Technical Report Documentation Page

1. Report No. SWUTC/13/600451-00065-1	2. Government Accession No.	3. Recipient's Catalog No.
4. Title and Subtitle GAME THEORY AND TRAFFIC ASSIGNMENT		5. Report Date September 2013 6. Performing Organization Code
7. Author(s) Tarun Rambha and Stephen D. Boyles		8. Performing Organization Report No.
9. Performing Organization Name and Address Center for Transportation Research University of Texas at Austin 1616 Guadalupe Street, Suite 4.200 Austin, Texas 78701		10. Work Unit No. (TRAIS) 11. Contract or Grant No. DTRT12-G-UTC06
12. Sponsoring Agency Name and Address Southwest Region University Transportation Center Texas Transportation Institute Texas A&M University System College Station, Texas 77843-3135		13. Type of Report and Period Covered Final Report 14. Sponsoring Agency Code

15. Supplementary Notes

Project supported by general revenues from the U.S. Department of Transportation, University of Transportation Centers Program, and general revenues from the State of Texas.

Project title: Game Theory and Traffic Assignment: Refinements, Stability, and Tractability

16. Abstract

Traffic assignment is used to determine the number of users on roadway links in a network. While this problem has been widely studied in transportation literature, its use of the concept of equilibrium has attracted considerable interest in the field of game theory. The approaches used in both transportation and game theory disciplines are explored, and the similarities and dissimilarities between the m are studied. In particular, treatment of multiple equilibrium solutions using equilibrium refinements and learning algorithms which convergence to equilibria under incomplete information and/or bounded rationality of players are discussed in detail.

17. Key Words	18. Distribution Statement
Congestion Games, Traffic Assignment, Learning in Games, Equilibrium Refinements.	No restrictions. This document is available to the public through NTIS: National Technical Information Service 5285 Port Royal Road Springfield, Virginia 22161

19. Security Classif.(of this report)	20. Security Classif.(of this page)	21. No. of Pages	22. Price
Unclassified	Unclassified	46	

Form DOT F 1700.7 (8-72) Reproduction of completed page authorized

Game Theory and Traffic Assignment

by

Tarun Rambha
Stephen D. Boyles

Research Report SWUTC/13/600451-00065-1

Southwest Region University Transportation Center
Center for Transportation Research
University of Texas at Austin
Austin, Texas 78712

August 2013

Disclaimer

Acknowledgments

The authors recognize that support for this research was provided by a grant from the U.S. Department of Transportation, University Transportation Centers Program to the Southwest Region University Transportation Center which is funded, in part, with general revenue funds from the State of Texas. Special thanks to Kai Yin for pointing us to useful literature on this topic.

Abstract

Traffic assignment is used to determine the number of users on roadway links in a network. While this problem has been widely studied in transportation literature, its use of the concept of equilibrium has attracted considerable interest in the field of game theory. The approaches used in both transportation and game theory disciplines are explored, and the similarities and dissimilarities between them are studied. In particular, treatment of multiple equilibrium solutions using equilibrium refinements and learning algorithms which convergence to equilibria under incomplete information and/or bounded rationality of players are discussed in detail.

Executive Summary

The traffic assignment problem (TAP) is one of the steps in the four-step transportation planning process and involves assigning travelers to different routes. This assignment is done based on the the user equilibrium (UE) principle according to which "*All used routes between an origin-destination (OD) pair are assumed to have equal and minimal travel times.*" This principle assumes travelers selfishly choose routes so as to minimize their travel time. The TAP has also been studied by economists, particularly in the field of game theory, due to the close resemblance of the UE principle to the concept of Nash Equilibrium (NE). In general, a game is modeled using three components: a set of players, their actions and their utility functions or payoffs. The TAP is modeled as a non-cooperative game and the NE is defined as a state in which no player can benefit by deviating from his/her current action. The games used to model traffic are also known as network or congestion games.

Among two popular ways to model the TAP as a game, the atomic version of a congestion game is widely studied and is a discrete version of the TAP in which the demand and flow on each link is constrained to be an integer. Travelers are non-cooperative players who choose from a common set of resources/road links. The payoff each player incurs in choosing a particular route is the travel time on it. In another approach to model the TAP as a game, each OD pair is considered a player who gets to distribute the OD demand along all possible paths between the OD pair. Travelers are assumed to be infinitesimally small and hence non-integral amounts of flow can be routed along a path. This is a non-atomic version of the congestion game. Congestion games fall under a class of games called potential games in which best responses for each player may be obtained by optimizing a single global objective or potential function. In other words, the utilities of each player may be substituted by the potential function.

In this report, two essential questions are addressed from a game theoretic perspective: which equilibrium is likely to be played in the presence of multiple NE solutions? and how do players know to play a NE?. Multiple NE solutions are distinguished using equilibrium refinements or solution concepts, which try to narrow down or "refine" the set of equilibrium solutions restricting them to more plausible/sensible ones. These refinements compute the NE, taking into account the possibility of deviations from the optimal strategy/path either by mistake or due to a lack of sufficient information on the payoffs. This feature distinguishes it from the idea of entropy maximization used in transportation literature for determining the most likely path flows.

The question of how players reach a NE is answered using a concept called learning in which players learn from repeated interaction and decide on an action based on the outcomes of previous rounds of play. Some of the learning algorithms used in congestion games are strikingly similar to algorithms used in the TAP and day-to-day dynamics. Learning algorithms not only answer to how a NE would be played but are also in a loose sense an answer to the question of which equilibrium would be played. Fictitious play is one such learning dynamic that is guaranteed to converge to a NE for congestion games. NE are steady states of a "fictitious" process involving repeatedly play of the original game in which, players best respond to their beliefs about the opponent's play (which are governed by the frequency of past actions) using a pure strategy.

A more realistic class of learning models (partly inspired from evolutionary game theory) relaxes the assumption of perfect rationality of players, and lets them err with some probability. The central idea in these models is to define a stochastic process using outcomes of a game as system states, and assume some dynamic which lets players move from one state to another. This process is then modeled as a Markov chain and its stationary or steady state distribution is used to study equilibrium solutions. As the probabilities of making mistakes get smaller (it is assumed that by repeated interactions players get more experienced) only a few states have positive limiting probabilities. Logit-response model is one such dynamic in which these limiting states with positive probability coincides with the argmin set of potential function.

This report reviews refinements and learning literature in game theory in the context of congestion games and presents examples to demonstrate its application in traffic models. Possible extensions of these approaches are also discussed.

Table of Contents

List of Figures

List of Tables

Chapter 1

Introduction

1.1 Traffic Assignment Problem

Urban transportation planning is traditionally carried out using a four-step method. The first three steps are used to estimate the number of travelers/users, their origin-destination (OD) pairs, and their mode of travel. The final step, also called *route choice* or *traffic assignment*, involves assigning travelers to different routes. This assignment procedure is done based on the concept of user equilibrium (UE), which states that "*All used routes between an OD pair have equal and minimal travel times.*" Proposed by Wardrop [24], the UE principle assumes that users selfishly choose routes so as to minimize their travel time. The equilibrium solution to the traffic assignment problem (TAP) can be expressed either in terms of link flows (volume of users on each roadway link) or path flows (volume of users on each path between every OD pair).

Consider a directed network $G = (V, E)$, where V and E are the set of nodes and arcs/links respectively. Assuming that the flow on an arc $e \in E$ is denoted by x_e, let the function $t_e(x_e)$ (also referred to as *link performance function* or *latency/delay function*) represent the travel time experienced by users on arc e. Suppose the set of OD pairs or zones is denoted by $Z \subseteq V^2$ and the demand between an OD pair $z \in Z$ is represented by d_z. We denote the set of paths between z by $P_z \subseteq E$ (we could consider only the set of simple paths, i.e., ones without any directed cycles) and the set of all paths in the network by $P = \cup_{z \in Z} P_z$. We will use the notation $e \in p$ to denote the set of links that belong to a path p. Assume that f_p denotes the flow on a path p. Let δ_{ep} represent the arc-path incidence variable, i.e., δ_{ep} is 1 if an arc e belongs to path p and is 0 otherwise. The following mathematical program, proposed by Beckmann et al. [2] (and hence popularly known as the Beckmann formulation), describes the traffic assignment problem (TAP):

$$\min \sum_{e \in E} \int_0^{x_e} t_e(x)\, dx \tag{1.1}$$

$$\text{s.t.} \quad x_e - \sum_{p \in P} \delta_{ep} f_p = 0 \qquad \forall\, e \in E \tag{1.2}$$

$$\sum_{p \in P_z} f_p = d_z \qquad \forall\, z \in Z \tag{1.3}$$

$$f_p \geq 0 \qquad \forall\, p \in P \tag{1.4}$$

$$x_e \geq 0 \qquad \forall\, e \in E \tag{1.5}$$

If the link performance functions are continuous, it is easy to verify that the objective of the TAP is continuous and differentiable. Hence, existence of an optimal solution follows directly from the fact that the objective is continuous and the constraints define a compact feasible region. If it is also assumed that the link performance functions are non-decreasing, the objective is convex and hence every equilibrium solution has equal link travel times. Furthermore, if the link performance functions are strictly increasing then there exists a unique solution in link flows to the TAP.

At low volumes, the travel time on a roadway link is usually insensitive to increase in flow but as more travelers use it, variability in driver behavior and speeds results in an increase in time taken to traverse the link. Hence, one expects that the link performance functions be non-decreasing. A widely used class of link performance functions are functions of the type $t_e(x_e) = t_e^0 \left(1 + \alpha(x_e/C_e)^\beta\right)$ (also known as the Bureau of Public Roads (BPR) function), where C_e and t_e^0 denote the capacity of link e and its free-flow travel time respectively, and α and β are parameters. Although these functions are strictly increasing, when carefully calibrated, they are almost flat at low volumes. Hence, an equilibrium solution obtained using BPR functions in unique in link flows.

