Coalitional game with opinion exchange

by

Bomin Jiang

Submitted to the Department of Mechanical Engineering
in partial fulfillment of the requirements for the degree of
Master of Science in Mechanical Engineering
at the
MASSACHUSETTS INSTITUTE OF TECHNOLOGY
June 2017

Author .... Signature redacted ....
Department of Mechanical Engineering
May 12, 2017

Certified by .... Signature redacted ....
Munther A. Dahleh
William A. Coolidge Professor of Electrical Engineering and Computer Science
Thesis Supervisor

Certified by .... Signature redacted ....
Mardavij Roozbehani
Principal Research Scientist, Laboratory for Information and Decision Systems
Thesis Supervisor

Certified by .... Signature redacted ....
Jean-Jacques Slotine
Professor of Mechanical Engineering
Thesis Supervisor

Accepted by .... Signature redacted
Rohan Abeyaratne
Chairman, Department Committee on Graduate Theses

Coalitional game with opinion exchange

by

Bomin Jiang

Submitted to the Department of Mechanical Engineering
on May 12, 2017, in partial fulfillment of the
requirements for the degree of
Master of Science in Mechanical Engineering

Abstract

Traditional coalitional game theory does not apply if different participants hold different opinions about the payoff function that corresponds to each subset of the coalition. In this thesis, we propose a framework in which players can exchange opinions about their views of payoff functions and then decide the distribution of the value of the grand coalition.

The main contributions of this thesis include:

1. This thesis proposes a coalitional game model with private payoff functions and opinion exchange. In current literature, no information exchange on private payoff functions has been considered.

2. When players are truth-telling, the problem of opinion consensus is decoupled from the coalitional game. However, interesting dynamics arise when players are strategic in the consensus phase. Assuming that all players are rational, we show that, if influential players are risk-averse, an efficient fusion of the distributed data is achieved at pure strategy Nash equilibrium, meaning that the average opinion will not drift as time goes on.

3. If the weighted average of private payoff functions is supermodular, then there exists a risk averse level such that the Bayesian core is non-empty.

4. Without the assumption that all players are rational, each player can use an algorithmic R-learning process, which gives the same result as the pure strategy Nash equilibrium with rational players.

Thesis Supervisor: Munther A. Dahleh
Title: William A. Coolidge Professor of Electrical Engineering and Computer Science

Thesis Supervisor: Mardavij Roozbehani
Title: Principal Research Scientist, Laboratory for Information and Decision Systems

Thesis Supervisor: Jean-Jacques Slotine
Title: Professor of Mechanical Engineering

Acknowledgments

This thesis would not have been possible without the support and assistance of many individuals and organizations. I would like to thank all those who have made this thesis possible.

First I would like to express my sincere gratitude to my main advisor, Prof. Munther Dahleh. He is a supervisor that adds up much value to his students. His comments and feedbacks are always very inspiring. Furthermore, he is a person that make things easier rather than more complicated; whenever we have different ideas on an issue, instead of forcing his opinion on me, he is always willing to have an open discussion.

Besides my main supervisor, I would like to express my gratitude to my co-advisor, Dr. Mardavij Roozbehani, who is always here to answer my questions when I need him the most. His experience in game theory and opinion dynamics ensures the completion of this project. Additionally, he is always willing to share his different opinions, which becomes a main driver of progress.

Furthermore, I would like to thank my fellow students. I really enjoy talking to you and got many inspiring ideas from our conversations.

Contents

List of Figures

Chapter 1

Introduction

1.1 Background and Literature

In recent years, the application of coalitional game theory in multi-agent systems has been receiving increasing attention. Those applications include task allocation [16], smart grids [14], transportation networks [13]. Although the theory of coalitional games has existed for a few decades, theory for the case of unrealized payoff functions (of subsets of players) is quite limited. In most of the literature that considers coalitional game theory, an oversimplified assumption is used, i.e., that all players agree on a common sub-coalition payoff function.

