mathbf{z}$ and producing movement samples $G(\mathbf{z})$. Comparing the joint states and cursor positions between the actual and generated future states shows that as the number of iterations increases, the model-generated samples get closer to obeying the dynamics of the actual toy model environment. Both the error between the actual and generated future joint states and the error between the actual and generated cursor position decrease during training (Fig. 2**C**-**D**). This demonstrates that the adversarial training process is capable of producing a generative model which reasonably reflects the actual dynamics of the environment solely based on training data.

The Forward and Inverse Model by Approximate Inference with GAN

To investigate whether the trained GAN model can be used as a forward model, we generated 100 $\mathbf{SUS}_{test}$ samples from the toy environment and masked the future states. Based on the

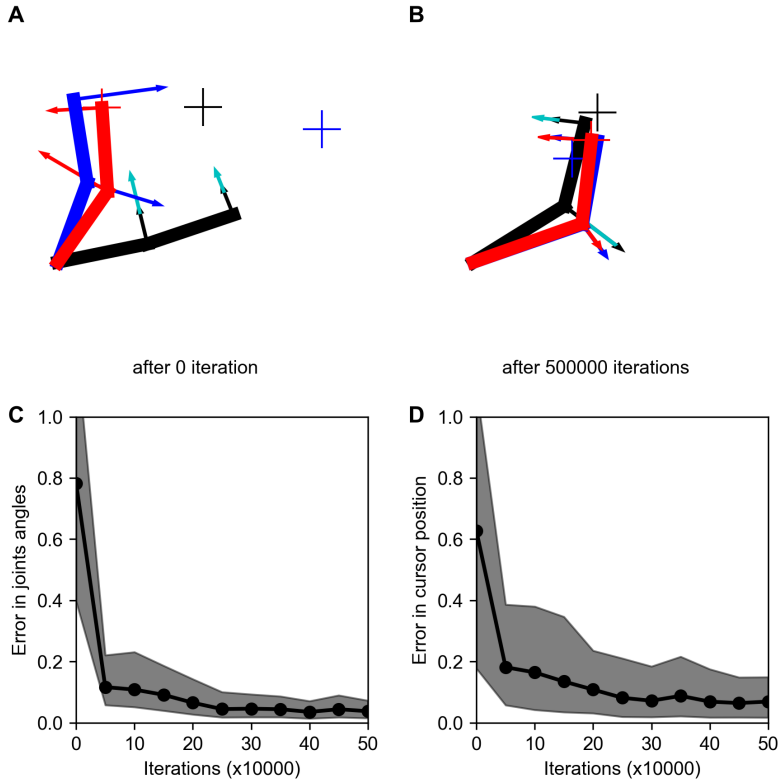

Fig 2. Training a Generative Model. Representative samples generated from (**A**) the untrained model and (**B**) the model after 500,000 training iterations. Both samples show the corresponding median error sample (in joint states) for the respective networks. **Black**: The current state generated from the generator, $S_t^g = (Q, \dot{Q}, x, y)_t^g$. The black cross is the generated cursor position, $(x, y)_t^g$. **Blue**: The future state generated from the generator, S_{t+1}^g. **Red**: The actual future state from the generated current state and action, $S_{t+1} \mid_{given} Q_t^g, \dot{Q}_t^g, U_t^g$. (**C**-**D**) The distributions of errors during training between the generated and the actual future state for (**C**) the joint state, and (**D**) the cursor position. The bold lines show the median error. The shade indicates the 5% and 95% quantiles.

knowledge of the current state and action(Fig. 3, black arms), the generative model is tasked with completing the sample for the perceptual and contextual loss. Compared to the naive model, the trained model is much better at predicting the future state of the arm (compare blue and red arms in Fig. 3**A**-**B**). The error for the future joint states between the actual and predicted future state (joint angle and angular velocities) decreases with increasing iterations (Fig. 3**C**). The same holds for the error between the predicted position of the cursors from the forward model and the actual future cursor position (Fig. 3**D**).