Remark. The assumption that the travel time on an arc depends only on the flow on it is also known as the *separability condition*. Relaxing this assumption leads to a more general traffic assignment formulation which can help model impacts of intersections. These problems are usually expressed as a Variational Inequality (VI) (see Smith [21] and Dafermos [7]). Another widely studied variant of the TAP is called dynamic traffic assignment (DTA), which captures the time-of-day variations in traffic (see Peeta and Ziliaskopoulos [16]). Although these models are more realistic they are harder to analyze as games, and hence we will restrict our attention to the version described in Section 1.1 thought the remainder of this report.

The mathematically appealing nature of the TAP has led to the development of several efficient algorithmic approaches to compute the unique equilibrium link flow solution. Two immediate questions of interest that may be raised are the following:

1. What are the equilibrium path flow solutions?

2. How does one guarantee that the equilibrium will be reached?

Let us address the first question. Algorithms to solve the traffic assignment problem predominantly compute link flows for a couple of reasons. First, the number of links in the network is a lot smaller than the number of paths (which, in fact, grows exponentially with network size); hence, link flow solutions are easier to compute. Second, link flow solutions are sufficient for most applications of transportation planning including congestion pricing and network design. However, for a few applications such as select link analysis and estimation of sub-network OD flows, path flow solutions are necessary. Unfortunately, one cannot construct a unique path flow solution from an equilibrium link flow solution.

Hence, researchers have tried to find a path flow solution that is more likely to occur. Zuylen and Willumsen [26] proposed the entropy maximization principle to estimate OD flows from traffic counts/link flows. Bell and Iida [3] described the use of this method to compute the most likely path flow solution from equilibrium link flows. More recently, Bar-Gera [1] developed efficient algorithms to identify the entropy-maximizing solution. The central idea in these methods is to distribute the demand between each OD pair across as many routes as possible so as to increase the number of permutations of users, subject to the constraint that the total number of travelers on each link equals the equilibrium link flow. It was shown by Bar-Gera [1] that entropy-maximization encompasses a behavioral property called the condition of proportionality.

Now consider the second question. Equilibrium states are generally normative, i.e., they prescribe an ideal situation in which every traveler is better off. To determine the equilibrium solution, it is necessary for travelers to be perfectly rational and have a complete knowledge of the link performance functions. However, when a large number of travelers interact, the extent of reasoning required to arrive at an equilibrium solution remains beyond one's human ability. In fact, in large networks it is very unlikely that travelers even realize when a system is at equilibrium. This issue has received considerable attention in literature and equilibrium is modeled as steady states of a stochastic process (see Smith [22] and Cascetta [6]). This approach also helps determine the rate at which a system converges after a network disruption and the role of advanced traveler information systems (ATIS) in reaching an equilibrium.

1.2 Game theoretic models for traffic

The TAP has also been widely studied by economists due to the close resemblance of the UE principle to the concept of Nash Equilibrium (NE). The TAP is modeled as a non-cooperative game and the NE is defined as a state in which no player can benefit by deviating from his/her current action. Two common approaches used are summarized in Table1.1. The problem of routing selfish users in a network is dubbed as *congestion games* or *network games*. It is also studied under a broader class of games called *population games* (ones that involve a large number of players).

Table 1.1: Game theoretic formulations of the TAP

Components	*Atomic Congestion Game*	*Non-atomic Congestion Game*
Players	Travelers	OD Pairs
Actions	Paths	Demand assignment across paths
Disutilities	Travel time on path	A Beckmann-like function

The atomic version of a congestion game is a discrete version of the TAP in which the demand and flow on each link is constrained to be an integer. It first proposed by Rosenthal [18] and later popularized by Monderer and Shapley [12]. Travelers are modeled as non-cooperative players who choose from a common set of resources/road links. The payoff each player incurs in choosing a particular route is the travel time on it. For most part, we will deal with atomic congestion games in this report as it is a finite normal form game, which lets us use several important theorems on existence of different types of equilibria. In another approach to model the TAP as a game, each OD pair is considered a player who gets to distribute the OD demand along all possible paths between the OD pair. Travelers are assumed to be infinitesimally small and hence non-integral amounts of flow can be routed along a path. This is a non-atomic version of the congestion game (see Devarajan [8]) and is very identical to the TAP but is broader in the sense that every NE to this game is an UE solution but not vice-versa.

The questions posed earlier in Section 1.1 have long been raised in game theory literature. The existence of multiple equilibria in games is widespread. In fact, for atomic congestion games, even strictly increasing cost functions can result in multiple link flow solution. In such cases, it is of interest to determine which of the equilibria is more likely to occur. A vast amount of literature has focused on this issue and has led to the development of several important behavioral concepts. These approaches, known as *equilibrium refinements* or *solution concepts*, try to narrow down or "refine" the set of equilibrium solutions restricting them to more plausible/sensible ones. Some of the popular equilibrium refinements include perfectness (Selten [19]), properness (Myerson [14]), and stable sets (Kohlberg and Mertens [11]). These refinements compute the NE, taking into account the possibility of deviations from the optimal strategy/path either by mistake or due to a lack of sufficient information on the payoffs. This feature distinguishes it from the idea of entropy maximization in determining the most likely path flows.

As far as the question of reaching an equilibrium is concerned, a popular approach followed by economists is the theory of learning in which players learn from repeated interaction and decide on an action based on the outcomes of previous rounds of play. Congestion games are a type of *potential game* for which several learning algorithms exhibit interesting convergence properties. Some of the learning algorithms used in congestion games are strikingly similar to algorithms used in the TAP and day-to-day dynamics. Learning algorithms not only answer to how a NE would be played but are also in a loose sense an answer to the question of which equilibrium would be played. In the context of congestion games, they also answer the question of multiple path flows as they operate in the space of path flows. However, the equilibrium solution to which they converge may depend on factors such as the initial/starting condition. Players in these learning models are modeled using the following three components: Inertia, Myopic behavior Mutations/Noise or Trembles

1.3 Organization of report

The rest of this report is organized as follows: Chapter 2 contains a description of congestion and potential games and the concept of equilibrium refinements. Chapter 3 surveys learning algorithms in congestion games. In Chapter 4, we exhibit an example of using a learning algorithm in conjunction with an equilibrium refinement to obtain a more stable/robust solution, and discuss the pointers to future research on this topic.

Chapter 2

Preliminaries

In this chapter, we review the concept of a NE, and define potential and congestion games. Equilibrium refinement techniques used to distinguish multiple equilibria are also discussed using a few examples.

A *normal form game* $\Gamma = (A_i, u_i)_{i \in N}$ is characterized by three components: A finite set of players $N = \{1, 2, \ldots, n\}$, set of actions or pure strategies for each player i(denoted by A_i), and utility functions or payoffs, $u_i : A \to \mathbb{R}$, where $A = \times_{i \in N} A_i$. Given a player i, the set of players $N \backslash \{i\}$ is denoted by $\{-i\}$. Γ is a *finite game* if the action space of each player is finite. A game which is not finite is referred to as an *infinite game*.

The set of mixed strategies for player i is denoted by S_i, and an element of S_i is represented by the vector s_i. The probability with which player i chooses an action a_i (or the marginal probability) is written as $s_i(a_i)$. In this report, we will usually reference players using subscripts and use superscripts for time steps or sequence indices. The expected utility for player i for a given mixed strategy $s \in S$, where $S = \times_i S_i$ is $u_i(s) = \sum_{a \in A} u_i(a) s(a) = \sum_{a \in A} u_i(a) \prod_{i \in N} s_i(a_i)$.

Definition 2.1. Given $s \in S$, the best-response correspondence of player i, denoted as $BR_i(s)$, is defined as $BR_i(s) = \text{argmax}_{s_i' \in S_i} u_i(s_i', s_{-i})$.

Definition 2.2. A strategy profile $s^* \in S$ is a Nash equilibrium (NE) $\Leftrightarrow$ $s_i^* \in BR_i(s^*)$, $\forall\, i \in N$. Alternately, s^* is a NE $\Leftrightarrow$ for each player i, $u_i(s_i^*, s_{-i}^*) \geq u_i(s_i, s_{-i}^*) \,\forall\, s_i \in S_i$.

Definition 2.3. Let $s \in S$ and let $\epsilon > 0$. s is an ϵ-Nash Equilibrium if for each i, $u_i(s) \geq u_i(s_i, s_{-i}) - \epsilon \,\forall\, s_i \in S$

The NE in the above definitions is also referred to as a *mixed strategy* NE. If each players mixed strategy is degenerate, the NE is said to be a *pure strategy* NE. A mixed strategy profile

s^* is a *strict* NE if each player's action is his/her unique best response, (i.e., $|BR_i(s^*)| = 1$) or if for each player i, $u_i(s_i^*, s_{-i}^*) > u_i(s_i, s_{-i}^*) \, \forall \, s_i \in S_i$. The existence of a NE in mixed strategies follows from Nash's famous theorem [15].

2.1 Potential games

Potential games are a special class of games in which best responses for each player may be obtained by optimizing a single global objective (*potential function*). In other words, the utilities of each player may be substituted by the potential function.

Definition 2.4 (w-Potential Game/Weighted Potential Game)**.** Let $w = (w_i)_{i \in N}$ be a vector of positive numbers called weights. $\pi : A \to \mathbb{R}$ is a *w-potential or weighted potential* and Γ is an *w-potential or weighted potential game*, if for every player i and $\forall \, a_{-i} \in A_{-i}$

$$u_i(a_i, a_{-i}) - u_i(a_i', a_{-i}) = w_i[\pi(a_i, a_{-i}) - \pi(a_i', a_{-i})], \, \forall \, a_i, a_i' \in A_i$$

If $w_i = 1 \, \forall \, i \in N$, then Γ is called an exact potential game or simply a potential game. It may be easily established that for potential games there exists a unique potential up to an additive constant.

Table 2.1: A potential game - prisoner's dilemma

	Cooperate	Defect
Cooperate	$1, 1$	$10, 0$
Defect	$0, 10$	$5, 5$

Consider the prisoner's dilemma with payoffs as shown in Table 2.1. The potential for this game may be written as $\pi = \left[\begin{smallmatrix} 6 & 5 \\ 5 & 0 \end{smallmatrix}\right]$, since the difference in payoffs for the row player, assuming the column player cooperates, is $1-0 = 6-5$; and if the column player defects, the difference is $10-5 = 5-0$. Similarly the differences in payoffs for the column player is $1-0 = 6-5$ and $10-5 = 5-0$.