Recently, researchers have started to look at this case in a variety of ways, e.g., using the model of Bayesian games, bargaining games, or repeated playing dynamic games. One paper [11] derived a model that generalizes coalitional games to a Bayesian framework using types. Furthermore, a Bayesian core contract is defined as the set of contracts of payoff distributions that are non-blocking under the expected value of payoffs of players, whether Ex ante, Ex interim, or Ex post. Note that non-blocking means one player is better off staying in the grand coalition, so this player would not block the formation of this grand coalition. Similarly, another paper [10] defined the concept of Bayesian core (A core is a set of payoff distributions such that every player is better off staying in the grand coalition) under uncertainty and gave a bargaining algorithm that converges to the Bayesian core, assuming that it exists. However, there are two practical issues with the setting. First, the theory says nothing when such a core does not exist; second, even if it does exist, people's individual observations, which are private information, are not used constructively because they do not exchange information on private payoff functions. By exchanging information, everyone can obtain a better estimate of the ground truth of the payoff function. In addition, players may not follow the algorithm suggested in the literature when they are strategic and want the algorithm to converge to some value in the core that favors them. Finally, a fair distribution, such as the Shapley value in the classical coalitional game model, is not well defined because a commonly-accepted payoff function may not exist. Another paper, [5], used a repeated playing model and assumed that players learn the actual state of the world as the game goes on, but, in practice, states may never converge if the game that is being played is changing rapidly over time or, even worse, if the game is only played once.

1.2 Contributions of this Thesis

In reality, the realization of a sub-coalition payoff function may involve opinion consensus, i.e., people's views of each other are affected by each other, and consensus eventually reveals the truth. However, to date, there has been virtually no work on the interplay between coalitional games and opinion consensus theory. This thesis takes an initial step in this direction and shows that this model gives rise to several interesting implications parallel to many social phenomena. As noted before, in this model, players obtain a better estimation of the ground truth of the payoff function by exchanging information; a fair value distribution (i.e., the Shapley value) is also well defined given some conditions for efficient opinion exchange that are stated in the thesis.

The proposed framework of the coalitional game with information exchange results in three interesting phenomena that relate to psychology and sociology. First, at the equilibrium of this game, each participant should be a little overconfident by exaggerating their own contribution in the coalition. Second, in a rational player setting, if the members' influences in a network are proportional to their risk-averse levels, the opinion exchange process is efficient, i.e., it is beneficial to an organization as a whole if more responsible people are taking more important positions. Gradual opinion exchange, instead of an instant opinion fusion, is necessary when players are not fully rational. Finally, in an environment with tremendous social pressure, players tend to switch positions in the opinion exchange process.

1.3 Real World Applications

Deciding equity distribution is a critical step in forming a startup company [12]. For a long period of time, it has been regarded as a problem that that is often solved case by case relying on experience. For example, [12] suggests that equity distribution should consider "past and future contributions," but those contributions are very subjective. To avoid this subjectivity, [17] argues that everyone who joins the startup at the same time should receive equal shares.

Recent years there are some theories and practices trying to deal with this problem in a systematic way. To date, the most suitable theory is the Shapley value in coalitional game theory, in which payoffs are distributed according to the contribution of each of the sub-coalitions and the three axioms of fairness [15]. An online tool, "Startup Equity Calculator," [9] implements this idea by asking the question "What will the company look like without this particular founder?", which essentially evaluates the contribution of each of the sub-coalitions.

However, the above ideas of the Shapley value assume that everyone will agree on the contribution of each of the sub-coalitions. In practice, different people have different opinions about the contribution of each of the sub-coalitions, hence a coalitional game theory with incomplete information is required, such as the one in this paper.

As another example, in the United States, passing legislation requires substantial effort and extensive lobbying and debates, and the same game is not played repeatedly. Thus, a repeated game model in the paper [4] is not applicable. In contemporary U.S. politics, in addition, it is usually the case that the Bayesian core, defined in [11] does not exist, because the two parties have strong prejudices about each other. Moreover, the opinion exchange process affects the outcomes substantially, so a model, such as an opinion consensus model, is required to capture its effect. At the end, each Senator and Representative has her or his own interests and cares almost exclusively about the welfare of his or her own constituents.

