論文の内容の要旨

Disentangling animal behavior with probabilistic generative models
（確率的生成モデルを用いた動物行動の解析）

森　啓太

Introduction

Animals exhibit a variety of behaviors that are adapted to their environment. Behavior is the final output of the nervous system and is a major factor that determines the survival of an individual. Understanding the generation and control processes of such diverse and important behaviors is a major goal not only for ethology, but also for neuroscience, ecology, evolution, and informatics, and it is a problem that has not yet been solved. The main goal of this research is to elucidate and reproduce the system that can generate and control various behaviors seen in nature.

Acquisition of the behavior of freely moving *Caenorhabditis elegans*

In this dissertation, the behavioral data of nematode *Caenorhabditis elegans* (*C. elegans*) was mainly used to confirm the validity of the method. *C. elegans* is a suitable experimental animal to test the effectiveness of this study because it shows stochastic and multiple behavioral patterns while the behavior is easy to measure. First, a behavioral data set was obtained to test the developed method, and the behavioral state was quantified by recording *C. elegans* freely moving in a two-dimensional plane. In addition, to investigate whether the stochastic behavioral responses to sensory stimuli can be appropriately modeled, I expressed channelrhodopsin (ChR2) in ASH neurons, which are sensory neurons to nociceptive stimuli, and activated them by randomly exposing them to blue light during behavioral recordings. From the acquired images, I quantified posture and speed as indicators of behavior using image processing.