Definition 2.5. The sequence $\gamma = (a^1, a^2, \ldots)$ is a *path* in A, if $\forall \, k \geq 2 \, \exists$ a unique $i : a^k = (a_i, a^{k-1})$ for some $a_i \neq a_i^{k-1} \in A_i$ (i.e., given an element of the sequence, the next strategy profile is obtained by letting a single player deviate).

Definition 2.6. For a finite path of action profiles $\gamma = (a^1, a^2, \ldots, a^K)$, let $I(\gamma, u) = \sum_{k=2}^{K} [u_{i_k}(a^k) - u_{i_k}(a^{k-1})]$, where i_k is the unique deviator at step k. We say γ is *closed* if $a^1 = a^K$. Further, if $a^l \neq a^k$ for every $l \neq k$, then γ is called a *simple closed path* (i.e., no outcome is revisited). The *length* of a simple closed path is the number of distinct strategy profiles in it.

Theorem 2.1 (Monderer and Shapley [12]). *The following claims are equivalent:*

(a) Γ *is a potential game*

(b) $I(\gamma, u) = 0$ *for every finite closed path* γ

(c) $I(\gamma, u) = 0$ *for every finite simple closed paths* γ *of length 4*

A path can be visualized as being traced by lattice points in a hyper-rectangle of strategy profiles allowing motion only along the axes. Note that a simple closed path of length 4 always involves only two players. Even for congestion games, the theorem is nontrivial as it requires the sum of benefits of each deviating player to be zero. This theorem serves as a useful tool to disprove the existence of a potential. For example, one could easily construct examples of non-separable congestion games that violate (c).

The discussion so far is closely related to the concept of potential used in physics. In finding the equilibrium of a system of bodies, a series of equations can be solved using the free body diagram for each body or an expression for the energy of the system (Lagrangian) can be constructed and minimized. Also, uniqueness of exact potential up to an additive constant is similar to the fact that potential energy of a body is a function of the reference level. Furthermore, Theorem 2.1 corresponds to the idea that a body does not gain/loose potential energy when taken along a closed path.

2.2 Congestion games

2.2.1 Atomic congestion games

We first prove that an atomic-congestion game is a potential game using condition (c) of Theorem 2.1. Monderer and Shapley [12] further show that every finite potential game is isomorphic to a congestion game i.e., one can construct a fictitious network with appropriate delay functions. Atomic congestion games are ones in which each traveler controls an indivisible unit of flow. The traffic assignment problem is treated as an n-person non-cooperative game (where $n = \sum_{z \in Z} d_z \in \mathbb{Z}$) in which pure strategies are the paths available to each traveler. The utilities may be defined as the negative of the travel time incurred on a path or simply as the travel time but with a max operator replaced with a min in the definitions of best responses and NE. We will follow the later approach. Equilibrium is defined as a state in which no traveler can decrease his/her total path travel time by unilaterally switching to another path.

Theorem 2.2. *Atomic congestion game is a potential game*

Proof. Without loss of generality, consider two players i and j. In order to create a simple closed path of length 4, we assume that player i shifts from path p_{i_1} to p_{i_2} and then player j shifts from p_{j_1} to p_{j_2} and so on (see Figure 2.1).

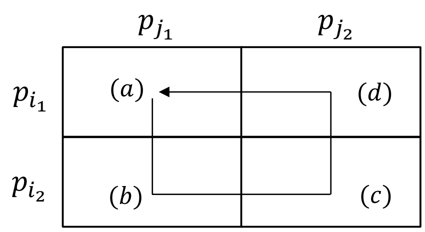

Figure 2.1: A closed cycle of length 4

We now verify if the differences in utilities for each deviating player sum to zero. This requires a careful examination of the changes to the link flows. Let the Venn diagram in Figure 2.2 represent the set of links that are common to the 4 paths under consideration.

Assume that $\{e_k\}_{k=1}^{15}$ represents a partition of the set of links that belong to the four paths. For example e_5, e_6, e_7 and e_8 represents the set of links that are common to paths p_{i_1} and p_{i_2}. We use a slightly different notation here for ease of explanation. Let $t_{e_k}(x_{e_k})$ represent the total travel time on the links in the set e_k, i.e., $t_{e_k}(x_{e_k}) = \sum_{e \in e_k} t_e(x_e)$, where x_{e_k} is a vector of number of players on each of the links in e_k. Also let $t_{e_k}(x_{e_k} + \delta)$ denote the sum of the travel times on all links in e_k after augmenting the flow on each link by δ units.

Figures 2.2(b),(c) and (d) represent the increase or decrease in the number of players on each set of links with respect to the base case in which players are playing (p_{i_1}, p_{j_1}) (Figure 2.2. The difference in utility for player i after deviating from p_{i_1} to p_{i_2} is given by Equation 2.1

$$t_{e_9}(x_{e_9}+1)+t_{e_{10}}(x_{e_{10}}+1)+t_{e_{11}}(x_{e_{11}}+1)+t_{e_{12}}(x_{e_{12}}+1)-t_{e_1}(x_{e_1})-t_{e_2}(x_{e_2})-t_{e_3}(x_{e_3})-t_{e_4}(x_{e_4}) \tag{2.1}$$

Similarly the benefits to the deviating player on the other three segments of the path are given by

$$t_{e_4}(x_{e_4})+t_{e_8}(x_{e_8}+1)+t_{e_{12}}(x_{e_{12}}+2)+t_{e_{15}}(x_{e_{15}}+1)-t_{e_2}(x_{e_2}-1)-t_{e_6}(x_{e_6})-t_{e_{10}}(x_{e_{10}}+1)-t_{e_{13}}(x_{e_{13}}) \tag{2.2}$$

$$t_{e_1}(x_{e_1})+t_{e_2}(x_{e_2}-1)+t_{e_3}(x_{e_3})+t_{e_4}(x_{e_4}+1)-t_{e_9}(x_{e_9}+1)-t_{e_{10}}(x_{e_{10}})-t_{e_{11}}(x_{e_{11}}+1)-t_{e_{12}}(x_{e_{12}}+2) \tag{2.3}$$

$$t_{e_2}(x_{e_2})+t_{e_6}(x_{e_6})+t_{e_{10}}(x_{e_{10}})+t_{e_{13}}(x_{e_{13}})-t_{e_4}(x_{e_4}+1)-t_{e_8}(x_{e_8}+1)-t_{e_{12}}(x_{e_{12}}+1)-t_{e_{15}}(x_{e_{15}}+1) \tag{2.4}$$

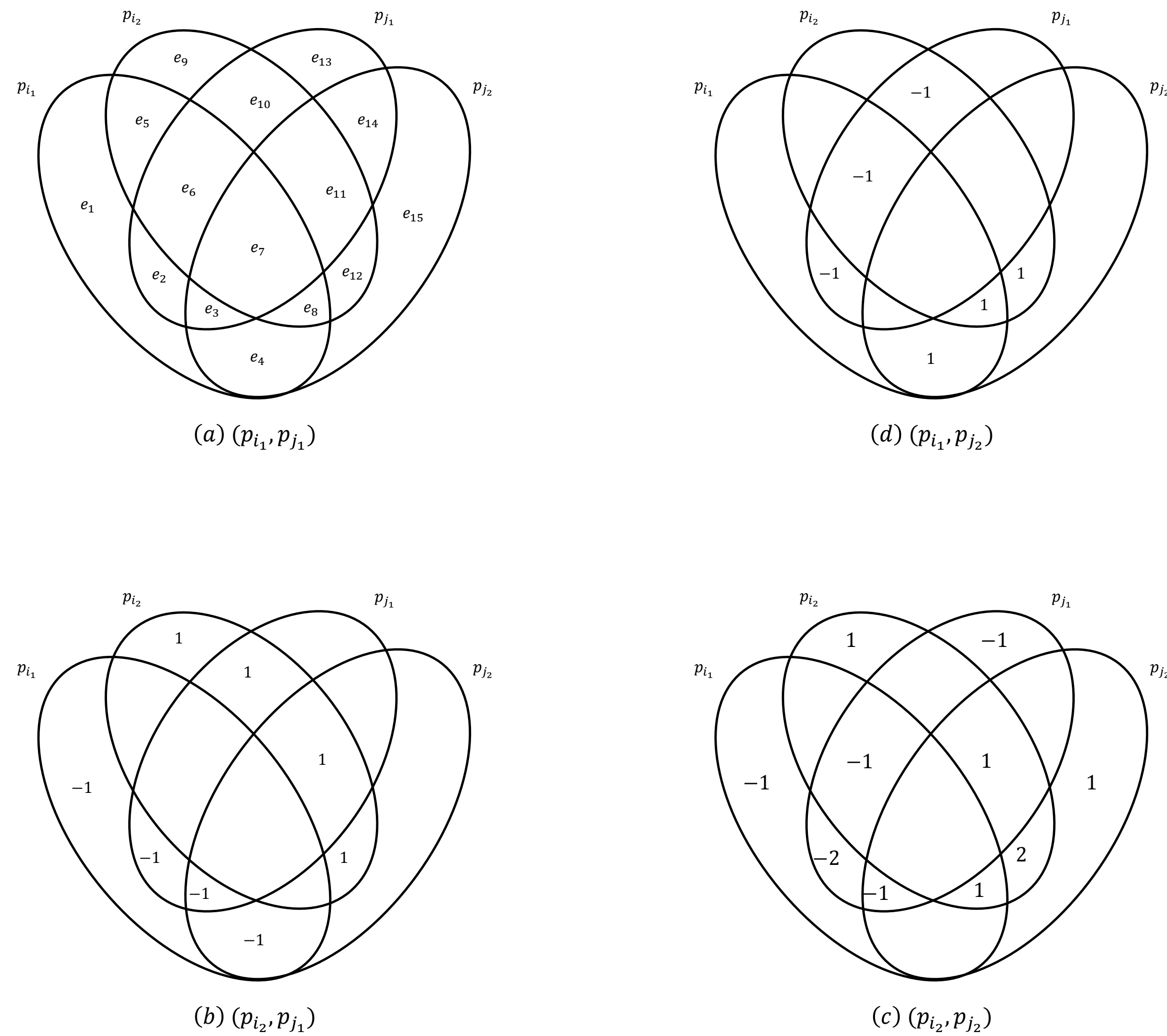

Figure 2.2: Change in the link flows in a simple closed cycle of length 4

Adding Equations 2.1-2.4, we observe that $I(\gamma, u) = 0$. Since the choice of players and their paths were arbitrary, the above result holds for all simple cycles of length 4. Hence, using Theorem 2.1, the atomic congestion game is a potential game.

