Predictive Motor Control Based on a Generative Adversarial Network

Movitz Lenninger[1,2◑], Whan-Hyuk Choi[3,4,5◑], Hansol Choi[1,4,5*],

1 Bernstein Center Freiburg, BioIII, University of Freiburg, Freiburg, Germany
2 Division of Information Science and Engineering, KTH Royal Institure of Technology
3 Dept. Mathematics, Kangwon National University, Chuncheon, Republic of Korea
4 Korean Genomics Center (KOGIC), Ulsan National Institute of Science and Technology (UNIST), Ulsan 44919, Republic of KoreaDept. Bio-Medical Engineering, KOGIC, UNIST, Ulsan, Republic of Korea
5 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

We propose a predictive motor control model based on a generative adversarial network. Predictive processing models of the brain propose the forward and inverse models of motor control as an inference process. The generative models are in the core of the predictive processing models as they give the formula for sample generation and mostly formalized using explicit generative models, in which the probabilistic distributions of states are represented by explicit structures and parameters. In this paper, we examine an alternative framework, employing an implicit generative model for motor control, based on a recently discovered machine learning technique, generative adversarial network, which has been shown successful in training generative models for multiple domains of data. We illustrate how a generative adversarial network can train an implicit generative model of a motor dynamics of a toy environment, and how the trained model can be used for forward and inverse motor control models. We suggest that the cortical-basal ganglia circuit may function as the model's generator and discriminator.

Introduction

Predictive processing models explain brains' motor control as an inference process [1–5], which use the generative models $P(x, z)$, which can be viewed as stochastic "recipes" for producing observed data [6]. When x is the observed data, and z is the latent variable, the system can generate observed samples x by first drawing the latent state z from the prior $P(z)$ and then drawing the data from the conditional distribution $P(x|z)$. When data comes, the predictive processing can invert the process to infer the latent cause of the observation $P(z|x)$. In motor control [3], the probabilistic density of the actions is inferred Bayesian fashion on the evidence of the present state and the assumed evidence of goal state achievement. Active inference defines motor production as an inference of the latent cause, which can explain the current state and the values of probable future states as prior of the inference [5]. Actions anticipated from the latent causes are realized by bodies to reduce the prediction error.

Most predictive processing models are formalized on the basis of explicit generative models [5, 6, 8, 9], which give explicit parametric specification of the distribution of observed

data [10] and latent states. For explicit modeling of brain activities, several problems must be solved, including how the system represents states and associated probability and how the models learn the structure and parameters of the target environment. The majority of predictive processing models concentrate on parametric learning and inference of generative models with *a priori* structures.

In this research, we offer a predictive processing model based on an implicit generative model as an alternative formalization. Implicit models do not explicitly express the probabilities of states and samples; rather, they create samples directly and let their distribution to resemble that of the training samples. Recent advances in Generative adversarial network (GAN) [11–17] provided a successful framework for training implicit generative models and inference using them in diverse domains of data and high dimensions. GAN is particularly intriguing since it does not require *a priori* structure of functions in order to train them.

We will demonstrate predictive processing with GAN implicit generative model. First, we will demonstrate how GAN may be used to train an implicit generative model of a toy environment's dynamics. Second, we will explain how approximation inference with a trained GAN can function as two crucial models for motor control: the forward model and the inverse model. The forward model predicts the future state based on the current state and action, whereas the inverse model determines the actions necessary to reach a desired state [1, 18–20]. We demonstrate that the GAN framework can provide a new idea for predictive processing: structural and parametric learning of the generative model and model inference. Additionally, we will explain the likely connection between the functional architecture of the Cortico-Basal Ganglia system and the generator-discriminator system of the GAN framework.

Materials and methods

A Toy Environment

Our toy arm environment 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. A cursor is at the end of the second rod. The agent can control its arm by changing the joints' angular acceleration. The environment is deterministic and defined in discrete time. If the current joint's states (angles and angular velocities with their corresponding cursor position) and the current action are determined, the future joint states at the following time point are determined. (Fig. 1 **B**, up left, D_{train}).