1.4 Thesis Struction

The rest of this thesis is organized as follows: Chapter II discusses existing models on coalitional games and opinion consensus. In this section, Proposition 1 shows that with large number of players, the Bayesian core as defined in [6] is likely to be empty, and Proposition 2 gives conditions for linear opinion consensus which could give non-empty Bayesian core. However, the linear opinion consensus models do not consider self-interested players, so a modified model with self-interested players is also discussed. In addition, Chapter III discusses system dynamics with self-interested players in coalitional games with opinion exchange. In this section, Theorem 1 gives conditions for which a coalitional game with opinion exchange is efficient. Furthermore, conditions for which a Bayesian core is non-empty are given in Proposition 4 with a consensus assumption, and Theorem 2 without the consensus assumption. Additionally, Chapter IV shows that an R-learning algorithm can provide a player the best strategy when other players are not rational, i.e., when other players' behaviors have to be learned. Finally, Chapter V gives concluding remarks and discusses further work.

Chapter 2

Coalitional games and opinion consensus models

In a classical coalitional game, it is assumed that the sub-coalition payoff function is common knowledge for all players. In this thesis, this oversimplified assumption is removed, and private sub-coalition payoff functions are allowed. Those different sub-coalition payoff functions represent different evaluations of other players' abilities, i.e., they are private opinions. As indicated in much of the social science literature, people's opinions can affect each other substantially [7]. Thus, such opinion exchange requires a new coalitional game model. My thesis, in particular, uses the linear opinion consensus model [1] as a tool to investigate opinion exchange in coalitional games. Informally, players first carry out opinion consensus, and then they play the classical coalitional game to decide the fair payoff distribution; a rigorous mathematical model of this process is given later in this thesis. However, the coalitional game with opinion exchange is more than a coalitional game after opinion consensus; during the opinion consensus process, each participant is incentivized by her or his final payoff in the coalitional game and may tell lies. That interaction generates a coupling between the coalitional game and the opinion consensus.

2.1 Notations and definitions

This subsection reviews notations and definitions used in classical coalitional games and opinion consensus models.

Definition. *[Supermodularity] Let $N = \{1, 2, \cdots, n\}$ be a set of consecutive integers. Suppose $f(\cdot)$: $2^N \to \mathbb{R}$ is a set function. The set function $f(\cdot)$ is supermodular iff one of the following equivalent conditions holds*

1. $\forall X \subseteq Y \subseteq N$ and $x \in N \backslash Y$, there holds $f(X \cup \{x\}) - f(X) \leq f(Y \cup \{x\}) - f(Y)$, or

2. $\forall X, Y \subseteq N$, there holds $f(X \cup Y) + f(X \cap Y) \geq f(X) + f(Y)$.

The set function is said to be strictly supermodular if the inequalities in the above two equations are strict.

Definition. *[Stochastic Matrix] A matrix $W = [w_{ij}]$ is called a stochastic matrix iff*

1. $\forall i, j$, $w_{ij} \geq 0$, and

2. $\forall i$, there holds $\sum_j w_{ij} = 1$

2.2 Coalitional game

Let $N = \{1, 2, \cdots, n\}$ be a set of n players. In the classical coalitional game setting, a subset $C \subseteq N$ is called a sub-coalition. A set function $v(C)$: $2^N \to \mathbb{R}$ of the subcoalition gives the payoff if sub-coalition C is formed. Note that the cardinality of $\{C | C \subseteq N\}$ is finite, so $v(C)$ can be represented as a vector, v.

In a coalitional game, we consider two major questions:

1. Is there a payoff allocation such that everyone is better-off in the grand coalition? (This problem is solved by the notion of core), and

2. Is there a payoff allocation which is fair to everyone? (This problem is solved by the notion of Shapley value).

The core of a coalitional game is the set of payoff allocation, g_i, $i \in N$:

$$\{g_i | \forall C \subseteq N, \sum_{i \in C} g_i \geq v(C), \text{ and } \sum_{i \in N} g_i = v(N)\}$$

If the core of a coalitional game is not empty, then the coalitional game has a stable solution, such that everyone is better off staying in the grand coalition. Furthermore, the core is non-empty as long as the payoff function is supermodular.