■

Notice that separability plays a vital role in the above proof and without this assumption $I(\gamma, u)$ is not necessarily zero. Also note that no other conditions on the travel time functions (such as monotonicity) are required.

Let us now formally define atomic congestion games. As before, let E represent the set of links. Also assume that the origin and destination node for traveler i is denoted by v_i^o and

v_i^d respectively. The set of strategies of a player is the set of routes available to him/her, i.e., $A_i = P_{(v_i^o, v_i^d)}$. Let $x_e(a)$ be the number of users on link $e \in E$, for a given action profile a, i.e., $x_e(a) = |\{i \in N : e \in a_i\}|$. The payoff for player i, $u_i(a) = \sum_{e \in a_i} t_e(x_e(a))$. The potential can be defined as $\pi(a) = \sum_{e \in \cup_i a_i} \sum_{k=1}^{x_e(a)} t_e(k)$.

Suppose x_{ie} is 1 if player i chooses a path with arc e and is 0 otherwise. Let for a node $v \in V$, the set of outgoing and incoming arcs be denoted by $E^+(v)$ and $E^-(v)$ respectively. Rosenthal [18] proved that an integer-version of the Beckmann formulation yields a pure strategy NE to the atomic congestion game.

Theorem 2.3 (Rosenthal [18])**.** *In games derived from network equilibrium models pure-strategy NE always exist. Furthermore, any solution to the following problem is a pure-strategy NE*

$$\min \sum_e \sum_{k=0}^{x_e} t_e(k) \tag{2.5}$$

$$\text{s.t.} \quad x_e = \sum_i x_{ie} \qquad \forall\, e \in E \tag{2.6}$$

$$\sum_{e \in E^+(v)} x_{ie} - \sum_{e \in E^-(v)} x_{ie} = \begin{cases} -1 & \forall\, i \in N, v = v_i^d \\ 0 & \forall\, i \in N, v \in V \backslash \{v_i^o, v_i^d\} \\ 1 & \forall\, i \in N, v = v_i^o \end{cases} \tag{2.7}$$

$$x_{ie} \in \{0, 1\} \qquad \forall\, i \in N, e \in E \tag{2.8}$$

Every pure strategy NE does not necessarily solve the above optimization problem, i.e. multiple pure strategy NE may exist at least one of which may be discovered by solving the above problem(see Rosenthal [18] for example).

2.2.2 Non-atomic congestion games

In the non-atomic version, OD pairs are the players and the action space player of i is an assignment of the OD demand d_i to the set of paths (P_i) between the OD pair i. As before, suppose f_p denotes the flow on path $p \in P$. The action space is infinite and is defined by a set of vectors of demand assignments $A_i = \{(f_p)_{p \in P_i} : \sum_{p \in P_i} f_p = d_i, f_p \geq 0\, \forall\, p \in P_i\}$. The payoff of player i is defined as

$$u_i = \sum_{e \in P_i} \int_0^{x_e} t_e(x)\, dx$$

There are subtle differences between the NE of this game and the Wardrop principle for UE. Recall that Wardrop principle/User Equilibrium (UE) may be defined as follows: For each $z \in Z$, there exist no $p, q \in P_z$ with $f_p, f_q > 0$ such that $\sum_{e \in p} t_e(x_e) > \sum_{e \in q} t_e(x_e)$. The following theorem was used to establish that every NE to the proposed game satisfies the Wardrop UE principle.

Theorem 2.4 (Devarajan [8]). *If, for a flow pattern satisfying* $\sum_{p \in P_i} f_p = d_i \, \forall \, i \in N$, *there exist for some* $z \in Z$, *paths* $p, q \in P_z$ *such that* $f_p, f_q > 0$ *and* $\sum_{e \in p} t_e(x_e) > \sum_{e \in q} t_e(x_e)$, *then* u_i *is lowered by transferring some flow* Δf *from* p *to* q.

The above theorem states that given a flow pattern that is *not* a UE, u_i can be lowered by shifting some flow. Hence using proof by contrapositive, this establishes that a flow pattern in which u_i cannot be lowered (i.e., one which is a NE) satisfies Wardrop's UE principle. However, the vice-versa need not be true.

Devarajan [8] points out that the conditions for the Wardrop's UE principle is not *strong* enough to guarantee a NE to the proposed game. Given a pure strategy, the UE principle requires that u_i does not change when flows are shifted between any pair in $(f_p)_{p \in P_i}$, however to get to a pure strategy NE, we can shift different amounts of flows between multiple paths at the same time. This process cannot be equated to a sequence of pair wise shifts as equilibrium may be disturbed after the first shift. The author further goes to show that, in the presence of strictly monotone link performance functions, a solution to the Beckmann formulation is a pure strategy NE to the proposed non-atomic congestion game.

2.3 Equilibrium Refinements

2.3.1 Potential as a refinement tool

The argmax set of the potential function (note that this set is the same for all potential functions that differ by an additive constant) gives a subset of the set of equilibrium strategy profiles. Hence, the concept of a potential can be used as a refinement tool. Monderer and Shapley [12] cite a few behavioral experiments in which players actually chose solutions that optimized the potential function, but they do not rule out the possibility of this being purely coincidental. While on the outset it is in-evident why one would expect players to play a potential maximizing outcome, we will see in Chapter 3 that several learning dynamics converge to potential maximizing solutions.

2.3.2 Trembling Hand Perfect equilibrium

As the set of the NE is usually not a singleton, it is desirable to impose appropriate behavioral assumptions to obtain equilibria that are more likely. The trembling-hand perfect (THP) equilibrium (Selten [19]) or simply referred to as perfect equilibrium is one such solution concept which refines the NE set. Consider the following 2-player matrix game to illustrate the concept of perfect equilibria.

Table 2.2: Trembling-hand perfect equilibrium

	L	R
T	$1,1$	$0,0$
B	$0,0$	$0,0$

Both (T, L) and (B, R) are the Nash equilibria of the above game. However, (B, R) is less likely to be an outcome of this game as it is optimal for the row player to switch to T if he/she believes that the column player may "tremble" with a small probability and choose L. For similar reasons, the column player also prefers to switch to L. Thus, (T, L) is the only perfect equilibrium of this game. THP equilibrium may be formally defined using the following definitions.

Definition 2.7. A subset of mixed strategies for each player S_i° is said to be totally mixed if $\forall\, s_i \in S_i^\circ, s_i(a_i) > 0 \,\forall\, a_i \in A_i$

Definition 2.8. A totally mixed strategy $s \in \times_{i \in N} S_i^\circ$ is said to be an ϵ-perfect equilibrium if it satisfies the following condition: $u_i(a_i, s_{-i}) < u_i(a'_i, s_{-i}) \Rightarrow s_i(a_i) \leq \epsilon$, $\forall\, i \in N, a_i, a'_i \in A_i$

The above definition implies that while all strategies are played with a positive probability, a higher probability is assigned only to the best response strategies. A perfect equilibrium solution is then defined as the limit of such ϵ-perfect equilibria.

Definition 2.9. $s \in S$ is a perfect equilibrium $\Leftrightarrow \exists$ sequences $\{\epsilon_k\}_{k=1}^{\infty}$ and $\{s^k\}_{k=1}^{\infty}$ such that

(a) $\epsilon_k > 0 \,\forall\, k$ and $\lim_{k\to\infty} \epsilon_k = 0$

(b) s^k is an ϵ_k-perfect equilibrium

(c) $\lim_{k\to\infty} s_i^k(a_i) = s_i(a_i)$, $\forall\, a_i \in A_i$

The definitions above are due to Myerson [14]. Selten's original definitions were defined for extensive form games and its agent normal form (in which each information set of an extensive form game is modeled as a player) and are slightly different but equivalent. The equivalence of the two definitions can be found in van Damme [23].

2.3.3 Proper equilibrium

Now consider the game (suggested by Myerson [14]) shown in Table 2.3, which is obtained by adding a strictly dominated strategy for each player.

Table 2.3: Proper equilibrium

	L	C	R
T	$1, 1$	$0, 0$	$-9, -9$
M	$0, 0$	$0, 0$	$-7, -7$
R	$-9, -9$	$-7, -7$	$-7, -7$

In this game it can be shown that both (T, L) and (M, C) are perfect, However, assuming that players do not play strictly dominated strategies this game is no different from the previous one. Hence, (M, C) is a less reasonable equilibrium. To address this issue, Myerson [14] proposed the concept of proper equilibrium in which when players tremble, they do so by assigning more probability to better strategies and the highest probability to the best ones. Although we do not discuss proper equilibria of congestion games in detail, we will point some similarities between some learning strategies and its potential use in Chapter 4.

Definition 2.10. A totally mixed strategy $s \in S^{\circ}$ is said to be an ϵ-proper equilibrium if it satisfies the following condition: $u_i(a_i, s_{-i}) < u_i(a_i', s_{-i}) \Rightarrow \frac{1}{\epsilon} s_i(a_i) < s_i(a_i'), \forall\, i \in N, a_i, a_i' \in A_i$.

Like in the case of perfect equilibria, a proper equilibrium is defined as the limit of a sequence of ϵ-proper equilibria.