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.
- $U = (u_1, u_2) \in \mathbb{R}^2$. 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 joints states, position and angular velocity of the joints.
- $S = (Q, \dot{Q}, x, y)$:The state of an 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. The future state

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

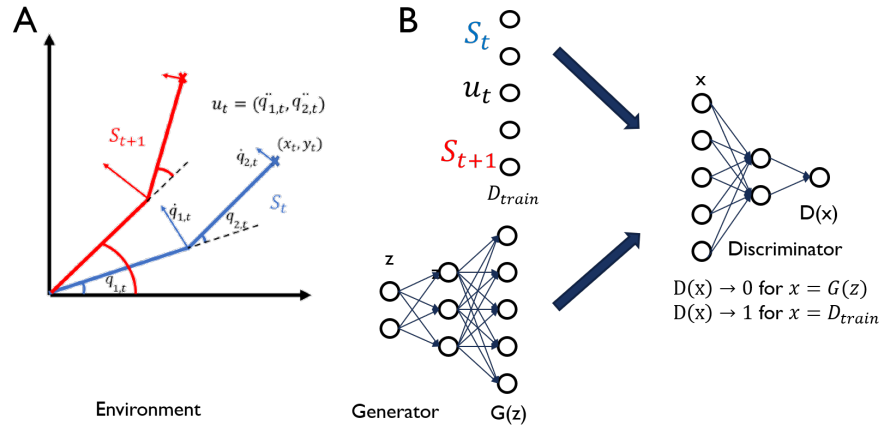

Fig 1. A toy environment and a GAN model. A A two-joint-arm system. A state S is composed of joint angles (q_1 and q_2), the angular velocities of the joints ($\dot{q}_1$ and $\dot{q}_2$, and the cursor position (x, y ; cross). The current state S_t (blue) and action U_t determines the future state S_{t+1}(red). **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.

- $Q_{t+1} = Q_t + \dot{Q}_{t+1}$
- $SUS = (q_{1,t}, q_{2,t}, \dot{q}_{1,t}, \dot{q}_{2,t}, x_{1,t}, y_{2,t}, u_{1,t}, u_{2,t}, q_{1,t+1}, q_{2,t+1}, \dot{q}_{1,t+1}, \dot{q}_{2,t+1}, x_{1,t+1}, y_{2,t+1})$.: A State-Action-State motor sequence.

Our toy environment has six degrees of freedom. When current joint states and actions are determined, current cursor position and future joint states and cursor positions are 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 action, $S_t and U_t$, are given.

Sampling training data

We uniformly sampled current joint states and the current actions, U_t. 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. Each dimension was normalized independently from -1 to 1: subtracted by the average of each dimension, and then divided by the maximum deviation from the average.

Training a generative model

A GAN model was trained for samples from our toy environment as previously reported [11, 13] (Fig 1B). In brief, We train two fully connected multilayer perceptron with leaky ReLU as non-linearity activation functions, a generator **G** and a discriminator 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 our toy environment motor sequence. D maps an input motor sequence to a 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 our actual toy environment

G and D are trained by optimizing the loss function (1):

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

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

We validated our generative model by testing whether the generated samples reflected the physical constraints of the environment. We randomly sampled 100 input z_{test} from the uniform distribution. The generated samples $G(z)$ can be divided into three parts: the current arm state, the current action, and the future arm state: S_t^g, U_t^g and S_{t+1}^g correspondingly (black arm in Fig. 2).

The future state S_{t+1}^g from a sample $\mathbf{SUS}^g$ generated from the trained generator should be close to the actual future state from the environment when the current state and action are as generated, $X_{t+1}|S_t^g, U_t^g$ and $Q_{t+1}, \dot{Q}_{t+1}|S_t^g, U_t^g$ (| indicates the left part is determined from the right part under the physical constraints of the toy environment. g indicates the value is from the generator). Each of which was compared 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

Our forward and inverse models of motor control use the GAN generative model trained for the toy environment. Our forward model estimates The main idea is to concern the motor control problems as an omitted data completion problem. The forward problem is to compute a future state of the environment,

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

where S_t^f, U_t^f are current states and actions which is given to the forward model. Our forward model approximately solves the problem as a sample completion problem. We concern we have a partially omitted sample SUS^f from our toy environment as below. (The black square below represents the partially omitted sample).