The same approximate inference method is used as an inverse model (Fig. 4). In the inverse model, the task is to predict the action when the agent's current state and target cursor positions are known (black arms and dots in Fig. 4, respectively). As the generative model trains, the inferred action by the inverse model becomes more accurate compared to the untrained model, that it moves the cursor closer to the target position (red crosses in Fig. 4).

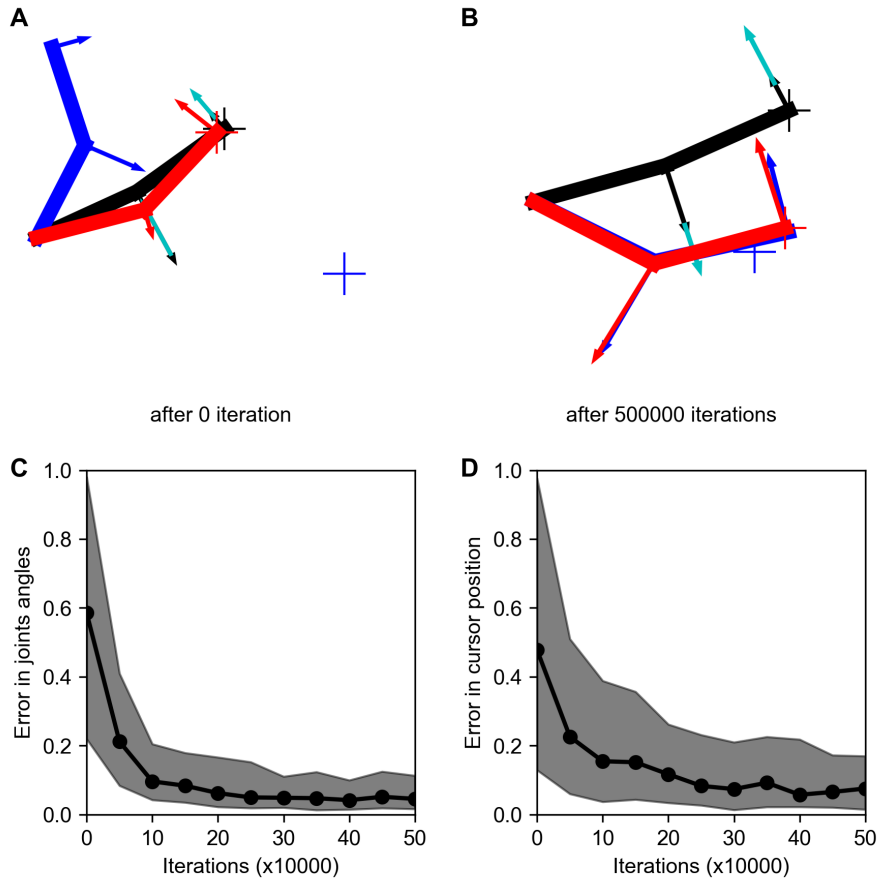

Fig 3. The forward model. Representative samples with forward model using (**A**) the untrained generative model and (**B**) the model after 500,000 training iterations. The samples with the median error in joint states are shown. Black arm: The current state of the arm. Blue: The future state predicted by our forward model, S^g_{t+1}. Blue arm: the future arm state predicted by our forward model from the current state and action, S^f_{t+1} from the forward model. Red arm: the actual future state of arm, $S_{t+1} \underset{given}{|} S^f_t, U^f_t$. (**C**-**D**) Change of the distribution of errors by training iteration between the actual and forward model predicted future state. (**C**) The errors in the joint state, and (**D**) the errors in the cursor position. The bold lines show the median error. The shade indicates the 5% and 95% quantiles.

Discussion

We have shown that GAN is capable of training a generative model for motor control sequences. The training lowered the error between the generated samples and the real constraints of the toy environment. Importantly, although minimal structural assumptions were imposed beforehand, the system could work adequately after training in the environment. The forward and inverse models of motor control were constructed as data completion processes, utilizing the trained GAN model. Interestingly, the same approximate inference architecture was used for both models. The forward and inverse models only differed in what information is considered the evidence and what information is omitted.