Definition 2.11. $s \in S$ is a proper equilibrium $\Leftrightarrow \exists$ sequences $\{\epsilon_k\}_{k=1}^{\infty}$ and $\{s^k\}_{k=1}^{\infty}$ such that

(a) $\epsilon_k > 0 \,\forall\, k$ and $\lim_{k\to\infty} \epsilon_k = 0$

(b) s^k is an ϵ_k-proper equilibrium

(c) $\lim_{k\to\infty} s_i^k(a_i) = s_i(a_i), \forall\, a_i \in A_i$

Using this definition one can rule out (M, C) in the game in Table 2.3. Several other equilibrium refinements have been proposed in game theory literature. Although some of them are more effective in further refining the equilibrium set, their existence in general is not guaranteed and hence haven't been discussed here. Both perfect and proper equilibria have been shown to exist for normal form games with finite action spaces. We point the reader to van Damme [23] for a discussion on some other equilibrium refinements.

Chapter 3

Learning in games

Learning is an approach in game theory which views Nash equilibria as steady states of a dynamic process. Normal form games or single-shot games are modeled as repeated games in which players respond according to some dynamic or strategy. In this chapter, we explore three common learning dynamics that exhibit some interesting properties in the context of potential games. We also discuss some similarities between these learning approaches and some models in transportation literature.

For the purpose of illustration, we will revisit the following example in each of the learning processes discussed in this chapter. Suppose two travelers wish to travel from node O to D in the network shown in Figure 3.1. The link performance functions are indicated on the arcs.

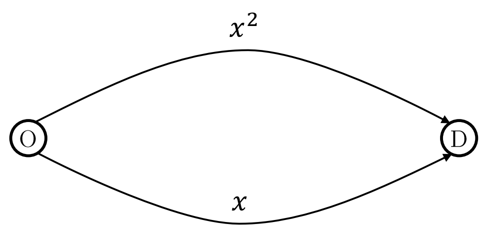

Figure 3.1: Example to demonstrate learning in networks

The problem of routing the travelers can be modeled as a 2-player game with payoff matrix as shown in Table 3.1. Let i and j denote the row and column player respectively. Each player has two paths to choose from: top (T) and bottom (B). The game has two pure strategy equilibria ((T, B) and (B, T)), and a mixed strategy NE in which each player chooses T and B with probability 1/4 and 3/4 respectively. This particular game is also an example of a coordination game in the sense that pure strategy NE are ones in which players do not choose the same action.

Table 3.1: Payoff matrix to demonstrate learning

	T	B
T	4,4	1,1
B	1,1	2,2

3.1 Finite Improvement Property

Proposed by Monderer and Shapley [12], finite improvement property is the simplest dynamic which is guaranteed to converge to a potential minimizing solution.

Definition 3.1. A path $\gamma = (a^1, a^2, \ldots)$ is an *improvement path* if $\forall\, k \geq 2, u_i(a^k) < u_i(a^{k-1})$, where $i \in N$ is the deviator at step k (i.e., the deviator is required to be strictly better off). If every improvement path generated by such myopic players is finite, then we say Γ has the *finite improvement property*(FIP).

The potential function, when minimized gives the set of pure-strategy NE ($a^* \in A$) for each player because $u_i(a_i^*, a_{-i}^*) < u_i(a_i, a_{-i}^*) \Leftrightarrow \pi(a_i^*, a_{-i}^*) < \pi(a_i, a_{-i}^*)\ \forall\, a_i \in A_i, i \in N$. Also, since A has a finite number of elements, the argmin set of the function π is non-empty and hence every finite potential game does have a pure-strategy NE.

Theorem 3.1 (Monderer and Shapley [12])**.** *Every finite potential game has the FIP*

For every improvement path γ, if the deviator benefits, the potential function value improves, i.e., $\pi(a^1) > \pi(a^2) > \pi(a^3) > \ldots$. As A is finite, the sequence γ has to be finite. Since atomic-congestion games are finite potential games, the FIP dynamic converges to an equilibrium solution.

Remark. FIP holds for a broader class of potential games called ordinal potential games. A game Γ is an *ordinal potential game* if it admits an *ordinal potential* function $\pi : A \to \mathbb{R}$ such that for every player i and $\forall\, a_{-i} \in A_{-i}\ u_i(a_i, a_{-i}) - u_i(a_i', a_{-i}) > 0 \Leftrightarrow \pi(a_i, a_{-i}) - \pi(a_i', a_{-i}) > 0,\ \forall\, a_i, a_i' \in A_i$.

Example. Consider the game described at the beginning of this chapter. Assume that the players begin by playing (B, B). Suppose player i deviates and chooses T. The path $((B, B), (T, B))$ is an improvement path as $1 < 2$. The new action profile (T, B) is a NE and no further improvement in any player's payoff is possible. Hence, the improvement path is finite. It can be easily verified that all other improvement paths are finite. Note that in general, improvement paths may cycle indefinitely (e.g., matching pennies).

3.2 Fictitious Play

Fictitious play is a learning dynamic that converges to a NE for certain classes of games. NE are steady states of a "fictitious" process involving repeatedly play of the original game in which, players best respond to their beliefs about the opponent's play (which are governed by the frequency of past actions) using a pure strategy. In other words, at each stage, every player picks a strategy that yields the maximum expected utility under the assumed mixture of opponents (which is based on past play). Fictitious play was originally proposed by Brown [5] for two-player zero-sum games and the convergence results were formally proved by Robinson [17]. While Brown [5] assumed that in each round of play, players take turns in an alternating manner and updates beliefs accordingly, Robinson supposed that beliefs are updated simultaneously. However, the convergence results remains valid in both cases but the rate of convergence may differ.

Monderer and Shapley [13] showed that every finite weighted potential game possess the fictitious play property (FPP). The n-player version of fictitious play requires players to myopically best respond to the empirical joint distribution of other players' actions. The fictitious play process is defined using a *belief path*, which is a sequence of vectors, each of which is a collection of mixed strategies, and the i^{th} component reflects others beliefs about player i's mixed strategy.

The process converges to equilibrium if this sequence of beliefs gets closer and closer to the actual set of equilibria. In other words, for every $\epsilon > 0$, the beliefs are in ϵ-equilibrium after a sufficient number of stages. A game is said to possess FPP if every fictitious play process converges in beliefs to equilibrium.

Let S^* and S^*_ϵ represent the set of mixed strategy equilibria and ϵ-equilibrium profiles of Γ. Suppose $\| \ \|$ be any Euclidean norm on S. For $\delta > 0$, let $B_\delta(S^*) = \{s \in S : \min_{s^* \in S^*} \|s - s^*\| < \delta\}$ represent a δ-ball around the set of mixed strategies. Consider a path in A, which is a sequence$\{a^t\}_{t=1}^{\infty}$ of elements of A.

Definition 3.2. A *belief path* is a sequence $\{s^t\}_{t=1}^{\infty}$ in S and is said to converge to equilibrium iff

(a) Every limit point of the sequence is an equilibrium.

(b) For every $\delta > 0 \, \exists \, T \in \mathbb{N} : s^t \in B_\delta(S^*), \, \forall \, t \geq T$.

(c) For every $\epsilon > 0 \, \exists \, T \in \mathbb{N}$ such that s^t is an ϵ-equilibrium for all $t \geq T$.

Given a path $\{a^t\}_{t=1}^{\infty}$, the belief path $\{s^t\}_{t=1}^{\infty}\,\forall\, t \geq 1$ and $i \in N$ is defined as

$$s_i^t(a_i) = \frac{1}{t}\sum_{k=1}^{t} \mathbf{1}_{\{a_i^k = a_i\}} \tag{3.1}$$

where $\mathbf{1}_{\{.\}}$ represents an indicator function. Note that

$$\begin{aligned} s_i^{t+1}(a_i) &= s_i^t(a_i) + \frac{1}{t+1}\left[\mathbf{1}_{\{a_i^{t+1}=a_i\}} - s_i^t(a_i)\right] \\ &= \frac{t}{t+1} s_i^t(a_i) + \frac{1}{t+1}\mathbf{1}_{\{a_i^{t+1}=a_i\}} \end{aligned} \tag{3.2}$$

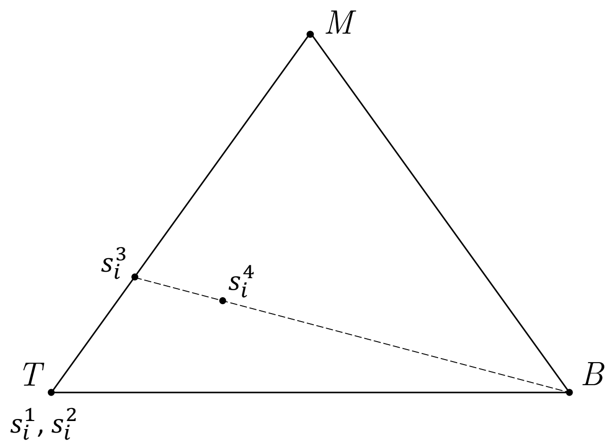

Figure 3.2: Mixed strategy space of Player i

For instance let player i have T, M and B in his/her strategy space and let $\{a_i^t\}$ be T, T, M, B, and so on. Then $\{s_i^t\}$ according to the above definition is $(1,0,0), (1,0,0), (\frac{2}{3}, \frac{1}{3}, 0), (\frac{1}{2}, \frac{1}{4}, \frac{1}{4})$, and so on(see Figure 3.2). Observe that the formula bears close resemblance with the link flow update process in method of successive averages (MSA)based algorithms (see Sheffi and Powell [20]).

Definition 3.3. $\{a^t\}_{t=1}^{\infty}$ is a fictitious play process if $\forall\, i \in N, t \geq 1, a_i^{t+1} \in BR_i(s^t)$, where $a_i^{t+1} \in A_i$. We say $\{a^t\}_{t=1}^{\infty}$ *converges in beliefs* to equilibrium if the associated belief path converges to equilibrium.

