Chapter 6
Strategy Selection and Outcome Evaluation of Change-Based Three-Way Decisions Based on Reinforcement Learning

Chunmao Jiang
Fujian University of Technology, China

ABSTRACT

In this chapter, we enhance the trisecting-acting-outcome (TAO) model of three-way decision-making (3WD) with a novel approach for strategy selection and outcome prediction using Q-learning in reinforcement learning. We reinterpret the changes in tripartition and actions in the TAO model as states and actions in reinforcement learning, respectively. The reward is quantified using cumulative prospect theory, and the Q-learning algorithm iteratively determines action sets that achieve target rewards efficiently. This method offers a cost-effective and psychologically attuned action set for predicting the utility in change-based 3WD, demonstrated through a practical example.

INTRODUCTION

The three-way decision (3WD) model is consistent with human cognition and offers a nuanced semantic framework for understanding decision-making processes. This approach categorizes decisions into three distinct categories: acceptance, non-commitment, and rejection, each derived from the positive, boundary, and negative regions of rough set approximations, respectively (Yao, 2009). In 2012, Yao refined this concept, proposing a model that divides a universal set into three segments to facilitate more precise classifications (Yao, 2012).

Over time, the scope of 3WD has expanded beyond its original probabilistic rough set basis, evolving into a broader conceptual framework, to form a complex research network (Yang & Li, 2019). This more expansive interpretation of 3WD emphasizes a triadic approach to thought and problem solving, a method prevalent across various disciplines and supported by cognitive theory. This approach simplifies complex

DOI: 10.4018/979-8-3693-1582-8.ch006

problems by dividing them into three interconnected but distinct components, reducing cognitive and information overload (Yao, 2019). In its evolution and development, a large number of research results were produced (Fang & Min, 2019; Liu & Liang, 2016; Li & Huang, 2017; Min & Zhu, 2012; Qian & Liu, 2020; Wang & Yao, 2018; Zhang & Li, 2020, Zhang & Pang, 2020; Yu & Wang, 2020; Yang & Li, 2019; Li & Huang, 2017; Zhan & Wang, 2023; Wang & Li, 2022; Wang & Ma, 2022).

A notable example of this expansive interpretation is the trisecting-acting-outcome (TAO) model, where 'trisecting' refers to dividing the universal set into three parts, and 'acting' is the process of selecting and providing the optimal action that fulfills the result and maintains them at a moderate level. On the one hand, based on the required outcome, the decision maker designed appropriate strategies for the tripartition to achieve the expected results. On the other hand, strategies are selected to maintain the outcome at a moderate level. The outcome refers to the effectiveness of the results of trisecting and acting.

Jiang and Yao (2018) introduced a quantitative evaluation method based on a general 3WD model, movement-based 3WD, and Jiang and Guo (2020) developed a probabilistic model for strategy selection based on information entropy. They also introduced the proportional utility function (PUF) for assessing outcomes from dual perspectives in movement-based 3WD (Jiang & Guo, 2021; Jiang & Guo, 2022). In addition, the change-based TAO model (Jiang & Duan, 2022) was proposed to measure effectiveness by integrating interval sets and distinctive utility functions. This framework uses interval sets to represent the impact of strategies or actions, using a utility measurement method to assess changes in three intervals and amalgamating these utilities for a comprehensive effectiveness evaluation.

However, some key issues still need to be adequately addressed when dealing with change-based three-way decisions. The goal of a change-based three-way decision is to change the object from the original tripartition to a more effective one. The three-way decision based on changes explains and answers one type of question: Why do we need to 'Trisecting and Acting'? That is, the three-way decision is intended to produce more satisfactory changes. In this chapter, we develop a method for selecting optimal strategies and calculating future utility based on change-based three-way decisions.

This chapter applies reinforcement learning to predict future utilities and generate a set of strategies to reach the target reward in the shortest cycle during future utility computation using the Q-learning algorithm. During the prediction process, altered tripartition can be treated as states in reinforcement learning. Multiple strategies for applying to the altered tripartition are available and these strategies are referred to as actions in reinforcement learning. Each time the strategy has changed the tripartition, it has produced a utility referred to as a reward for reinforcement learning.

The Motives and Contribution of This Chapter

The primary motivation behind this research is to address critical issues in change-based three-way decisions (C-3WD) and to improve the understanding of the Trisecting-Acting-Outcome (TAO) model in decision-making. While previous studies have mainly focused on the 'trisecting' and 'Acting' phases, little attention has been given to the 'outcome.' Additionally, there is a need to find strategies that can achieve desired utility efficiently in a finite cycle, reducing trial-and-error costs.

To address these challenges, this research proposes a method that leverages reinforcement learning to predict future utilities and generate optimal strategies for change-based three-way decisions. Key contributions of this work include the following:

Related Content