The Shapley value is a payoff allocation, g_i, derived from three fairness principles, i.e., symmetry, linearity, and null player. This value is given by

$$g_i = d_i(v) = \sum_{C \subseteq N \setminus \{i\}} \frac{|C|!(n - |C| - 1)!}{n!} (v(C \cup \{i\}) - v(C))$$

The Shapley value defines a fair distribution of the total payoff $v(N)$.

Example 1. *Suppose there are three players playing a coalitional game. They have a common sub-coalition payoff function* $v(\emptyset) = 0$, $v(1) = v(2) = v(3) = 0.2$, $v(1,2) = v(1,3) = v(2,3) = 0.6$, *and* $v(1,2,3) = 1$.

We fisrt verify that the subset-payoff function $v(C)$: $2^N \to \mathbb{R}$ is supermodular:

$$v(1) + v(2) < v(1,2)$$

$$v(1) + v(3) < v(1,3)$$

$$v(2) + v(3) < v(2,3)$$

$$v(1,2) + v(3) < v(1,2,3)$$

$$v(2,3) + v(1) < v(1,2,3)$$

$$v(1,3) + v(2) < v(1,2,3)$$

Note the sub-coalition payoff function can be written in the form of a vector

$$[v(\emptyset), v(1), v(2), v(3), v(1,2), v(1,3), v(2,3), v(1,2,3)]^T$$

and the condition of supermodularity can be regarded as a set of linear constraints inposed on the vector.

Then we demonstrate the notation of core. Consider the payoff allocation

$$g_1 = g_2 = g_3 = \frac{1}{3}$$

The above payoff allocation satisfies the budget constraint

$$g_1 + g_2 + g_3 = v(1,2,3)$$

and the rationality constraint (i.e., everyone is better staying in the grand coalition)

$$v(i) < g_i$$

$$v(i,j) < g_i + g_j$$

hence it is a stable solution of the coalitional game and an element in the core.

Now we consider the Shapley value $d_i(v)$. Shapley value is defined by three fairness principles, i.e.,

1. Symmetry: if $\forall S, v(S \cap \{i\}) = v(S \cap \{j\})$, then $d_i = d_j$.

2. Linearity: if two coalition games with subcoalition payoff functions v and w are combined, then the allocated payoff d is also combined:

$$d_i(v + w) = d_i(v) + d_i(w)$$

3. Null player: if $\forall S, v(S \cap \{i\}) = v(S)$, then $d_i = 0$

One can verify that the Shapley value

$$g = d(v) = \begin{bmatrix} -\frac{1}{3} & \frac{1}{3} & -\frac{1}{6} & -\frac{1}{6} & \frac{1}{6} & -\frac{1}{3} & \frac{1}{6} & \frac{1}{3} \\ -\frac{1}{3} & -\frac{1}{6} & \frac{1}{3} & -\frac{1}{6} & \frac{1}{6} & \frac{1}{6} & -\frac{1}{3} & \frac{1}{3} \\ -\frac{1}{3} & -\frac{1}{6} & -\frac{1}{6} & \frac{1}{3} & -\frac{1}{3} & \frac{1}{6} & \frac{1}{6} & \frac{1}{3} \end{bmatrix} \begin{bmatrix} v(\emptyset) \\ v(1) \\ v(2) \\ v(3) \\ v(1,2) \\ v(2,3) \\ v(1,3) \\ v(1,2,3) \end{bmatrix} = \begin{bmatrix} \frac{1}{3} \\ \frac{1}{3} \\ \frac{1}{3} \end{bmatrix}$$

satisfies the three properties above. Note that this Shapley value is in the core of the game; this is not a coincidence. If the subcoalition payoff function $v(C)$ of a game is supermodular, then the Shapley value is in the core of this game. Suppose v is the vectorized form of $v(\cdot)$, and we let $v(\emptyset) = 0, v(N) = 1$, then we can also write the Shapley value in the form

$$g_i = \frac{1}{n} + d_i^T v$$

with the plausible property

$$\sum_i d_i = 0$$

We will use the above property in our derivation later.