We hypothesize that the functional architecture of Cortico-Basal Ganglial (BG) system can be interpreted as the generator and discriminator in our GAN predictive processing architecture:

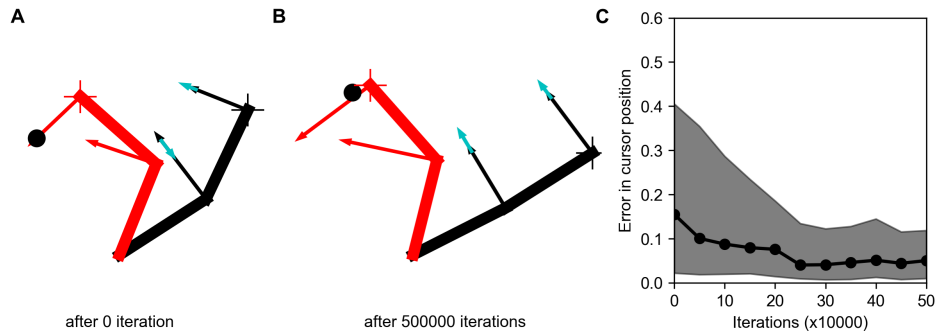

Fig 4. The inverse model. Representative arm movement using our inverse model with a generative model after (**A**) 0 iteration and (**B**) 500,000 iterations. The samples with the median error between the target cursor position and actual cursor positions are shown. Black arm: The current state of the arm. Black dot: the target cursor position in the future state. Red arm: the future arm state when the arm moved with the action from our inverse model, $S_{t+1}|S_t U_t^i$. (**C**) The error between the target cursor position and the actual cursor position. The bold lines show the median error. The shade indicates the 5% and 95% quantiles.

the neocortex as the generator and the BG as the discriminator.

Firstly, the discriminator requires (1) to receive signals from different sources of inputs to discriminate actual motion from internally generated, (2) to be capable of reducing the high-dimensional inputs to a lower-dimensional evaluation value, and (3) to send the evaluation signal to both the generator and the discriminator to teach and instruct.

The Basal Ganglia (BG) system may have those desired characteristics. (1) The dense input neurons of BG receive inputs from the diverse areas of the brain [22–25], especially from the neocortex and BG reduces the dimensions of the inputs to the activity of the smaller dimension of DAergic neurons [24, 25]. (2) We suggest that the phasic DAergic signal may act as the discriminator signal. Phasic DAergic activity has been suggested to function as a temporal difference signal (i.e., the difference between the predicted and arrived state value) for reinforcement learning [26–28]. Importantly, recent studies suggest alternative roles of DA neurons to represent non-reward information [29, 30] and state prediction errors [31–36]. Gardner et al. [37] suggested that the phasic DA signal may not only encode the reward prediction error but perhaps acts as a general prediction error signal in a diverse feature space. The prediction error can be interpreted as an evaluation signal of the generated samples, as suggested by Pfau and Vinaylas [38]. (3) The DA projects back to the basal ganglia as well as to diverse brain areas, including the neocortex (see [39] for a review). Thus, the discriminatory signal can be accessible both to the discriminator network and a generator network. These functional structures of dimensional reduction, encoding of low-dimensional evaluation signals about how realistic a sample is, and signal projections of the evaluation signals are all necessary (even though not sufficient) properties for a discriminator.

The generator requires (1) to receive evaluation input from the discriminator and to be trained from the signal, (2) to generate potential samples for the control sequences, and (3) to map abstract latent input into a specific sample. We propose that the neocortex may have the necessary properties.

(1) The neocortex receives DAergic projections from the BG, which produce synaptic triad [40] and modifies the synapse by the prediction error signal. The DAergic prediction error signal may teach the neocortex area. (2) The motor cortex activity is correlated with the current state and actions [44–48] as well as imaginary motor movement [41, 42]. The imagination of motor movement is one way of motor sample generation, which is a possible behavior by a generator. (3) The cortical area is suggested to have a functional hierarchy ranging from abstract higher causes to more concrete sensory-motor states for visual perception [49], motor

control [15, 50, 51], and decision making [52, 53].

