Q-LEARNING IN AXELROD'S IPD TOURNAMENT

ABSTRACT. In this paper I replicate and advance the experimental work done by Sandholm, Tuomas W., and Robert H. Crites (Biosystems 37.1 (1996): 147-166) into evaluating the use of reinforcement learning techniques in the iterated prisoner's dilemma (IPD). Specifically, I further their investigation of the impact of varying the learning rate, discount rate, and exploration schemes of q-learning strategies for the IPD. Additionally, where their work stopped at examining IPD engagement between two individual strategies, I have extended my experimentation to cover the more general setting of an IPD-tournament as popularized by Robert Axelrod in his 1984 book: The evolution of Cooperation.

1. MOTIVATION

Politicians and the general population largely rely on polls and predictions made by various experts for the prediction of electoral results. The failure of recent predictions made by such experts in the United Kingdom and the United States motivates the search for better tools and techniques. This paper presents a series of techniques based on on-line learning to tackle this problem.

2. PRELIMINARIES

Let $I \subseteq \mathbb{R}$ be an open interval and $f\colon I \to \mathbb{R}$ a convex function. It is known that f is Lipschitz continuous on any interval $[a, b] \subset I$, f admits both left and right derivatives f'_- and f'_+ over I, both non-decreasing, and f is differentiable everywhere except for a set that is at most countable.

3. TAYLOR-TYPE THEOREMS

Lemma 1. *Let $I \subseteq \mathbb{R}$ be an open interval and $f\colon I \to \mathbb{R}$ a convex function. Then, the following holds for all $a, b \in I$:*

$$f(b) - f(a) = \int_a^b f'(t)\,dt = \int_a^b f'_+(t)\,dt = \int_a^b f'_-(t)\,dt.$$

Proof. The result are based on [**?**]. Since f is Lipschitz continuous over the closed interval in I containing a and b, it is absolutely continuous, which proves the first equality. The second equality is clear for Lebesgue integrals since $f'(t) = f'_+(t) = f'_-(t)$ for all but at most a countable set of points. $\square$