Neither the core nor the Shapley value is well defined without a payoff function, $v(C)$, which is commonly accepted among all players. However, the notion of a core is generalized in [4] and [11] as a Bayesian core, where private sub-coalition payoff functions are assumed. Note that the definitions of the Bayesian core are different in the above two thesiss, and the definition in our thesis is similar to that in [4]. In our thesis, the Bayesian core is defined as the set of value distributions, g_i, such that every player is better off staying in the grand coalition. Mathematically, "better off" is defined by

$$\forall i, \forall C \subsetneqq N, \sum_{j \in C} g_j \geq v_i(C) \tag{2.1}$$

where v_i is private information of player i, characterizing his or her unique opinion of the game. Furthermore, there holds the budget constraint

$$\forall i, \sum_{j \in N} g_j = v_i(N) \tag{2.2}$$

Now, a value distribution, g_i, is in the Bayesian core iff both (2.1) and (2.2) hold. The problem with the setting of the Bayesian coalitional game is that, in many cases, the Bayesian core is empty even though the core is not empty for each player i. That is particularly true if the number of players, n, is large, as illustrated by Proposition 1.

Proposition 1. *Suppose that there is a strictly supermodular ground truth payoff function $v(C) : 2^\Omega \to \mathbb{R}$ (Note one can represent the function as an m-vector v).*

Further suppose each player's opinion $v_i \sim N(v, \Sigma_i)$ is a sample from the ground truth payoff function, where $N(v, \Sigma_i)$ represents a truncated normal distribution with support $v_i \in \{V | V$ vectorize $v_i(C)$ and $v_i(C)$ is supermodular$\}$, and Σ_i is a diagonal matrix with diagonal entries $\sigma_i^2 > 0$. As the number of players increases, i.e. $n \to \infty$, the Bayesian core defined by (2.1) and (2.2) is empty with probability 1.

Proof. Proof by contradiction. Let S_1 and S_2 be partitions of N. i.e., $S_1 \cap S_1 = \emptyset$ and $S_1 \cup S_2 = N$. Because each player takes a sample from a Gaussian distribution,

$$\lim_{n \to \infty} \mathbb{P}\{\exists i, j, k, \text{ s.t. } v_i(S_1) > \frac{1}{2} v_k(N) \text{ and } v_j(S_2) > \frac{1}{2} v_k(N)\} \to 1$$

The "better off" condition defined by (2.1) gives

$$v_k(N) < v_i(S_1) + v_j(S_2) \leq \sum_{p \in S_1} g_p + \sum_{p \in S_2} g_p = \sum_{p \in N} g_p$$

Now the inequality $\sum_{p \in N} g_p > v_k(N)$ contradicts the budget constraint defined by (2.2). □

In the above proposition, even if each sampled value function has non-empty core, the game itself has empty Bayesian core.

2.3 Opinion consensus

Suppose a graph $\mathcal{G} = \{N, E\}$ characterizes the opinion influence among a set of players N. There is an edge $e_{ij} \in E$ with weight $w_{ij} \in (0, 1)$ if player i has a influence on player j's opinion. If there is not an edge between i and j, we set $w_{ij} = 0$. Furthermore, $w_{ii} = 1 - \sum_{j \neq i} w_{ij} \geq 0$. At each time instance t_k, player i hold an opinion of the function $v_i(C)[k]$. In fact, because the cardinality of $\{C|C \subseteq N\}$ is finite, one can consider $v_i(C)[k]$ as a vector $v_i[k]$.

