Chapter 2
The Explainable Model to Multi-Objective Reinforcement Learning Toward an Autonomous Smart System

Tomohiro Yamaguchi
Nara College, National Institute of Technology, Japan

ABSTRACT

The mission of this chapter is to add an explainable model to multi-goal reinforcement learning toward an autonomous smart system to design both complex behaviors and complex decision making friendly for a human user. At the front of the introduction section, and a relation between reinforcement learning including an explainable model and a smart system is described. To realize the explainable model, this chapter formalizes the exploration of various behaviors toward sub-goal states efficiently and in a systematic way in order to collect complex behaviors from a start state towards the main goal state. However, it incurs significant learning costs in previous learning methods, such as behavior cloning. Therefore, this chapter proposes a novel multi-goal reinforcement learning method based on the iterative loop-action selection strategy. As a result, the complex behavior sequence is learned with a given sub-goal sequence as a sequence of macro actions. This chapter reports the preliminary work carried out under the OpenAIGym learning environment with the CartPoleSwingUp task.

INTRODUCTION

The mission of this chapter is to add an explainable model to multi-goal reinforcement learning (Drugan 2017)(Yamaguchi 2022) toward an autonomous smart system to design both complex behaviors and complex decision making friendly for a human user. At the front of the introduction, a relation between reinforcement learning including an explainable model and a smart system is described.

DOI: 10.4018/978-1-6684-7684-0.ch002

Smart systems incorporate functions of sensing, actuation, and control in order to describe and analyze a situation, and make decisions based on the available data. In most cases the "smartness" of the system can be attributed to autonomous operation. To make systems smarter, third-generation smart systems combine "intelligence". To realize this, DARPA formulated the explainable artificial intelligence (XAI) program with the goal to enable end users to better understand, trust, and effectively manage artificially intelligent systems (Gunning, 2021). **Table 1** shows the issues on a smart system with explainable AI (XAI) mainly on learning capabilities. The research field of a smart system is expanded from data analytics to autonomous control task. This chapter focuses on the latter task. The characteristics of the autonomous control is that the model has multiple states and it handles state transitions with actions. Reinforcement learning is one of the machine learning methods to learn autonomous control associated with multiple decision making problems including state transitions.

Table 1. It is the issues on a smart system with explainable AI

machine learning task	machine learning method	problem
data analytics	supervised learning	classification
	unsupervised learning	clustering
autonomous control	reinforcement learning	multiple stage decision making

Therefore, this chapter describes the exploration of various behaviors toward sub-goal states in an efficient and systematic manner through the use of repeated loop-action. There are three main features. The first one is that the iterative loop-action simplifies the result of an action sequence. The second feature is that it explores the action space with a given sub-goal set as a guide for collecting behaviors toward one of the sub-goals. The third feature is that a macro action is learned from a collected behavior, which consists of a certain length of repeated loop-actions between two sub-goal states.

In recent years, deep reinforcement learning research has made significant progress by utilizing images as perceptual input (Mnih 2015). This has expanded the scope of application for reinforcement learning to include TV game play simulators, such as the Atari 2600 TV game machine from the 1970s. In several Atari 2600 TV games, deep reinforcement learners have successfully surpassed the highest score achieved by human experts using evaluation criteria such as maximizing the game score (Mnih 2015).

Reinforcement learning (RL) is a popular algorithm used for automatically solving sequential decision problems. It is commonly modeled as Markov decision processes (MDPs) (Puterman 1994). While there are numerous RL methods available, many of them focus on single-objective settings where the agent's goal is to determine a single solution based on an optimality criterion. However, it has become evident that learning becomes challenging in games that require context-based learning, such as anticipating future events (e.g., obtaining keys for later stages) or making choices at branching points that do not immediately generate rewards. In such cases, imitation learning methods that leverage action sequences from human experts, particularly those capable of handling context like the Decision Transformer (DT) method proposed by Chen (2021), are expected to be beneficial. Nevertheless, when dealing with unknown goals or environments, applying imitation learning can be difficult due to the lack of prior knowledge about the model's behavior sequences.

Related Content