Entropy Regularization for
Scalable, Safe and Robust Reinforcement Learning

Name

Lingwei Zhu

Laboratory’ｓ name

Robot Learning Lab

Supervisor’ｓ name

Takamitsu Matsubara

Abstract

The last decade has been an era of Reinforcement Learning (RL). It has witnessed tremendous successes in lab robotics or games such as Go. Basically, RL describes how an agent should act in response to a given environment in order to maximize its feedback or minimize the penalty. Such mechanism has been verified to exist in animals and humans. Ample examples exist in biology such as a dog learning to adapt to a process of ringing first followed by food. When the dog fully adapts to such process, saliva is detected when only ringing is presented without the food. This learning to adapt can be abstracted and described in a self-contained mathematical formulation known as the Reinforcement Learning (RL) where the agent learns to make decision that maximizes its return (food in the case above). However, unlike animals that have the notion of efficiency and risk-awareness, the standard RL formulation offers only a way to reach the goal but not about how the goal is reached. As a consequence, the RL agent might take extraorbitantly long to reach the goal, and/or incur unacceptable cost alone its path. Furthermore, the classic RL algorithms are greedy, hence prone to errors and noises. Due to the above-mentioned issues, currently RL falls short of realistic applications compared to other branches of contemporal machine learning such as supervised learning that has applications permeated in our daily life. In this thesis, I attempt to bridge the gap by proposing novel and useful theories based on entropy regularization and validate them on practical scenarios such as industrial plants or robotics. A commercial-level large-scale chemical process simulator for a Vinyl Acetate Monomer plant is frequently used as the testbed. This thesis mainly tackles the issues of scalability, safety and robustness in RL via the tool of entropy. I hope the thesis could play an important role in the coming resurgence of interest in pratical RL deployment and applications.