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

The inverse problem is to find a U^i which can achieve a goal state S_{t+1}^i from the current state current state S_t^i. Our inverse model also concerns the problem as a sample completion problem :

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

Yeh et al. ([13]) demonstrated how to infer the omitted part of data using GAN generative model in images completion problem. The main idea is to find an optimal input $\hat{z}$, which minimizes perceptual loss L_p and the contextual loss L_c. The perceptual loss keeps the $G(\hat{z})$ evaluated as a real sample by the discriminator. The contextual loss keeps the sample $G(\hat{z})$ to be close to the given evidence.

$$L_p(z) = D(G(z))$$

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

$\mathbf{SUS}_{given}$ is the evidence given to the model. 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.

$$M_f = (1, 1, 1, 1, 1, 1, 1, 1, 0, 0, 0, 0, 0, 0)$$

$$M_i = (1, 1, 1, 1, 1, 1, 0, 0, 0, 0, 0, 0, 1, 1)$$

The total loss is $L = L_p + \lambda L_c$ where λ is a weight parameter.

We started with 256 different random inputs, zs . Back propagation was used to determine the gradient that should be used to update each of z to reduce its $L(z)$. The iteration was repeated 1,000 times. The $\hat{z}$ with the smallest loss was chosen. $G(\hat{z})$ was generated to find the future state for the forward model and the action for the inverse model.

We validated the forward model by testing whether the predicted future state is close to the actual future state. We compared the cursor position and the joint state(angular position and velocities) from our forward model and the ones determined by actual toy environment. We validated 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 $x, y|S_t^i, U_t^i$.

Results

Training Generative Model for a Two-Arm System

Trained GAN model for the samples from our toy environment could produce samples which are close to the real ones. As the number of iterations increases, our generative model generated samples close to the actual toy model environment. The error between the actual and generated future joint states decreases (Fig.2 (AC)), and the error between the actual and generated cursor position also decreased (Fig. 2 (BD)). The generative adversarial process could let our generative model produce samples reflecting the dynamics of the environment. The generator received 1000 random inputs $\mathbf{z}$ and produced the samples, $G(\mathbf{z})$s. We compared the joints states and the cursor positions between the actual and generated future joint states. Fig. 2 shows the representative samples from the naive and trained model with median error on the cursor position. The training let the actual and generated future state closer compared to the Naive model.

The Forward and Inverse Model by Approximate Inference with GAN

The trained GAN model could be used as a forward model. We gave 100 test samples with the current state and action to the forward model (Fig.3, black arms), and it predicted the future state. The predicted future state from the forward model (Fig.3, blue arms) becomes close to the actual future state from the environment (Fig.3, red arms). The error for the future joint states between the actual and predicted future state (joint angle and angular velocities) decreased as the iteration number increased (Fig. 3C). The error between the predicted position of the cursors (Fig. 3, blue crosses; Fig3 D) from the forward model and actual future cursor position (Fig. 3, red cross) also decreased.

The same approximate inference method was used as an inverse model (Fig. 4). The inference was to predict the action when the agent's current state (black arms in Fig. 4) and target cursor positions (black dot in Fig. 4) are given. As the generative model trained, the action from the inverse model moved the cursor close to the target (red crosses in Fig. 4) compared to the untrained model or the random movement of the arm.

Discussion

We showed that GAN is capable of training a generative model for motor control sequences. The training lowered the error of the generated samples compared to the real constraints of the toy environment. The forward and inverse models of motor control were constructed as the data completion process utilizing the trained GAN model. Same approximate inference architecture were used for the both of the model. Only the evidence and target dimension were different each

