Chapter VIII

Multi-Agent Evolutionary Game Dynamics and Reinforcement Learning Applied to Online Optimization for the Traffic Policy

Yuya Sasaki
Utah State University, USA

Nicholas S. Flann
Utah State University, USA

Paul W. Box
Commonwealth Scientific and Industrial Research Organization, Australia

ABSTRACT

This chapter demonstrates an application of agent-based selection dynamics to the traffic assignment problem. We introduce an evolutionary dynamic approach that acquires payoff data from multi-agent reinforcement learning to enable an adaptive optimization of traffic assignment, provided that classical theories of traffic user equilibrium pose the problem as one of global optimization. We then show how these data can be employed to define the conditions for evolutionary stability and Nash equilibria. The validity of this method is demonstrated by studies in traffic network.
modeling, including an integrated application using geographic information systems applied to a complex road network in the San Francisco Bay area.

INTRODUCTION

When we think about alternative driving routes from an origin to a destination, there are several factors of travel time to consider, such as the distance, speed limit, and possible congestion. Although people appear to have an incentive to use the shortest distance routes, this does not happen in reality, because the supply function of traffic roads exhibits an increasing cost nature. In other words, we have monotonically increasing travel time with respect to increasing traffic flow volume.

Assume a simple example, where two agents travel from an origin to a destination, with only two paths available, a short path and a long path. Also assume that the long path has greater capacity, with more lanes than the shorter path. The supply function of the longer path has a flatter slope and a higher intercept (free-flow travel time) than the short path. Suppose that the two agents (denoted by A and B) make decisions simultaneously. Then the outcome cost matrix for the four possible pure-strategy combinations will look like Table 1. If both choose the greedy strategy (the short path), the tragic consequence is that both get trapped in severe congestion. The noncooperative equilibria in this pure strategy setting are thus the symmetric pair of (A:1 B: 2) and (A: 2 B: 1) in Table 1. Whenever some agents choose the shortest paths, others have to compromise by making inferior choices. The characteristics of traffic behaviors illustrated by this example will be analyzed more generally in the succeeding sections.

Notations and Definitions of Traffic Networks

A traffic network consists of nodes and arcs, where an arc is always closed by a pair of nodes. We use the term “O-D pair” to refer to a pair of origin and destination nodes in an agent’s trip. A path is an ordered sequence of arcs that connects two nodes that are not necessarily adjacent to each other. Thus, an O-D path refers to an ordered sequence of arcs that connects the origin and destination nodes. Let I, J, and R denote sets of arcs, nodes, and paths in a network, respectively. R must be defined for each O-D pair, and I and J can be global for all the O-D pairs. Since we will focus first on the case of one O-D pair (later in this chapter we will consider multiple origins and one destination), we will not use any subscript or superscript to R. Let \( I_j \) denote the set of arcs radiating from node \( j \in J \), where \( I_j \) also represents the set of available strategies for agents being at node \( j \). An arc is a component of the network in the entire system, although it can also be perceived as a “strategy,” “next move,” or “action” from decision-makers’ viewpoint. We will exclusively use the term *arc* when the network structure is discussed, and the

<table>
<thead>
<tr>
<th>A’s Action</th>
<th>B’s Action</th>
<th>short path</th>
<th>long path</th>
</tr>
</thead>
<tbody>
<tr>
<td>short path</td>
<td>A: 5 B: 5</td>
<td>A: 1 B: 2</td>
<td></td>
</tr>
<tr>
<td>long path</td>
<td>A: 2 B: 1</td>
<td>A: 3 B: 3</td>
<td></td>
</tr>
</tbody>
</table>

Table 1. Example payoff matrix of four combinations of choices. (Note. The values are travel time in any time unit).
Pricing Basket Options with Optimum Wavelet Correlation Measures
Christopher Zapart, Satoshi Kishino and Tsutomu Mishina (2006). Computational Economics: A Perspective from Computational Intelligence (pp. 34-61).
www.igi-global.com/chapter/pricing-basket-options-optimum-wavelet/6779?camid=4v1a