Learning Dynamics of Reinforcement Learning with Psychical Satiation

T12M006 Hiroki Hamaguchi
Advisers Hiroshi Nakano
Graduate Course of Applied Mathematics and Informatics
Graduate School of Science and Technology
Ryukoku University

Abstract

There are two difficulties in implementing reinforcement learning. One is how learning space should be defined. The other is the trade–off problem between exploration and exploitation. We introduce a function approximation method and psychical satiation effect into reinforcement learning to solve these difficulties. Our psychical satiation effect extends applicable problem area of reinforcement learning. Furthermore, we try to interpret our several result as the learning dynamics.

On continuous state space, performance of reinforcement learning depends on description ability for the learning space. In other words, we should describe the continuous space to a finite number of parameters. We propose a self–organized online approximate method for continuous state space. In our method, a continuous state spaces is approximated by a finite set of local linear function modules which are learned from individual input–output pairs. The local function modules are newly created by approximation error that exceeds the threshold, and united into one module by similarity between the neighboring local function modules. Also, we propose a nearest neighbor algorithm for dynamic topological database, and measure it against a existing kd–Tree method.

In our psychical satiation method, an activity coefficient for each state–action pair is added to Q–Learning method. Each activity coefficient increases while the state–action pair is selected. The state–action pair temporarily fall into disuse if its activity coefficient exceeds the threshold value. Our psychical satiation method is able to get out of any local solutions. Therefore, learning process with our method has a dynamics on the local solution network. We have analyzed the local solution network dynamics with simple discrete Toy models.

Our result provides a novel point of view on reinforcement learning process differ from Goal–oriented learning paradigm. Our method is shown effectiveness of optimization method. Furthermore, our method is expected that provides dynamical point of view.