The place for the inference process of the forward and inverse model is also needed. First of all, the generator needs to generate samples matching with the contextual cues. The affordance competition [45, 54, 55] suggests a process of contexts limit the sample space. A signal of target states can activate the neurons correlated with the potential actions to achieve the target [45]. When multiple actions are possible for the target signals, the neurons correlated with competing actions are activated together, and the neural correlates for the irrelevant actions for the given context are reduced.

Secondly, another criterion to generate actions in our model is the perceptual loss, which the discriminator evaluates the samples generated by the generator by a trained model. We hypothesize that the thalamic output from BG may have the necessary properties. The BG projects to the thalamus as one of the circuit's outputs. Interestingly, the thalamic output of BG has been suggested to be correlated with decision-making among multiple choices [31, 56–58]. This evaluation is not only limited for the value of the decision but but also related with with perceptual decision-making [59, 60]. With the function of DA as general pre diction error signal explained above, we suggest that the BG might function as general evaluator for a generated samples based on the size of reward and how probable the sample is.

We wanted to focus on the functional architecture of the system. How the evaluation signal of the generated samples can teach the generator and discriminator. Our GAN framework for modeling brain functions needs a lot of details to be explained. One of the limitations is the use of back-propagation in the learning and inference process, which is still a topic of debate and may not be feasible in real neural systems. Another algorithm, such as the recently suggested forward-forward algorithm [**?**], may provide alternative options but still need further studies.

Conclusion

We have suggested a predictive processing model for motor control based on an implicit generative model using generative adversarial networks and suggested a potential relationship with brain systems. We hope this opens for the formalization of predictive models using implicit generative models.

Acknowledgments

We thank Prof. Dr. Carsten Mehring, University of Freiburg, for the comments that greatly improved the manuscript. Whan-Hyuk Choi is supported by the National Research Foundation of Korea(NRF) grant funded by the Korea government(NRF-2022R1C1C2011689).

References

1. Pickering MJ, Clark A. Getting ahead: forward models and their place in cognitive architecture. Trends in Cognitive Sciences. 2014;18(9):451–456. doi:10.1016/j.tics.2014.05.006.

2. Clark A. Whatever next? Predictive brains, situated agents, and the future of cognitive science. Behav Brain Sci. 2012;.

3. Botvinick M, Toussaint M. Planning as inference. Trends in Cognitive Sciences. 2012;16(10):485–488. doi:10.1016/j.tics.2012.08.006.

4. Solway A, Botvinick MM. Goal-directed decision making as probabilistic inference: A computational framework and potential neural correlates. Psychological Review. 2012;119(1):120–154. doi:10.1037/a0026435.

5. Friston KJ, Daunizeau J, Kiebel SJ. Reinforcement Learning or Active Inference? PLoS ONE. 2009;4(7):e6421. doi:10.1371/journal.pone.0006421.

6. Gershman SJ. The Generative Adversarial Brain. Frontiers in Artificial Intelligence. 2019;2. doi:10.3389/frai.2019.00018.

7. Blei DM, Kucukelbir A, McAuliffe JD. Variational Inference: A Review for Statisticians. Journal of the American Statistical Association. 2017;112(518):859–877. doi:10.1080/01621459.2017.1285773.

8. Gershman SJ, Beck JM. In: Complex Probabilistic Inference. John Wiley and Sons Ltd; 2017. p. 453–466. Available from: `https://onlinelibrary.wiley.com/doi/abs/10.1002/9781119159193.ch33`.

9. Bogacz R. A tutorial on the free-energy framework for modelling perception and learning. Journal of Mathematical Psychology. 2017;76:198–211. doi:10.1016/j.jmp.2015.11.003.

