Title	Emergence of Reaching using Predictive Learning as Sensorimotor Development in Complex Dynamics
Author(s)	Kim, Dongmin; Kanazawa, Hoshinori; Kuniyoshi, Yasuo
Citation	The 11th International Symposium on Adaptive Motion of Animals and Machines (AMAM2023). 2023, p. 144-145
Version Type	VoR
URL	https://doi.org/10.18910/92309
rights	
Note	

Osaka University Knowledge Archive : OUKA

https://ir.library.osaka-u.ac.jp/

Osaka University

Emergence of Reaching using Predictive Learning as Sensorimotor Development in Complex Dynamics

Dongmin Kim[1], Hoshinori Kanazawa[1], Yasuo Kuniyoshi[1]
[1]Graduate School of Information Science and Technology, The University of Tokyo, Japan
{*d-kim, kanazawa, kuniyosh*}*@isi.imi.i.u-tokyo.ac.jp*

1 Introduction

Reaching is a fundamental skill that is essential for both animals and robots. In animals, it enables interaction with the environment and performing tasks such as obtaining food and defending themselves. For robots, reaching also provides the ability to interact with the physical world. Unlike animals, traditional robots require training with explicit rewards to develop reaching. In contrast, animals can acquire the skill of reaching through sensorimotor experience and self-exploration in early developmental process.

Theories suggest that the Central Nervous System (CNS) has an internal predictive model for motor planning or control [1]. This hypothesis was later supported by the discovery that the interaction between a forward and inverse predictive model enables internal feedback control, leading to the emergence of reaching [2]. The forward model predicts the next sensory state based on the current sensory state and motor command, while the inverse model predicts the current motor command based on the current and next sensory states. Previous research simulated the emergence of reaching with a sensorimotor predictive model through random movements [3]. However, this research was limited to a simple physics space and a joint angle-based control model.

To evaluate the hypothesis that reaching can emerge through predictive learning in environments with more complex dynamics, we conducted an experiment using a 2-link arm model with muscle actuators that result in non-linear dynamics in the MuJoCo physics simulator (Figure 1a).

2 Methods

2.1 Forward and inverse predictive model

We implemented two sensorimotor predictors, the forward and inverse model (Figure 1b). The hand position (end-effector position of the 2-link arm) was used as the sensory state, and muscle activation as the motor command.

The forward model encodes the current sensory state v_t and motor command a_t into internally represented data h_{t+1} using an encoder (linear layer) and a Long Short-Term Memory (LSTM) layer. The predicted next sensory state v'_{t+1} is then decoded by the decoder (linear layer). The inverse model encodes the next sensory state v_{t+1} and combines it with the internally represented sensory state h_t. The predicted current motor command a'_t is then decoded by the decoder (linear layer). The predicted sensory state and motor command were compared with the actual data in the simulator and updated to minimize the prediction error.

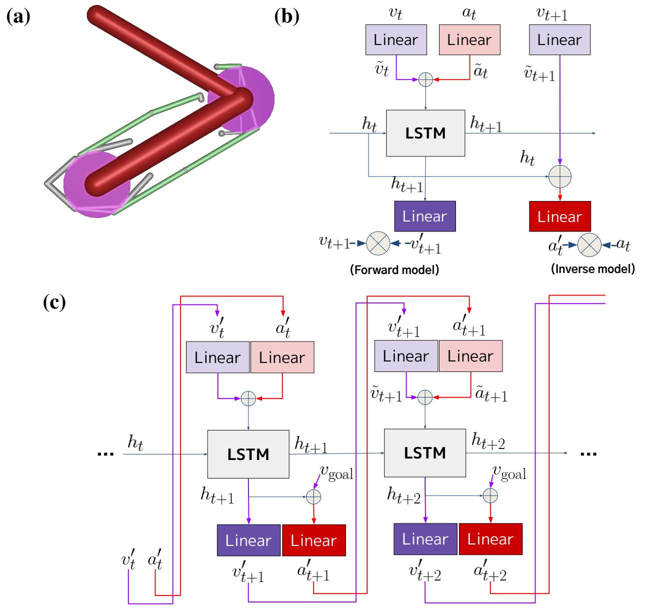

Figure 1: Sensorimotor predictive model for emergence of reaching. (a) 2-link arm model with muscle actuator. (b) Forward and inverse predictors. (c) Loop structure to emerge reaching.