In the classical opinion consensus literature, all players are truth-telling. When all players are truth- telling, the problem of opinion consensus is decoupled from the coalitional game. Players just update their opinions according to the linear opinion dynamics defined by

$$v_i[k] = \sum_j w_{ij} v_j[k-1]$$

In the above system, opinion consensus can be achieved, i.e., $\forall i$, the limit $\lim_{k\to\infty} v_i[k]$ exists and $\forall i, j, \lim_{k\to\infty} v_i[k] = \lim_{k\to\infty} v_j[k]$, iff the stochastic matrix $W = [w_{ij}]$ has one eigenvalue of 1 and all other eigenvalues are strictly in the unit disk. If the opinion consensus can be achieved, one can define a consensused payoff function $v = \lim_{k\to\infty} v_i[k]$. After the opinion consensus, players can play the coalitional game and a grand coalition exists iff the stable core of v is non-empty. A set of sufficient conditions for which the stable core of v is non-empty is given by Lemma 2 below.

Proposition 2. *There exists a stable coalition under the consensused payoff function v if all of the prior payoff functions $v_i[1]$ are supermodular set functions.*

Proof. First, note that if $v_i[k-1]$ is supermodular, then $v_i[k] = \sum_j w_{ij} v_j[k-1]$ is also supermodular. Since $v_i[1]$ is supermodular, by induction, it follows that $\forall k$, $v_i[k]$ is supermodular.

Because the set of supermodular set functions is closed, the consensused payoff function $v = \lim_{k\to\infty} v_i[k]$ is supermodular. Since the core is non-empty as long as the payoff function is supermodular, a stable coalition exists under the consensused payoff function. □

In reality, however, because players are incentivized by the payoff distributions in the coalitional game, they do not necessarily tell the truth during the opinion consensus process. To incorporate this strategic aspect of opinion consensus, a better model is required. Assume that, at each time instance, player i reveals an opinion, $x_i[k]$, as a decision variable at stage k, that may or may not be equal to $v_i[k]$. Define $\theta \in (0,1)$ as a trust parameter. Now each player updates his or her opinion according to the linear opinion dynamics defined by

$$v_i[k] = \theta \sum_j w_{ij} x_j[k-1] + (1-\theta) v_i[k-1] \tag{2.3}$$

Proposition 3. *Suppose $v_i[k]$, $x_i[k]$, w_{ij} and θ are defined as above, and the opinion dynamics follow* (2.3)*. If*

$$\forall i, \lim_{k \to \infty} x_i[k] = x$$

then

$$\forall i, \lim_{k \to \infty} v_i[k] = x$$

The reverse is not true.

Proof. We first prove that if $\forall i$, $\lim_{k\to\infty} x_i[k] = x$, then $\forall i$, $\lim_{k\to\infty} v_i[k] = x$. We want to show that $\forall \epsilon > 0$, $\exists n \in \mathbb{N}^+$, s.t. $\forall i$, $\forall k \geq n$, $\|v_i[k] - x\|_\infty < \epsilon$.

Let ϵ be given. Because $\lim_{k\to\infty} x_i[k] = x$, $\exists n_1 \in \mathbb{N}^+$ s.t. $\forall i$, $\forall k \geq n_1$, $\|x_i[k] - x\|_\infty < \frac{\epsilon}{2}$. Without loss of generality, one can only consider the jth component of $v_i[k]$, $x_i[k]$ and x, denoted by $(\cdot)_j$. When $k \geq n_1$, one consider the first order feedback system with reference $(x)_j + \frac{\epsilon}{2}$.

$$(\overline{v}_i[k])_j = \theta \left[(x)_j + \frac{\epsilon}{2} \right] + (1-\theta) \left(\overline{v}_i[k-1] \right)_j \tag{2.4}$$

