Acquirement of Stand-up Behavior by Reinforcement Learning on Multiple Joint Model

T08M001 Kazuma Asai
Advisers Hiroshi Nakano
Graduate Course of Applied Mathematics and Informatics
Graduate School of Science and Technology
Ryukoku University

Abstract

Reinforcement learning is one of machine learning techniques. It is applicable to various types of learning problems which can not get desired input-output pairs.

Locomotion learning of robot resembling a human in its shape is a complicated probrem. The mechanical structure is intricate, and its environment is unpredictable. Reinforcement learning without, explicit desired input-output pairs is suitable for the complicated robot control.

Q-learning is a reinforcement learning technique. Learning algorithm of Q-learning is such a simple, and easy to understand. We have applied Q-learning to stand-up problem on our multiple joint model in this study. Real environment has any uncertainty factors. We aim to stabilize stand-up behavior of the multiple joint model on real environment.

First, we have described the multiple joint model that imitate lower part of human body. The multiple joint model has three joints. Three joints are placed each applicable joint parts of human body that are waist, knees and ankles.

Next, we tried to get a stand-up behavior on our multiple joint model with Q-learning. State value of Q-learning is defined by the degrees of each joint angle and their angular velocities. Action value is defined by given torque to each joints of our model.

We have confirmed that Q-learning makes a stand-up behavior on our multiple joint model by use of one(two joints fixed), and by two joints(one joint fixed). However, number of states increses steeply with number of movable joints. It means that the learning steps to get a desired stand-up behavior increases steeply with the number of movable joints.

In this study, we modified the ε-greedy action choice method in Q-learning to hasten its learning speed. Parameter value ε of the ε-greedy method gives the probability of randomly action choice. The modified ε value with learning step T is

$$\varepsilon(T) = \varepsilon_0 \cdot \exp(-T/\beta)$$

thus, ε_0 is initial vale of $\varepsilon$with T = 0 and β is the decay constant.

We have succeeded in reduction of learning steps with $\varepsilon(T)$, and its acquired stand-up behavior is equivalent to the behavior with fixed ε.

Last of all, we have confirmed the stability of obtained standing physical position. We tried to put various external forces to our multiple joint model which postures is obtained standing position. However, it was instable against several external forces. Then, we have given an additional learning to our model. Its learning algorithm and parameters on additional learning does not been changed, but only initial condition is different. After additional learning, we have obtained a stable standing posture against the external forces.

Our result suggests that the stable standing position which is learning with the reinforcement learning can deal on uncertainty environment.