The Research Report of the Faculty of Engineering, Kagoshima University, No.55 (2013)

Doctoral Dissertation, Graduate School of Science and Engineering, Kagoshima University, Doctorate Vol.378

Study on Reinforcement Learning using Voronoi Diagram in Continuous State Space

Kathy Thi Aung

1. Introduction

There are several kinds of learning methods but reinforcement learning (RL) is the most suitable method in machine learning that deals with the decision to take an action using an agent at discrete time steps, and it is expected that would be useful anywhere in the future. Q-learning can apply in many practical applications but it works only for discrete state space, and difficult to handle in continuous state space because of curse of dimensionality problem therefore it needs to discretize the state space into a lot of smaller discrete regions when we treat such continuous cases.

This dissertation proposes VQE (Voronoi Q-value Element) to be able to apply the Q-learning method on continuous state space. As a method of division, we apply Voronoi diagram which is a general space division. Nevertheless, Voronoi diagram has a lot of flexibility thus a method of position determination of VQE becomes a problem.

2. Results and Discussions

We examines the learning performance of various strategies using the coincidence of action space and state space model that called BugPos model, and non-coincidence of action space and state space model that called BugMain model based on reward values in a single-agent environment, and decide how to act in certain state. In order to test our hypotheses, we experimented by rotating the agent's actions angle and angle of rotation of VQE by the angle in 5 times interval between 0 degrees and 90 degrees in which VQE are arranged in a lattice structure. Moreover, a random arrangement of VQEs experiment also conducted to correctly evaluate the optimal Q-values for state and action pairs in order to deal with continuous-valued inputs. As a result of our experiments using 4-actions BugPos model, the performance is good when the case of angle of VQE and angle of action have shifted 45 degrees has expressed.

We present the addition method of VQE as a position determination method in order to realize a Voronoi region since the performance of Q-Learning changes according to the arrangement of VQE. Moreover, the simulation was performed in BaitViewWorld experiment model and the learning performance was examined. And we also present a new adaptive segmentation of continuous state space based on vector quantization algorithm such as LBG (Linde-Buzo-Gray) for high-dimensional continuous state spaces. We constructed our single-agent model in continuous state and discrete actions spaces using Q-learning function. Moreover, the study of the resulting state space partition reveals in a Voronoi tessellation. In addition, the experimental results show that this proposed method can partition the continuous state space appropriately into Voronoi regions according to not only the number of actions, and achieve a good performance of reward based learning tasks compared with other approaches such as square partition lattice on discrete state space.

3. Conclusions

This thesis mainly presents a study on state space partitioning using Vorornoi diagram based on Q-Learning algorithm with the use of VQE. Here we manly present a study on solving curse of dimensionality problem conducting the normal Q-Learning on continuous state space in a single-agent environment. We aim by this research to speed up the learning efficiency in different situations as well as decrease the learning time. In order to do that we proposes VQE in various method such as Voronoi space division, rotating VQE, addition of VQE, addition and integration of VQE, etc. in several versions. In addition we present a better performance of learning for the algorithm.

Date of Conferment of the Degree: March 25 2013
Kagoshima University