According to Definition 3.3 each player chooses an action at stage $(t+1)$ that is a best response to their beliefs at stage t, after which all players update their beliefs using Equation

3.1. Note, however, that at every stage each player picks a pure strategy. Theorem 3.2 establishes that congestion games have FPP.

Theorem 3.2 (Monderer and Shapley [13])**.** *Every game with identical payoff functions has the fictitious play property.*

The actual equilibrium solution to which fictitious play process converges depends on factors such as initial beliefs, tie-breaking rules and the manner in which beliefs are updated (i.e., sequential or simultaneous). However, Theorem 3.2 holds irrespective of a specific choice of the above mentioned factors. We will demonstrate this feature suing different tie-breaking rules using the following example.

Example. Let us revisit the example in Figure 3.1 to study the convergence of fictitious play. Assume that the action profile (T, T) is being played in the first round of play. Table 3.2 summarizes the results of the fictitious play process. Ties were broken in favor of the top path (T) for both players. As can be seen from the table, the process does not converge to a pure strategy but converges in beliefs to the mixed strategy NE of the game (i.e., $s_i^\infty(T) = s_j^\infty(T) = 0.25$).

Table 3.2: Fictitious play with a tie-breaking rule

t	*Player i*					*Player j*				
	$u_i(T, s_j^t)$	$u_i(B, s_j^t)$	a_i^t	$s_i^t(T)$	$s_i^t(B)$	$u_j(T, s_i^t)$	$u_j(B, s_i^t)$	a_j^t	$s_j^t(T)$	$s_j^t(B)$
1	4.000	1.000	T	1.000	0.000	4.000	1.000	T	1.000	0.000
2	2.500	1.500	B	0.500	0.500	2.500	1.500	B	0.500	0.500
3	2.000	1.667	B	0.333	0.667	2.000	1.667	B	0.333	0.667
4	1.750	1.750	B	0.250	0.750	1.750	1.750	B	0.250	0.750
5	2.200	1.600	T	0.400	0.600	2.200	1.600	T	0.400	0.600
6	2.000	1.667	B	0.333	0.667	2.000	1.667	B	0.333	0.667
7	1.857	1.714	B	0.286	0.714	1.857	1.714	B	0.286	0.714
8	1.750	1.750	B	0.250	0.750	1.750	1.750	B	0.250	0.750
9	2.000	1.667	T	0.333	0.667	2.000	1.667	T	0.333	0.667
10	1.900	1.700	B	0.300	0.700	1.900	1.700	B	0.300	0.700
11	1.818	1.727	B	0.273	0.727	1.818	1.727	B	0.273	0.727
12	1.750	1.750	B	0.250	0.750	1.750	1.750	B	0.250	0.750
13	1.923	1.692	T	0.308	0.692	1.923	1.692	T	0.308	0.692
14	1.857	1.714	B	0.286	0.714	1.857	1.714	B	0.286	0.714
15	1.800	1.733	B	0.267	0.733	1.800	1.733	B	0.267	0.733
16	1.750	1.750	B	0.250	0.750	1.750	1.750	B	0.250	0.750
17	1.882	1.706	T	0.294	0.706	1.882	1.706	T	0.294	0.706
18	1.833	1.722	B	0.278	0.722	1.833	1.722	B	0.278	0.722
⋮	⋮	⋮	⋮	⋮	⋮	⋮	⋮	⋮	⋮	⋮
125	1.768	1.744	T	0.256	0.744	1.768	1.744	T	0.256	0.744

Now consider a variant of the fictitious play process in which ties are arbitrarily resolved, an instance of which is shown in Table 3.3. Unlike the previous example, this process was found to converge to a pure strategy NE (B, T).

Table 3.3: Fictitious play process with arbitrary resolution of ties

t	*Player i*					*Player j*				
	$u_i(T, s_j^t)$	$u_i(B, s_j^t)$	a_i^t	$s_i^t(T)$	$s_i^t(B)$	$u_j(T, s_i^t)$	$u_j(B, s_i^t)$	a_j^t	$s_j^t(T)$	$s_j^t(B)$
1	4.000	1.000	T	1.000	0.000	4.000	1.000	T	1.000	0.000
2	2.500	1.500	B	0.500	0.500	2.500	1.500	B	0.500	0.500
3	2.000	1.667	B	0.333	0.667	2.000	1.667	B	0.333	0.667
4	1.750	1.750	B	0.250	0.750	1.750	1.750	B	0.250	0.750
5	2.200	1.600	B	0.200	0.800	1.600	1.800	T	0.400	0.600
6	2.500	1.500	B	0.167	0.833	1.500	1.833	T	0.500	0.500
7	2.714	1.429	B	0.143	0.857	1.429	1.857	T	0.571	0.429
8	2.875	1.375	B	0.125	0.875	1.375	1.875	T	0.625	0.375
9	3.000	1.333	B	0.111	0.889	1.333	1.889	T	0.667	0.333
10	3.100	1.300	B	0.100	0.900	1.300	1.900	T	0.700	0.300
11	3.182	1.273	B	0.091	0.909	1.273	1.909	T	0.727	0.273
12	3.250	1.250	B	0.083	0.917	1.250	1.917	T	0.750	0.250
13	3.308	1.231	B	0.077	0.923	1.231	1.923	T	0.769	0.231
14	3.357	1.214	B	0.071	0.929	1.214	1.929	T	0.786	0.214
15	3.400	1.200	B	0.067	0.933	1.200	1.933	T	0.800	0.200
16	3.438	1.188	B	0.063	0.938	1.188	1.938	T	0.813	0.188
17	3.471	1.176	B	0.059	0.941	1.176	1.941	T	0.824	0.176
18	3.500	1.167	B	0.056	0.944	1.167	1.944	T	0.833	0.167
⋮	⋮	⋮	⋮	⋮	⋮	⋮	⋮	⋮	⋮	⋮
125	3.928	1.024	B	0.008	0.992	1.024	1.992	T	0.976	0.024

3.3 Logit-response models

Fictitious play process is also called a best response mechanism as at each stage, players are assumed to make perfectly rational choices given their beliefs about their opponents. A more realistic class of learning models (partly inspired from evolutionary game theory) relaxes this assumption of perfect rationality. Like in fictitious play, players engage in repeated play of a stage game but may now err with some probability. The central idea in these models is to define a stochastic process using action profiles as system states, and assume some dynamic which lets players move from one state to another. This process is modeled as a Markov chain and its stationary or steady state distribution is used to study equilibrium solutions. Literature surveyed in this section is primarily from Blume [4], Young [25], and Kandori et al. [10]

Players in these learning models are modeled using the following three components:

1. *Inertia*: While fictitious play may require travelers to change routes in each round of play, in reality, travelers are likely to switch routes rather infrequently. Hence, players are assumed to possess inertia in the context of route switching and shift route only when an opportunity to change routes is presented. Such an opportunity is called a strategy revision opportunity. For the time duration between two such successive opportunities, players are assumed to continue playing their chosen action.

2. *Myopic behavior*: Utilities in most repeated games in game theory literature are modeled using discounted costs. However in transportation networks, travelers are focused on minimizing their current travel time and long run implication of their choices are not considered in the decision making process. Hence, players are modeled to behave myopically and try to choose actions that optimize their present payoffs.

3. *Mutations/Noise or Trembles*: Players are assumed to tremble or make mistakes while choosing an action. Depending on the probabilities that are assigned to the strategies that are not best responses, different learning algorithms can be constructed. While Young [25] and Kandori et al. [10] use a model in which all strategies which are not best responses are chosen with some ϵ probability, Blume [4] assumes a log-linear rule in selecting actions (and hence the name logit learning). The act of making a mistake is motivated in several ways. From an evolutionary game theoretic standpoint it is treated as mutations. Alternately, it could be that players lack the ability to evaluate actions and pick the optimal one. It could also be reasoned as a result of some noise/randomness in the model parameters (such as link performance functions). In fact, the latter interpretation of logit learning makes it almost identical to stochastic UE models (see Cascetta [6]). The link travel times in stochastic UE models are random which equates to a logit type expression for path choice probabilities under appropriate assumptions on the error terms. This approach in interpreting mistakes in learning models is similar to Harsayani's [9] purification theorem according to which mixed strategies are shown to be equivalent to pure strategy NE in games with incomplete information/noise.

Consider a set of n players, where $n \in \mathbb{Z}$. Logit learning process is modeled as a continuous time Markov chain (CTMC) $\{\mathcal{N}(t)\}_{t\geq 0}$. The states are simply the set of action profiles A. We assume that the sojourn times of each player i, τ_i, are exponentially distributed with parameter λ(assume that players are equipped with an independent exponential clock with rate λ). Since we deal with continuous distributions, no two players revise their strategies simultaneously.

When player i gets to choose (say at time t), he/she does so using a log-linear choice rule as described in Equation 3.3. Assume at time t all other players are playing the action profile a_{-i}^t. Then the probability with which player i chooses $a_i^t \in A_i$ is given by

$$\xi_i^\epsilon(a_i^t, a_{-i}^t) = \frac{\exp\left(-u_i(a_i^t, a_{-i}^t)/\epsilon\right)}{\sum_{a_i \in A_i} \exp\left(-u_i(a_i, a_{-i}^t)/\epsilon\right)} \tag{3.3}$$

Note that a key difference in logit learning and the previously discussed dynamics is that the time at which players get to revise their strategies is not discrete but a continuous variable. It has to be emphasized here that player i does not play a mixed strategy but chooses a single action a_i^t that is randomly drawn from the above distribution. The value of ϵ defines the extent of making a mistake or extent of irrationality. As ϵ tends to zero, players begin to choose best reposes with greater probabilities. Once a player i decides to play a particular action a_i^t, he/she sticks to it until presented with another strategy revision opportunity.

For any positive ϵ, the process $\{\mathcal{N}(t)\}_{t\geq 0}$ is irreducible and recurrent (since all states can communicate with each other). Hence, a unique steady state/limiting distribution that has all states in its support exists. However, as ϵ tends to zero, i.e., as the probability of making mistakes get smaller (it is assumed that by repeated interactions players get more experienced) only a few states have positive limiting probabilities. These states constitute what is termed a *stochastically stable set*.