2.2 Emergence of reaching

The loop of the forward and inverse predictors was used to generate a sequence of motor commands from the target sensory state (Figure 1c). This loop emerged reaching without directly learning it, but rather by predicting the trajectory.

The LSTM was initialized with the initial sensory state and null motor command for 20 iterations. The reaching goal state v_{goal} was then fed to the inverse predictor, which predicted the motor command a'_{t+1} to move towards the goal. The forward predictor was updated with the previous predicted sensorimotor state v'_{t+1} and motor command a'_{t+1}, generating the next internally represented sensory state h_{t+2}. This process was repeated to generate a longer trajectory.

2.3 Motor babbling with 2-link muscle arm model

To test reaching in a complex environment, we used a 2-link muscle arm model in the MuJoCo physics simulator. We scaled up the original MuJoCo arm26 model to the size used in previous research [3]. Motor babbling was performed by randomly generating motor commands. First, a base motor command $\tilde{a}$ in the range of 0 to 1 was randomly generated for muscles in each episode. Then, the command was made to vary over time by applying a cosine wave using Eq. (1):

$$a = \frac{1}{2}[\cos(p_1 + p_2(\frac{2\pi}{n})t) + 1]\tilde{a} \quad (1)$$

where p_1 and p_2 are random variables generated uniformly in the range of 0 to 1, t is current step, and n is total step.

3 Experiments and Results

We collected 12,000 episodes of sensorimotor data, each episode simulating 5,000 steps of motor babbling with a timestep of 0.001 seconds. The training set ratio was 0.95, and the kinematics, including arm length, range of motion, and degree of freedom, were kept the same as in [3].

The neural model was updated to use LSTM instead of Recurrent Neural Network (RNN) and a deeper multi-layer perceptron (MLP) was used, compared to [3]. The LSTM layer had 200 nodes, and the encoder and decoder each had 2 linear layers, each with 200 nodes and a Rectified Linear Unit (ReLU) function between the layers. A sigmoid function was used as the activation function for each decoder. The entire model was trained for 100 epochs with a batch size of 64, using the ADAM optimizer with a learning rate of 0.001.

3.1 Forward and inverse prediction

The forward model predicted the next hand position v'_{t+1} based on the current hand position v_t and muscle activation a_t. The result showed that the predicted v'_{t+1} was approximately same as the actual v_{t+1} (mean L2 norm between v_{t+1} and v'_{t+1} was 0.024 ± 0.015 m; Figure 2).

The inverse model was also tested, which predicted the current muscle activation a'_t based on the current and next hand positions v_t, v_{t+1}. The small time step of the simulation compared to previous research [3] made it difficult to visualize the results (mean L2 norm between v_{t+1} and $\hat{v}_{t+1} \leftarrow (v_t, a'_t)$ was $6.867 * 10^{-6} \pm 9.818 * 10^{-6}$ m). Instead, we calculated the L2 norm between a_t and a'_t and found it to be 0.74 ± 0.31. With the model having 6 muscles and activation range from 0 to 1, the maximum L2 norm is 2.45, so the inverse model was considered to have approximately learned the inverse dynamics.

3.2 Reaching and pointing

We tested the reaching and pointing by looping the forward and inverse models. The results showed that the agent was able to approach the target in the correct direction (Figure 3a). It was important to note that the agent generated the trajectory using only its predictive models after the initial hand position and target position were given, without access to the actual hand position during the reaching.