10. Mohamed S, Lakshminarayanan B. Learning in Implicit Generative Models. arXiv:161003483 [cs, stat]. 2017;.

11. Pouget A, Beck JM, Ma WJ, Latham PE. Probabilistic brains: knowns and unknowns. Nature Neuroscience. 2013;16(99):1170–1178. doi:10.1038/nn.3495.

12. Goodfellow I, Pouget-Abadie J, Mirza M, Xu B, Warde-Farley D, Ozair S, et al. In: Generative Adversarial Nets. Curran Associates, Inc.; 2014. p. 2672–2680. Available from: `http://papers.nips.cc/paper/5423-generative-adversarial-nets.pdf`.

13. Arjovsky M, Chintala S, Bottou L. Wasserstein Generative Adversarial Networks. In: Proceedings of the 34th International Conference on Machine Learning. PMLR; 2017. p. 214–223. Available from: `https://proceedings.mlr.press/v70/arjovsky17a.html`.

14. Yeh R, Chen C, Lim TY, Hasegawa-Johnson M, Do MN. Semantic Image Inpainting with Perceptual and Contextual Losses. arXiv:160707539 [cs]. 2016;.

15. Dumoulin V, Belghazi I, Poole B, Mastropietro O, Lamb A, Arjovsky M, et al. Adversarially Learned Inference. arXiv:160600704 [cs, stat]. 2017;.

16. Donahue J, Krähenbühl P, Darrell T. Adversarial Feature Learning. arXiv:160509782 [cs, stat]. 2017;.

17. Radford A, Metz L, Chintala S. Unsupervised Representation Learning with Deep Convolutional Generative Adversarial Networks. arXiv:151106434 [cs]. 2015;.

18. Berthelot D, Schumm T, Metz L. BEGAN: Boundary Equilibrium Generative Adversarial Networks. arXiv:170310717 [cs, stat]. 2017;.

19. Jordan MI, Rumelhart DE. Forward Models: Supervised Learning with a Distal Teacher. Cognitive Science. 1992;16(3):307–354. doi:10.1207/s15516709cog1603_1.

20. Desmurget M, Grafton S. Forward modeling allows feedback control for fast reaching movements. Trends in Cognitive Sciences. 2000;4(11):423–431. doi:10.1016/S1364-6613(00)01537-0.

21. Franklin D, Wolpert D. Computational Mechanisms of Sensorimotor Control. Neuron. 2011;72(3):425–442. doi:10.1016/j.neuron.2011.10.006.

22. Gurney K, Prescott TJ, Redgrave P. A computational model of action selection in the basal ganglia. I. A new functional anatomy. Biological Cybernetics. 2001;84(6):401–410. doi:10.1007/PL00007984.

23. Bar-Gad I, Bergman H. Stepping out of the box: information processing in the neural networks of the basal ganglia. Current Opinion in Neurobiology. 2001;11(6):689–695. doi:10.1016/S0959-4388(01)00270-7.

24. Bar-Gad I, Morris G, Bergman H. Information processing, dimensionality reduction and reinforcement learning in the basal ganglia. Progress in Neurobiology. 2003;71(6):439–473. doi:10.1016/j.pneurobio.2003.12.001.

25. Watabe-Uchida M, Zhu L, Ogawa S, Vamanrao A, Uchida N. Whole-Brain Mapping of Direct Inputs to Midbrain Dopamine Neurons. Neuron. 2012;74(5):858–873. doi:10.1016/j.neuron.2012.03.017.

26. Schultz W, Dayan P, Montague PR. A neural substrate of prediction and reward. Science (New York, NY). 1997;275(5306):1593–9.

27. Montague PR, Hyman SE, Cohen JD. Computational roles for dopamine in behavioural control. Nature. 2004;431(7010):760–767. doi:10.1038/nature03015.

28. Houk JC, Adams JL, Barto AG. In: A model of how the basal ganglia generate and use neural signals that predict reinforcement. Computational neuroscience. The MIT Press; 1995. p. 249–270.

29. Tian J, Huang R, Cohen J, Osakada F, Kobak D, Machens C, et al. Distributed and Mixed Information in Monosynaptic Inputs to Dopamine Neurons. Neuron. 2016;91(6):1374–1389. doi:10.1016/j.neuron.2016.08.018.