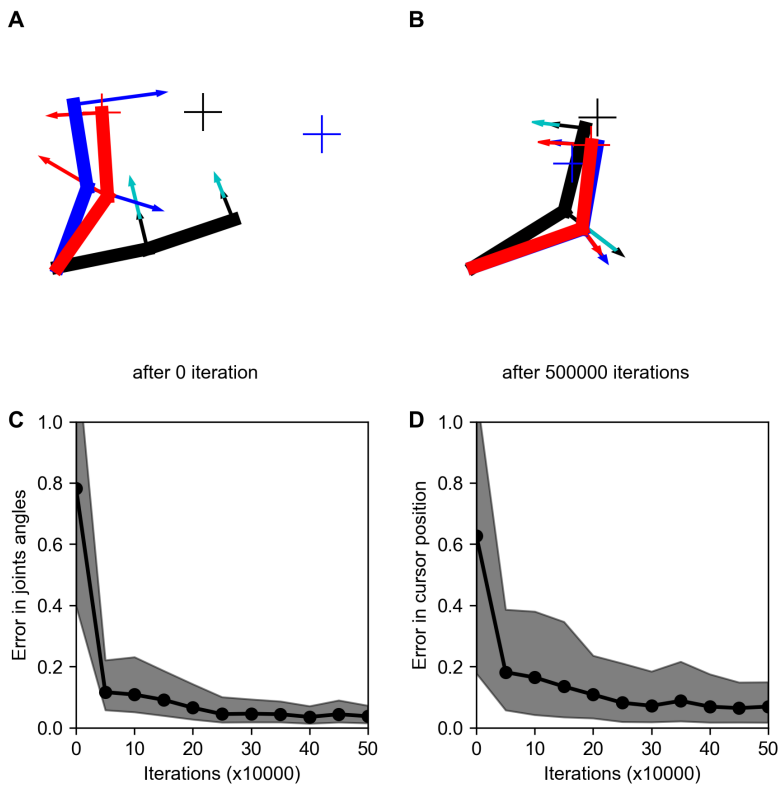

Fig 2. Training a Generative Model. Representative samples generated from **A** the untrained model and **B** the model after 500,000 training iterations. The samples with the median error in joint states are shown. **Black**: The current state generated from the generator, $S_t^g = (Q, \dot{Q}, x, y)_t^g$. 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, $S_{t+1} \underset{given}{|} Q_t^g, \dot{Q}_t^g, U_t^g$. **CD** Change of the distribution of errors by training iteration between the generated and the actual future state by iteration of the difference in **C** the joint state and **D** the cursor position. The bold lines show the median error. The shade indicates the 5% and 95% quantiles.

other. The system accepts currently determined conditions (current states and actions in the forward model and current states and target states in the inverse model) as evidence and is able to infer the missing domain of knowledge.

We hypothesize that the functional architecture of Cortico-Basal Ganglial system of the brains can be interpreted as the the generator and discriminator of our GAN predictive processing architecture. First, the discriminators (1) must receive signals from distinct sources of inputs. (2) A computational function is necessary to reduce high-dimensional inputs to a lower-dimensional evaluation value. (3) The evaluation signal must instruct both the generator and the discriminator.

Basal Ganglia (BG) system may have those characteristics. (1) The dense input neurons of BG receive inputs from the diverse area of the brains [21–24], especially from the cortical area. BG reduces the dimensions of the inputs to the activity of the smaller number of DAergic neurons [23, 24]. (2) The function of phasic dopaminergic activity has been suggested as the temporal difference signals [25]: the difference between the predicted and the arrived value of a state, which drives reinforcement learning in the brains. Importantly, even though DA signal is widely known to represent the reward prediction error [25–27], recent studies report the