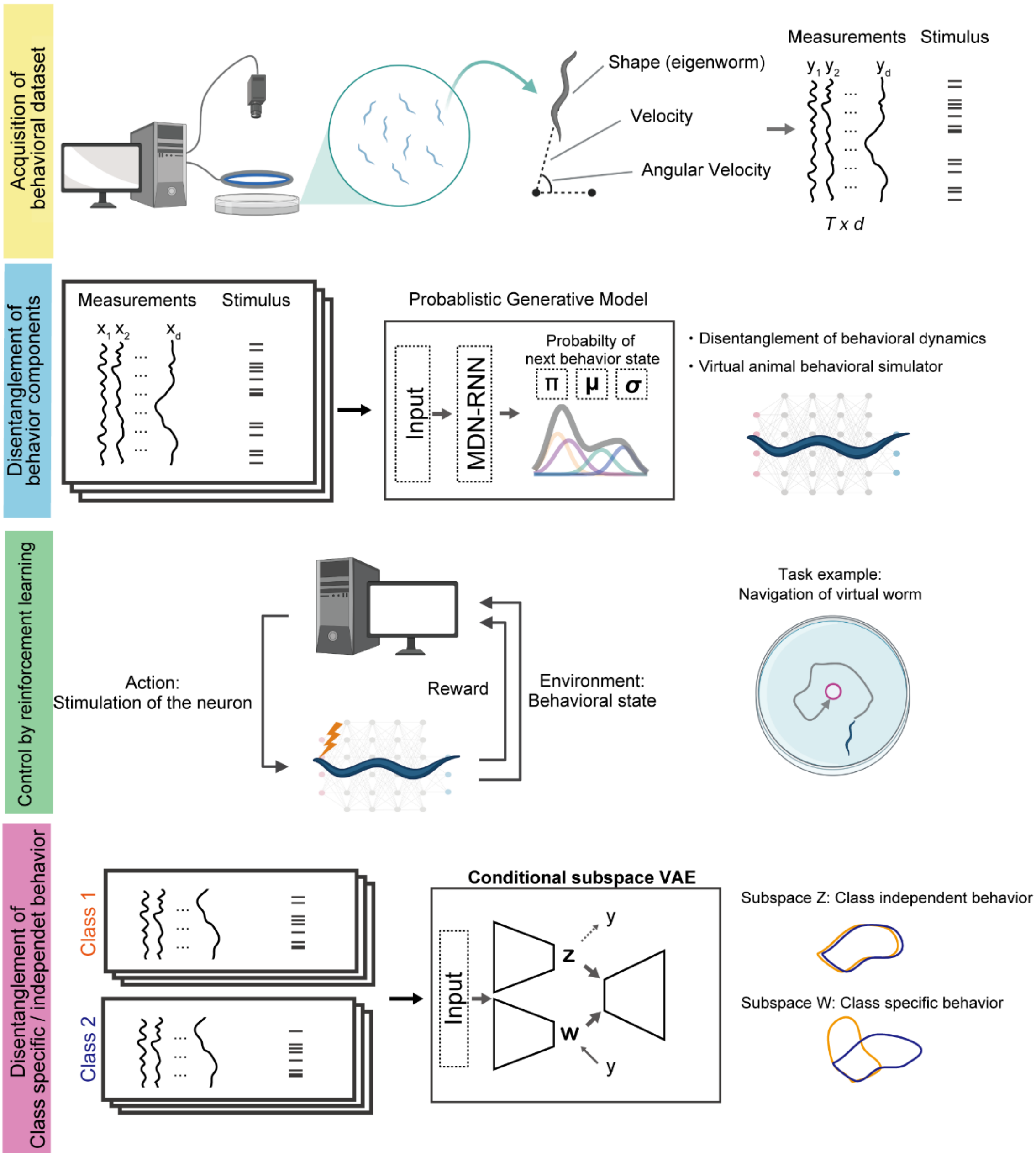

Figure Schematic illustration of this study

Chapter 1 Probabilistic generative neural networks disentangle dynamics of animal behavior

In Chapter 1, I aimed to develop a virtual animal model which can both reproduce stochastic animal behavior and represent various behaviors in a disentangled latent space. In order to model a phenomenon that takes multiple states and is stochastic, a mixture density network - recurrent neural network (MDN-RNN), is employed. The MDN-RNN is a simulation-based modeling method to model time series data with stochastic behavior by using a probability distribution as the output part of the RNN. Using MDN-RNN, I trained a simulator to predict the behavior of *C. elegans* 0.2 seconds later based on the behavior of the past 20 seconds. After training the simulator using MDN-RNN, it was confirmed that the simulator generated behaviors with the same behavioral states and dynamics as those of the real *C. elegans* through analysis using t-SNE and time-delay embedding methods.

To understand the behavior that consists of multiple stochastic states, I analyzed the internal representation of the model that can generate the behavior similar to that of a real *C. elegans*. The results showed that each MDN component represented different behaviors. As a result, it was found that the dynamics of different behavioral states were extracted and represented in each component.

Chapter 2 Behavior control via reinforcement learning

In Chapter 2, I developed a computational control mechanism for behavior, and showed that a machine can extract behavioral strategies found in nature by automatically searching for a control mechanism that suits the desired task. In order to achieve this, I applied reinforcement learning as a control algorithm, and aimed to replace some of the computational mechanisms of the nervous system with a computer in order to accomplish the tasks that animals perform in nature. It was shown that the computer can acquire and reproduce behavioral strategies similar to those actually performed by animals without prior information.

Appendix A Disentangling animal behavior via temporal conditional-subspace VAE

In Appendix A, I aimed to analyze the topology of the dynamics behind the behaviors of animals belonging to different classes by separating the behaviors that are characteristic of each class from the behaviors that are common regardless of the class. For example, when the behaviors of wild-type animals and model animals for psychiatric disorders are acquired, it is important to extract the behaviors exhibited only by the disease model animals and investigate the generation mechanism of the behaviors to clarify the disease. To achieve this goal, I applied the conditional subspace - variational autoencoder (CS-VAE), which takes behavioral data consisting of multiple groups as input and divides them into elements characteristic of the group to which they belong and elements common to all groups in the middle layer of the VAE. The CS - VAE takes behavioral data consisting of multiple groups as input and divides it into elements characteristic of the group to which it belongs and elements common to all groups in the middle layer of the VAE. This is achieved by minimizing the amount of mutual information between the labels of the groups to which they belong and the features of the latent space in the middle layer. The effectiveness of this method has been verified using toy models and will be verified using animal behavior in the future.

Conclusion

In this dissertation, I have succeeded in separating animal behaviors with probabilistic and multiple control states by using representation learning of deep generative models. I also succeeded in automatically learning a policy to control a virtual animal by reinforcement learning. In this study, *C. elegans* was used as a model animal, but this model can be applied to other animal species as well. I aim to further develop these methods to elucidate the process of generating behavior from neural activity in an interpretable manner.