30. Noudoost B, Moore T. Control of visual cortical signals by prefrontal dopamine. Nature. 2011;474(7351):372–375. doi:10.1038/nature09995.

31. Wei W, Rubin JE, Wang XJ. Role of the Indirect Pathway of the Basal Ganglia in Perceptual Decision Making. Journal of Neuroscience. 2015;35(9):4052–4064. doi:10.1523/JNEUROSCI.3611-14.2015.

32. Engelhard B, Finkelstein J, Cox J, Fleming W, Jang HJ, Ornelas S, et al. Specialized coding of sensory, motor and cognitive variables in VTA dopamine neurons. Nature. 2019;570(77627762):509–513. doi:10.1038/s41586-019-1261-9.

33. Schwartenbeck P, FitzGerald THB, Dolan R. Neural signals encoding shifts in beliefs. NeuroImage. 2016;125:578–586. doi:10.1016/j.neuroimage.2015.10.067.

34. Suarez JA, Howard JD, Schoenbaum G, Kahnt T. Sensory prediction errors in the human midbrain signal identity violations independent of perceptual distance. eLife. 2019;8:e43962. doi:10.7554/eLife.43962.

35. Takahashi YK, Batchelor HM, Liu B, Khanna A, Morales M, Schoenbaum G. Dopamine Neurons Respond to Errors in the Prediction of Sensory Features of Expected Rewards. Neuron. 2017;95(6):1395–1405.e3. doi:10.1016/j.neuron.2017.08.025.

36. Nakahara H, Itoh H, Kawagoe R, Takikawa Y, Hikosaka O. Dopamine Neurons Can Represent Context-Dependent Prediction Error. Neuron. 2004;41(2):269–280. doi:10.1016/S0896-6273(03)00869-9.

37. Gardner MPH, Schoenbaum G, Gershman SJ. Rethinking dopamine as generalized prediction error. Proceedings of the Royal Society B: Biological Sciences. 2018;285(1891):20181645. doi:10.1098/rspb.2018.1645.

38. Pfau D, Vinyals O. Connecting Generative Adversarial Networks and Actor-Critic Methods. arXiv:2212.13345 [cs, LG]. 2016;.

39. Puig MV, Rose J, Schmidt R, Freund N. Dopamine modulation of learning and memory in the prefrontal cortex: insights from studies in primates, rodents, and birds. Frontiers in Neural Circuits. 2014;8. doi:10.3389/fncir.2014.00093.

40. Yao WD, Spealman RD, Zhang J. Dopaminergic signaling in dendritic spines. Biochemical Pharmacology. 2008;75(11):2055–2069. doi:10.1016/j.bcp.2008.01.018.

41. Schnitzler A, Salenius S, Salmelin R, Jousmäki V, Hari R. Involvement of Primary Motor Cortex in Motor Imagery: A Neuromagnetic Study. NeuroImage. 1997;6(3):201–208. doi:10.1006/nimg.1997.0286.

42. Doud AJ, Lucas JP, Pisansky MT, He B. Continuous Three-Dimensional Control of a Virtual Helicopter Using a Motor Imagery Based Brain-Computer Interface. arXiv:1610.01945 doi:10.1371/journal.pone.0026322.

43. Naselaris T, Olman CA, Stansbury DE, Ugurbil K, Gallant JL. A voxel-wise encoding model for early visual areas decodes mental images of remembered scenes. NeuroImage. 2015;105:215–228. doi:10.1016/j.neuroimage.2014.10.018.

44. Fried I, Mukamel R, Kreiman G. Internally Generated Preactivation of Single Neurons in Human Medial Frontal Cortex Predicts Volition. Neuron. 2011;69(3):548–562. doi:10.1016/j.neuron.2010.11.045.

45. Cisek P, Kalaska JF. Neural Correlates of Reaching Decisions in Dorsal Premotor Cortex: Specification of Multiple Direction Choices and Final Selection of Action. Neuron. 2005;45(5):801–814. doi:10.1016/j.neuron.2005.01.027.