Since the the rate at which a player gets to revise his/her strategy is λ and the probabilities with which they change their strategies are given by ξ's, the the transition rate from state a to a', $q_{aa'}$, is given by

$$q_{aa'} = \begin{cases} \lambda \xi_i^\epsilon(a_i, a_{-i}) & \text{if for some } i \in N,\ a_i \neq a'_i, a_{-i} = a'_{-i} \\ \sum_{i \in N} \lambda \xi_i^\epsilon(a_i, a_{-i}) & \text{if } a' = a \\ 0 & \text{otherwise} \end{cases} \tag{3.4}$$

For a given ϵ, let the long run or steady state probabilities of finding the Markov chain in state a be denoted by ρ_a^ϵ.

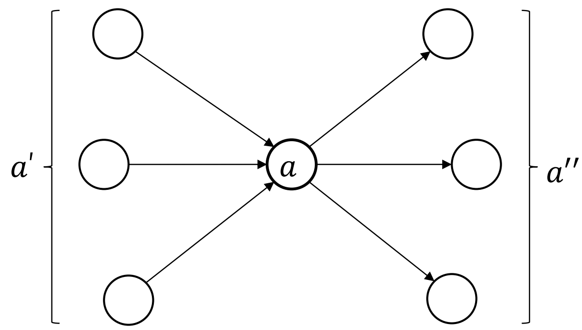

Figure 3.3: Global balance equations

The global balance equations for this CTMC may be written as follows:

$$\sum_{a'' \in A} \rho_a^\epsilon q_{aa''} = \sum_{a' \in A} \rho_{a'}^\epsilon q_{a'a} \, \forall \, a \in A \tag{3.5}$$

$$\sum_{a \in A} \rho_a^\epsilon = 1 \tag{3.6}$$

Blume showed that as ϵ's tend to zero, stochastically stable set coincides with the argmin set of the potential function.

Theorem 3.3 (Blume [4])**.** $\lim_{\epsilon \to 0} \sum_{a \in \operatorname{argmin} \pi(.)} \rho_a^\epsilon = 1$

Remark. Convergence of the logit learning model holds for any potential game and is not restricted to congestion games.

Example. We now demonstrate the logit learning model using the example presented in Figure 3.1. The probabilities of selecting an action for a given value of ϵ are given by

$$\xi_i^\epsilon(T,T) = \frac{\exp(-4/\epsilon)}{\exp(-4/\epsilon) + \exp(-1/\epsilon)} \qquad \xi_i^\epsilon(B,T) = \frac{\exp(-1/\epsilon)}{\exp(-4/\epsilon) + \exp(-1/\epsilon)} \tag{3.7}$$

$$\xi_i^\epsilon(T,B) = \frac{\exp(-1/\epsilon)}{\exp(-1/\epsilon) + \exp(-2/\epsilon)} \qquad \xi_i^\epsilon(B,B) = \frac{\exp(-2/\epsilon)}{\exp(-1/\epsilon) + \exp(-2/\epsilon)} \tag{3.8}$$

$$\xi_j^\epsilon(T,T) = \frac{\exp(-4/\epsilon)}{\exp(-4/\epsilon) + \exp(-1/\epsilon)} \qquad \xi_j^\epsilon(T,B) = \frac{\exp(-1/\epsilon)}{\exp(-4/\epsilon) + \exp(-1/\epsilon)} \tag{3.9}$$

$$\xi_j^\epsilon(B,T) = \frac{\exp(-1/\epsilon)}{\exp(-1/\epsilon) + \exp(-2/\epsilon)} \qquad \xi_j^\epsilon(B,B) = \frac{\exp(-2/\epsilon)}{\exp(-1/\epsilon) + \exp(-2/\epsilon)} \tag{3.10}$$

The CTMC has four states (T,T), (B,T), (T,B) and (B,B) which we call 1,2,3 and 4 respectively. The transition diagram and the associated transition rates are shown in Figure 3.4.

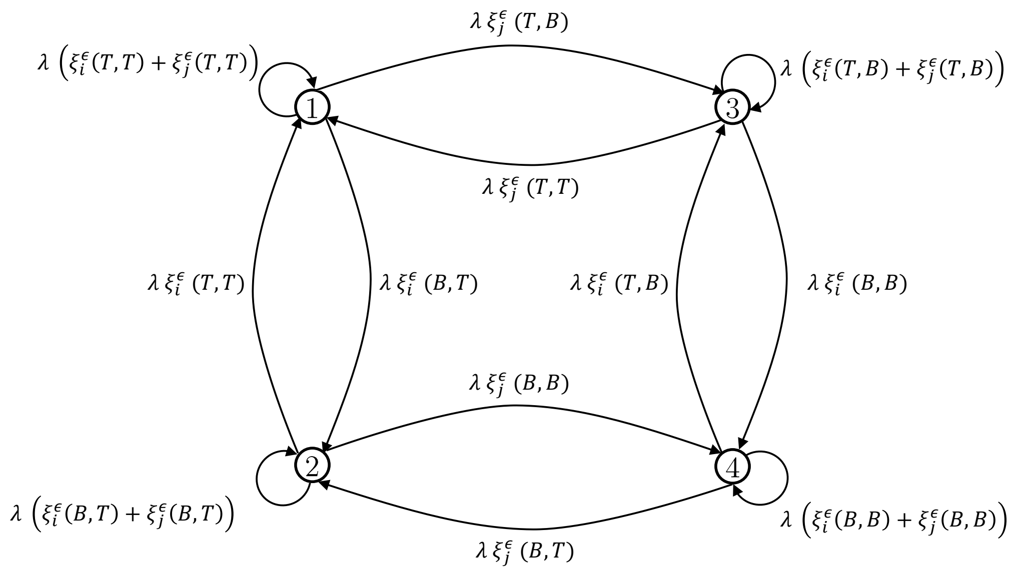

Figure 3.4: Transition diagram for a logit learning process

The global balance equations are then constructed and solved to obtain the long run proportion of time spent in each state.

$$\rho_1^\epsilon + \rho_2^\epsilon + \rho_3^\epsilon + \rho_4^\epsilon = 1 \tag{3.11}$$

$$\rho_1^\epsilon \left(\xi_j^\epsilon(T,B) + \xi_i^\epsilon(B,T)\right) = \rho_2^\epsilon \xi_i^\epsilon(T,T) + \rho_3^\epsilon \xi_j^\epsilon(T,T) \tag{3.12}$$

$$\rho_2^\epsilon \left(\xi_i^\epsilon(T,T) + \xi_j^\epsilon(B,B)\right) = \rho_1^\epsilon \xi_i^\epsilon(B,T) + \rho_4^\epsilon \xi_j^\epsilon(B,T) \tag{3.13}$$

$$\rho_3^\epsilon \left(\xi_j^\epsilon(T,T) + \xi_i^\epsilon(B,B)\right) = \rho_1^\epsilon \xi_j^\epsilon(T,B) + \rho_4^\epsilon \xi_i^\epsilon(T,B) \tag{3.14}$$

$$\rho_4^\epsilon \left(\xi_j^\epsilon(B,T) + \xi_i^\epsilon(T,B)\right) = \rho_2^\epsilon \xi_j^\epsilon(B,B) + \rho_3^\epsilon \xi_i^\epsilon(B,B) \tag{3.15}$$

Solving the balance equations for a given value of ϵ yields the steady state probabilities of finding the system in that state. As ϵtends to zero, the support of the steady state probabilities is identical to the potential minimizing action profiles. The two possible pure strategy NE for this game are the ones each route has one traveler. Table 3.4 summarizes the behavior of the steady state probabilities for different values of ϵ. As can be seen from the table, the steady state probability of each of the two pure strategy NE is 0.5 for low values of ϵ.

Table 3.4: Convergence of logit learning

ϵ	1	0.5	0.33	0.25	0.2	0.1	...	0.01
ρ_1^ϵ	0.020593	0.001159	6.02E-05	3.04E-06	1.52E-07	4.68E-14	...	2.6E-131
ρ_2^ϵ	0.413622	0.467768	0.487826	0.495461	0.498321	0.499989	...	0.5
ρ_3^ϵ	0.413622	0.467768	0.487826	0.495461	0.498321	0.499989	...	0.5
ρ_4^ϵ	0.152163	0.063305	0.024287	0.009075	0.003358	2.27E-05	...	1.86E-44

Chapter 4

Discussion

4.1 Applications of refinements and learning algorithms

As seen earlier, in games with multiple equilibria, one could use refinement techniques to eliminate equilibria which are not sensible. Here we illustrate an example of a congestion game with an equilibrium that is not perfect. Fictitious play process is then modified to guarantee convergence to a perfect equilibrium solution.

Consider a game proposed by Monderer and Shapley [12] in which two travelers are on an undirected rectangular block (with links numbered from 1 to 4 as shown in Figure 4.1). Let player i travel from top left to bottom right and player j travel from top right to bottom left.

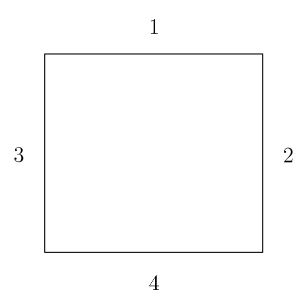

Figure 4.1: Network to demonstrate THP solutions

Suppose the delay functions of this game are $t_1(1) = 2, t_1(2) = 4, t_2(1) = 3, t_2(2) = 6, t_3(1) = 7, t_3(2) = 9, t_4(1) = 5$ and $t_4(2) = 6$. Then the utilities of the game can be written as shown in Table 4.1.