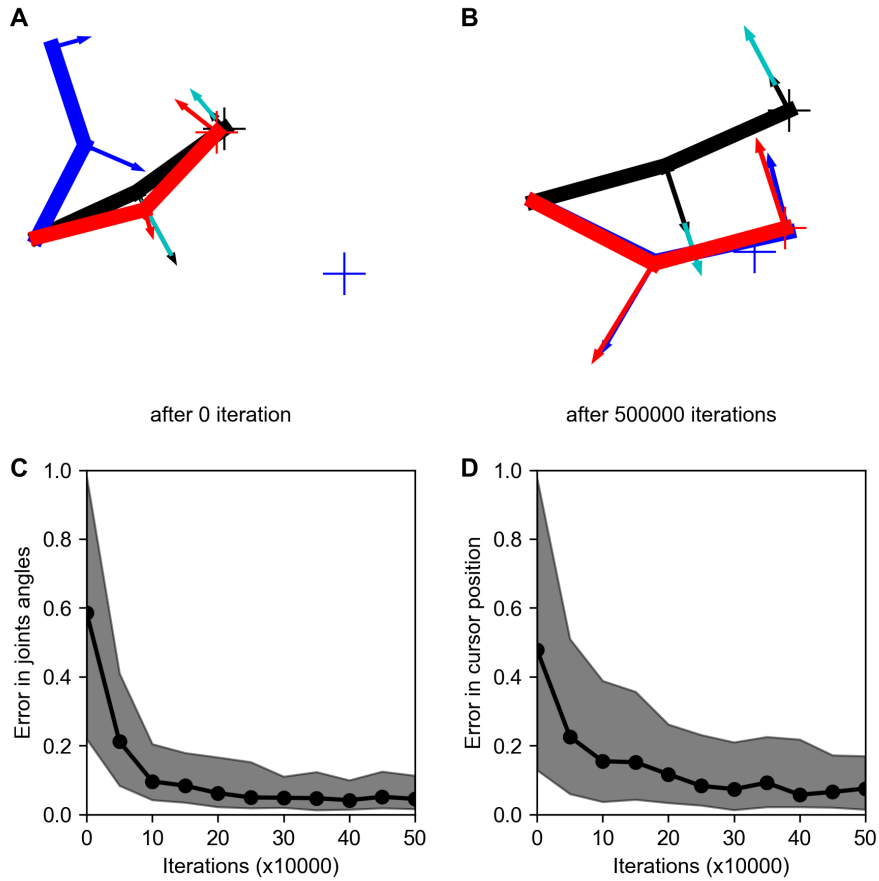

Fig 3. The forward model. Representative samples with forward model using **A** the untrained generative model and **B** the model after 500,000 training iterations.The samples with the median error in joint states are shown. Black arm: The current state of the arm. Blue: The future state predicted by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

functional role of DA neurons to represent the non-reward information [28, 29] and the state prediction errors [30–35]. Gardner et al. [36] suggested that the phasic DA signal may not only encode the reward prediction error but function as a general prediction error in diverse feature spaces. The DAergic state prediction error may behave as the evaluation signal from the discriminator because it encodes how a generated activity is close to the real states. (3) The DA signal projects back to the basal ganglia itself and also send projection to a diverse area, including the cortices ([37] for review). These functional structures of dimension reduction, the encoding of smaller dimensional evaluation signals about how realistic a sample is, and trained from the evaluation signals are the necessary (even though not sufficient) property for the discriminators.

The generator requires (1) to receive evaluation input from the discriminator and be trained from the signal. (2) should be able to generate samples that closely resemble the actual

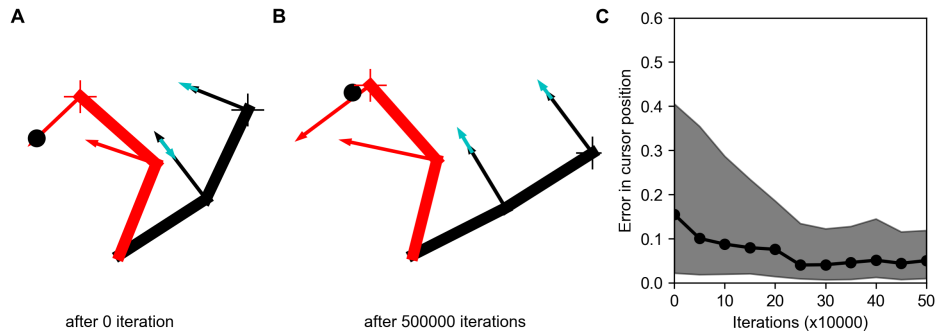

Fig 4. The inverse model. Representative arm movement using our inverse model with a generative model after **A** 0 iteration and **B** 500,000 iterations. The samples with the median error between the target cursor position and actual cursor positions are shown. Black arm: The current state of the arm. 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** Change of the error between the target cursor position and the actual cursor position. **C** the difference in the cursor position. The bold lines show the median error. The shade indicates the 5% and 95% quantiles.

distribution and (3) represent the states of controls. (4) Mapping the latent input causes to a particular sample. We propose that the cortical region may possess the required features.