46. Li N, Daie K, Svoboda K, Druckmann S. Robust neuronal dynamics in premotor cortex during motor planning. Nature. 2016;532(7600):459–464. doi:10.1038/nature17643.

47. Churchland MM, Cunningham JP, Kaufman MT, Ryu SI, Shenoy KV. Cortical Preparatory Activity: Representation of Movement or First Cog in a Dynamical Machine? Neuron. 2010;68(3):387–400. doi:10.1016/j.neuron.2010.09.015.

48. Churchland MM, Cunningham JP, Kaufman MT, Foster JD, Nuyujukian P, Ryu SI, et al. Neural population dynamics during reaching. Nature. 2012;487(7405):51–56. doi:10.1038/nature11129.

49. DiCarlo J, Zoccolan D, Rust N. How Does the Brain Solve Visual Object Recognition? Neuron. 2012;73(3):415–434. doi:10.1016/j.neuron.2012.01.010.

50. DeWolf T, Eliasmith C. The neural optimal control hierarchy for motor control. Journal of Neural Engineering. 2011;8(6):065009. doi:10.1088/1741-2560/8/6/065009.

51. Tani J. Exploring Robotic Minds: Actions, Symbols, and Consciousness as Self-Organizing Dynamic Phenomena. Oxford University Press; 2016.

52. Ribas-Fernandes JF, Solway A, Diuk C, McGuire J, Barto A, Niv Y, et al. A Neural Signature of Hierarchical Reinforcement Learning. Neuron. 2011;71(2):370–379. doi:10.1016/j.neuron.2011.05.042.

53. Ribas-Fernandes JJF, Shahnazian D, Holroyd CB, Botvinick MM. Subgoal- and Goal-related Reward Prediction Errors in Medial Prefrontal Cortex. Journal of Cognitive Neuroscience. 2018;31(1):8–23. doi:10.1162/jocn_a_01341.

54. Pezzulo G, Cisek P. Navigating the Affordance Landscape: Feedback Control as a Process Model of Behavior and Cognition. Trends in Cognitive Sciences. 2016;20(6):414–424. doi:10.1016/j.tics.2016.03.013.

55. Cisek P. Cortical mechanisms of action selection: the affordance competition hypothesis. Philosophical Transactions of the Royal Society B: Biological Sciences. 2007;362(1485):1585–1599. doi:10.1098/rstb.2007.2054.

56. Dunovan K, Verstynen T. Believer-Skeptic Meets Actor-Critic: Rethinking the Role of Basal Ganglia Pathways during Decision-Making and Reinforcement Learning. Frontiers in Neuroscience. 2016;10. doi:10.3389/fnins.2016.00106.

57. Gurney KN, Humphries MD, Redgrave P. A New Framework for Cortico-Striatal Plasticity: Behavioural Theory Meets In Vitro Data at the Reinforcement-Action Interface. PLOS Biology. 2015;13(1):e1002034. doi:10.1371/journal.pbio.1002034.

58. Keuken MC, Maanen LV, Bogacz R, Schäfer A, Neumann J, Turner R, et al. The subthalamic nucleus during decision-making with multiple alternatives. Human Brain Mapping. 2015;36(10):4041–4052. doi:10.1002/hbm.22896.

59. Ding L, Gold J. The Basal Ganglia's Contributions to Perceptual Decision Making. Neuron. 2013;79(4):640–649. doi:10.1016/j.neuron.2013.07.042.

60. Lak A, Nomoto K, Keramati M, Sakagami M, Kepecs A. Midbrain Dopamine Neurons Signal Belief in Choice Accuracy during a Perceptual Decision. Current Biology. 2017;27(6):821–832. doi:10.1016/j.cub.2017.02.026.

61. Hinton G. The Forward-Forward Algorithm: Some Preliminary Investigations. 2022;(arXiv:2212.13345). doi:10.48550/arXiv.2212.13345.