Table 4.1: Payoffs in a congestion game with perfect equilibria

	Path 1-3	Path 2-4
Path 1-2	$t_1(2)+t_2(1), t_1(2)+t_3(1)$	$t_2(2)+t_1(1), t_2(2)+t_4(1)$
Path 3-4	$t_3(2)+t_4(1), c_3(2)+t_1(1)$	$t_4(2)+t_3(1), t_4(2)+t_2(1)$

	Path 1-3	Path 2-4
Path 1-2	6,11	8,11
Path 3-4	14,11	13,9

A potential to this game is $\pi = [\begin{smallmatrix} 16 & 16 \\ 23 & 21 \end{smallmatrix}]$. This game has two pure strategy equilibria $((1-2, 1-3), (1-2, 2-4))$ and no unique mixed strategy equilibrium. However, only (1-2,2-4) is perfect (if players were playing (1-2,1-3), . As noted earlier, fictitious play in congestion games can lead to different equilibrium solutions based on choice of initial beliefs, starting solution etc. For instance if player i and j chose paths 1-2 and 1-3 respectively, they would end up playing the same strategy indefinitely for certain tie-breaking rules. However, if each player's initial beliefs are totally mixed, it can be established that a fictitious play process that converges to a pure strategy NE converges to a perfect equilibrium.

Theorem 4.1. *A fictitious play process that converges to a pure strategy NE, converges to a perfect equilibrium if the initial beliefs are totally mixed.*

Proof. Let the initial set of beliefs for each player $i \in N$, be $s_i^1 \in S_i^\circ$. For $t \geq 2$, assume players use Equation 3.1 to update their beliefs. Since the fictitious play process is assumed to converge to a pure strategy equilibrium, $\exists$ a $T \geq 2$ such that $\forall\, t \geq T$, $a^t = a^T$, where a^T is the equilibrium solution. Since s^1 is totally mixed, s^T is also totally mixed $\Rightarrow \exists\, \epsilon \geq 0$ such that s^T is an ϵ-perfect equilibrium. For each player $i \in N$, $t \geq T$ such that $a_i \neq a_i^T$ $s_i^{t+1}(a_i) = \frac{t}{t+1} s_i^t(a_i) + \frac{1}{t+1} \mathbf{1}_{\{a_i^{t+1} = a_i\}} = \frac{t}{t+1} s_i^t(a_i)$. Hence, we have $s_i^T(a_i) \leq \epsilon$, $s_i^{T+1}(a_i) \leq \frac{T}{T+1}\epsilon$, $s_i^{T+2}(a_i) \leq \frac{T}{T+2}\epsilon$ and so on. Let $\epsilon_k = \frac{\epsilon T}{T+k}$. Clearly, the sequence of ϵ's shrink to zero and $\lim_{t\to\infty} s^k = a^k = a^T$. Hence, the modified fictitious play process converges to a perfect equilibrium. ∎

Note the above example is a 2-player game and hence the set of perfect equilibria and those that are admissible by deletion of weakly dominated strategies are equivalent. But in general, trembling-hand perfectness is stricter than iterated admissibility. The role of reducing the set of plausible equilibria using the concept of perfectness may still be limited in congestion games. One might be able to restrict the set of equilibrium solutions to a greater degree if we can find a way to construct the set of proper equilibria. Although, the notion of proper equilibria resembles learning algorithms such as logit-response (which assign higher probabilities to better response strategies), they may not be ϵ-proper equilibria at any stage as the updates are carried out in an asynchronous manner. Unlike entropy maximization methods, refinements may help find equilibria that are more robust to perturbations and those that are sensitive to travel times. A potential relationship between different equilibrium refinements in congestion games is shown in the following figure.

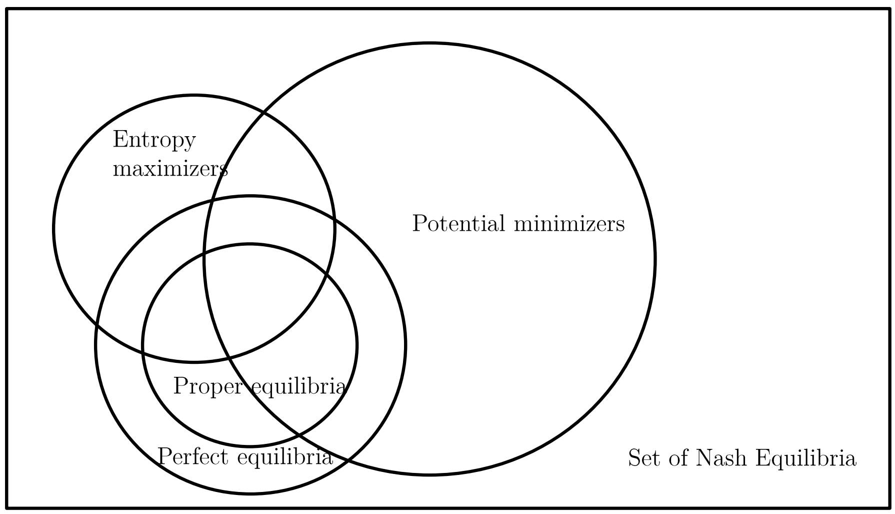

Figure 4.2: Relationship between different refinements

4.2 Limitations of game theoretic models

Although the computation of NE of games is generally difficult, properties of potential games makes it easier to analyze games with a large number of players. While learning algorithms are known to converge to the NE, the rate at which they converge may be slow. However, it is important to realize that the primary purpose of these algorithms is to understand how players might end up choosing the NE, but not to determine the NE solutions.

Another major issue with the use of learning algorithms is that they deal in the space of path flows, and hence becomes computationally intractable with increase in network size. Hence, efficient enumeration methods may be developed to eliminate the inclusion of dominated strategies or paths in players' action spaces. Also, as the convergence of most learning algorithms depends on the error structure, properties of models which extend existing learning methods require further research.

4.3 Conclusion

Transportation network modeling and game theoretic treatment of congestion games have a lot in common. In this report, we studied concepts from game theory that have been used to address the TAP. Two major issues were addressed: the multiple equilibrium solutions to the TAP and of converge to equilibrium with agents who are not necessarily rational and do not have full information. In particular we discussed the properties of potential games and associated learning algorithms such as fictitious play and logit learning models. As an aside, game theoretic approaches to distinguish multiple equilibria were discussed and possible applications of it in the context of congestion games were briefly illustrated. These equilibrium refinements incorporate certain behavioral aspects of decision making and are hence very different than conventional entropy maximization methods. Multiple equilibria in transportation networks and their likelihood of occurrence, and extensions to learning algorithms discussed in this report are interesting topics worth exploring.

Bibliography

[1] H. Bar-Gera. Primal method for determining the most likely route flows in large road networks. *Transportation Science*, 40(3):269–286, 2006.

[2] M. Beckmann, C. McGuire, and C. B. Winsten. Studies in the economics of transportation. Technical report, 1956.

[3] M. G. Bell and Y. Iida. *Transportation network analysis*. Wiley, New York, 1997.

[4] L. Blume. Population games. Game Theory and Information 9607001, EconWPA, July 1996.

[5] G. W. Brown. Iterative solution of games by fictitious play. In T. Koopmans, editor, *Activity Analysis of Production and Allocation*, pages 374–376. Wiley, New York, 1951.

[6] E. Cascetta. A stochastic process approach to the analysis of temporal dynamics in transportation networks. *Transportation Research Part B: Methodological*, 23(1):1 – 17, 1989.

[7] S. C. Dafermos. Traffic equilibrium and variational inequalities. *Transportation Science*, 14(1):42–54, 1980.

[8] S. Devarajan. A note of network equilibrium and noncooperative games. *Transportation Research Part B: Methodological*, 15(6):421 – 426, 1981.

[9] J. Harsanyi. Games with randomly disturbed payoffs: A new rationale for mixed-strategy equilibrium points. *International Journal of Game Theory*, 2(1):1–23, 1973.

[10] M. Kandori, G. J. Mailath, and R. Rob. Learning, mutation, and long run equilibria in games. *Econometrica*, 61(1):pp. 29–56, 1993.

[11] E. Kohlberg and J.-F. Mertens. On the strategic stability of equilibria. *Econometrica*, 54(5):pp. 1003–1037, 1986.

[12] D. Monderer and L. S. Shapley. Potential games. *Games and Economic Behavior*, 14 (1):124 – 143, 1996.

[13] D. Monderer and L. S. Shapley. Fictitious play property for games with identical interests. *Journal of Economic Theory*, 68(1):258 – 265, 1996.

[14] R. Myerson. Refinements of the nash equilibrium concept. *International Journal of Game Theory*, 7(2):73–80, 1978.

[15] J. Nash. Non-cooperative games. *Annals of Mathematics*, 54(2):pp. 286–295, 1951.

[16] S. Peeta and A. Ziliaskopoulos. Foundations of dynamic traffic assignment: The past, the present and the future. *Networks and Spatial Economics*, 1(3-4):233–265, 2001.

[17] J. Robinson. An iterative method of solving a game. *Annals of Mathematics*, 54(2):pp. 296–301, 1951.

[18] R. W. Rosenthal. The network equilibrium problem in integers. *Networks*, 3(1):53–59, 1973.

[19] R. Selten. Reexamination of the perfectness concept for equilibrium points in extensive games. *International Journal of Game Theory*, 4(1):25–55, 1975.

[20] Y. Sheffi and W. B. Powell. An algorithm for the equilibrium assignment problem with random link times. *Networks*, 12(2):191–207, 1982.

[21] M. J. Smith. The existence, uniqueness and stability of traffic equilibria. *Transportation Research Part B: Methodological*, 13(4):295–304, 1979.

[22] M. J. Smith. The stability of a dynamic model of traffic assignmentan application of a method of lyapunov. *Transportation Science*, 18(3):245–252, 1984.

[23] E. van Damme. *Refinements of the Nash Equilibrium Concept*, volume 219 of *Lecture Notes in Economics and Mathematical Systems*. Springer, 1983.

[24] J. Wardrop. Some theoretical aspects of road traffic research. *Proceedings of the Institution of Civil Engineers, Part II*, 1:352–362, 1952.

[25] H. P. Young. The evolution of conventions. *Econometrica*, 61(1):pp. 57–84, 1993.

[26] H. J. V. Zuylen and L. G. Willumsen. The most likely trip matrix estimated from traffic counts. *Transportation Research Part B: Methodological*, 14(3):281 – 293, 1980.