We also tested a pointing action where the target was placed at an unreachable position (Figure 3b). The agent tended to point towards the target, consistent with previous research [3].

4 Conclusion and Future Work

This paper investigated the use of predictive learning for developing reaching in complex environments with non-linear actuator dynamics. Our 2-link arm model with a muscle actuator successfully reached targets and pointed towards unreachable ones through forward and inverse predictive models. The

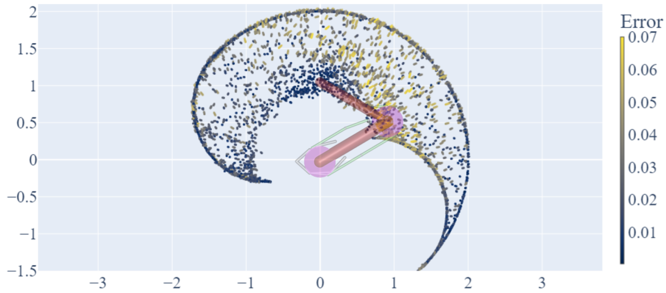

Figure 2: Forward prediction test. Drawn random sampled 5,000 steps from test set. Each line connects the next hand position and the predicted hand position.

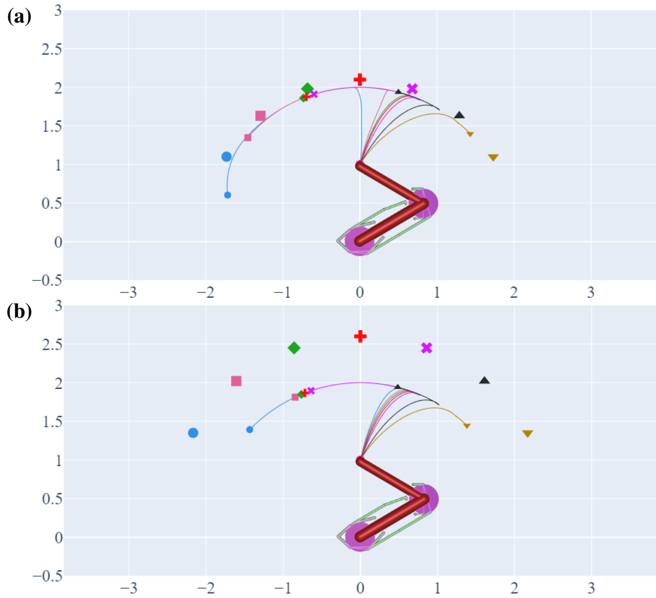

Figure 3: Trajectory of reaching and pointing. (a) Reaching to the targets placed at a reachable distance. (b) Pointing towards targets with far distance.

results indicate the potential of predictive learning for reaching in complex settings and its applications in developing adaptive, interactive robots.

Potential future work could involve exploring additional models, such as infant models [4], to better understand reaching development. Examining the impact of simultaneous body development, e.g., changes in arm length, on learning could also offer valuable insights into reaching behavior and adaptive robotic systems.

References

[1] D. M. Wolpert, Z. Ghahramani, and M. I. Jordan, "An internal model for sensorimotor integration," *Science*, Vol.269, No. 5232, pp.1880–1882, 1995.

[2] M. Desmurget and S. Grafton, "Forward modeling allows feedback control for fast reaching movements," *Trends in cognitive sciences*, Vol.4, No. 11, pp.423–431, 2000.

[3] N. Takemura, T. Inui, and T. Fukui, "A neural network model for development of reaching and pointing based on the interaction of forward and inverse transformations," *Developmental science*, Vol.21, No. 3, p.e12565, 2018.

[4] D. Kim, H. Kanazawa, and Y. Kuniyoshi, "Simulating a human fetus in soft uterus," proceeding in *2022 IEEE International Conference on Development and Learning (ICDL)*, pp.135–141, 2022.