The cortical area has (1) projection from the BG, which produce synaptic triad [38] and modifies the synapse. (2) The motor cortex activities are found to be correlated with the imaginary motor movement [39]. The motor imagery can be decoded from the signal [40]. Another indirect evidence is from the studies about visual perception. The visual cortical was suggested as the place of the generative models in the spatio-temporal sensory-motor representation of visual processing [6, 41–43] and active during visual imagination, which is to produce samples over the scene [44]. (3) The cortical area is involved in the representation of the current and future states and self-generated imagery. The motor area activities are correlated with the current state and actions during the planning [45–48], and action [49]. (4) The cortical area is suggested to have a functional hierarchy from abstract higher causes to the more concrete sensory-motor states for visual perception [50], the motor control [14, 51, 52], and decision making [53, 54].

The place for the inference process of the forward and inverse model is needed. The generator needs to generate samples with perceptual loss. The affordance competition theory of cortex [46, 55, 56] suggests that the conditional cues can activate relevant actions [46], which is required function for computing contextual loss. The second criterion is perceptual loss computed by the discriminator. Secondly, the discriminator needs to encode the perceptual loss and the Thalamic output from basal ganglia had been suggested to encode decision among multiple choices [30, 57–59] and perceptual decision-making [60, 61]. This shows that the cortical area and BG has the characteristics required for the inference process.

Conclusion

We suggested predictive processing for motor control based on an implicit generative model using GAN and suggested the potential relationship with brain systems. This opens the formalization of the predictive model with implicit generative models.

Acknowledgments

We thank Prof. Dr. Carsten Mehring, University of Freiburg, for the comments that greatly improved the manuscript.

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. 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`.
12. Arjovsky M, Chintala S, Bottou L. Wasserstein GAN. arXiv:170107875 [cs, stat]. 2017;.
13. Yeh R, Chen C, Lim TY, Hasegawa-Johnson M, Do MN. Semantic Image Inpainting with Perceptual and Contextual Losses. arXiv:160707539 [cs]. 2016;.
14. Dumoulin V, Belghazi I, Poole B, Mastropietro O, Lamb A, Arjovsky M, et al. Adversarially Learned Inference. arXiv:160600704 [cs, stat]. 2017;.
15. Donahue J, Krähenbühl P, Darrell T. Adversarial Feature Learning. arXiv:160509782 [cs, stat]. 2017;.

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

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

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

19. 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.

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

21. 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.

22. 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.

23. 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.

24. 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.

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

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

27. 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.

28. 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.

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

30. 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.

31. 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.

32. 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.

33. 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.

34. 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.

35. 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.

36. 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.

37. 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.

38. 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.

39. 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.

40. Doud AJ, Lucas JP, Pisansky MT, He B. Continuous Three-Dimensional Control of a Virtual Helicopter Using a Motor Imagery Based Brain-Computer Interface. PLoS ONE. 2011;6(10). doi:10.1371/journal.pone.0026322.

41. Yuille A, Kersten D. Vision as Bayesian inference: analysis by synthesis? Trends in Cognitive Sciences. 2006;10(7):301–308. doi:10.1016/j.tics.2006.05.002.

42. Kersten D, Mamassian P, Yuille A. Object Perception as Bayesian Inference. Annual Review of Psychology. 2004;55(1):271–304. doi:10.1146/annurev.psych.55.090902.142005.

43. Friston K, Adams RA, Perrinet L, Breakspear M. Perceptions as hypotheses: saccades as experiments. Frontiers in Perception Science. 2012;3:151. doi:10.3389/fpsyg.2012.00151.

44. 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.

45. 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.

46. 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.

47. 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.

48. 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.

49. 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.

50. 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.

51. 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.

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

53. 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.

54. 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.

55. 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.

56. 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.

57. 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.

58. 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.

59. 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.

60. 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.

61. 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.