Note that $(v_i[k])_j$ is upper bounded by the above feedback system in (2.4) because $(x_i[k])_j < (x)_j + \frac{\epsilon}{2}$. Furthermore, the above feedback system (2.4) converges to $(x)_j + \frac{\epsilon}{2}$, i.e. $\exists n_2$ s.t. $\forall i$, $\forall k \geq n_2$, $\left|(\overline{v}_i[k])_j - \left[(x)_j + \frac{\epsilon}{2}\right]\right| < \frac{\epsilon}{2}$. Now choose $n_3 = \max\{n_1, n_2\}$, there holds $\forall i$, $\forall k \geq n_3$, $(v_i[k])_j \leq (\overline{v}_i[k])_j < \left[(x)_j + \frac{\epsilon}{2}\right] + \frac{\epsilon}{2} = (x)_j + \epsilon$. Similarly one can find a lower bound system and show that $(v_i[k])_j > (x)_j - \epsilon$. Hence $\exists n \in \mathbb{N}^+$, s.t. $\forall i$, $\forall k \geq n$, $\|v_i[k] - x\|_\infty < \epsilon$.

However, the reverse is not true when $[w_{ij}]$ is singular. □

When there are strategic players, the above propositions shows that consensus in the expressed opinion ensures the consensus of true opinion. Since the true opinion is not directly measurable, the above property shows a way to infer true opinion.

Chapter 3

System dynamics with strategic players

As pointed out in Section 2.3, the system dynamics are trivial when all players are truth-telling because the opinion consensus and coalitional game are decoupled. This chapter discusses the opinion dynamics when players may tell lies to get themselves better payoff distributions. We refer to such players as strategic players.

3.1 Enforcing effective information exchange

In the rational player setting, if telling a lie has no cost, the game becomes a cheap-talk game [8]. Players will not trust any information, and there is no efficient information exchange. Similar problems exist under the cognitive hierarchical model [2], where the opinions will not reach consensus because the second level players again form a cheap-talk game, even though the first level players may tell the truth. We would like to investigate this type of bounded rationality models in future work.

In our coalitional game setting, each player's private knowledge of the sub-coalition payoff function can be viewed as a sample of the ground truth. People enter the opinion consensus to acquire information on other samples, and hence acquire a better understanding of the ground truth. However, revealing false information to others will introduce bias and also undermine trust among the players. Hence, it is useful to introduce a disutility when false information is revealed so that players become risk-averse, and effective information exchange is established.

3.2 Rational and risk-averse players

Consider a ground truth payoff function $v(\cdot)$. Suppose it is normalized, i.e. $v(\emptyset) = 0$ and $v(N) = 1$. When we write it in its vector form, each entries in v are $v(C), \emptyset \subsetneq C \subsetneq N$ (hence it is an m-vector, $m = 2^N - 2$). Note $v(C), \emptyset \subsetneq C \subsetneq N$ is unknown to players. Further suppose that each player's private initial opinion at time instance $k = 0$ is an i.i.d. sample $v_i[0] \sim N(v, \Sigma_i)$, where $\Sigma_i = \sigma_i^2 I$ and I is the identity matrix. Define weight of opinions as $t_i = \frac{1}{\sigma_i^2} / \sum_i \frac{1}{\sigma_i^2}$, then $\hat{v}[0] = \mathbb{E}[v_i[0]] = \sum_i t_i v_i[0]$ is the ML estimator of v. Note that $\mathbb{E}[\cdot]$ denotes weighted average. In addition, assume the influence among players, defined by $W = [w_{ij}]$, satisfies $\lim_{k\to\infty} W^k = 1_n \begin{bmatrix} t_1 & t_2 & \cdots & t_n \end{bmatrix}$, when 1_n denotes n dimensional column-1-vector. Note $[w_{ij}]$ and t_i are also common knowledge; the only private information to player i is its opinion $v_i[k]$.

Suppose the payoff $v(N)$ is allocated according to Shapley value $d_i(\hat{v}[K])$ of the average opinion $\hat{v}[K]$ at step K. Because the Shapley ratio defines a linear function, we can also refer this final payment to player i as $\frac{1}{n} + d_i^T \hat{v}[K]$, where d_i^T is a m-vector. Note the property of Shapley value gives $\sum_i d_i^T = 0$. If every player is truth telling, then the system reaches consensus, i.e. $\forall i, \lim_{k\to\infty} v_i[k] = \hat{v}$, and the final payoff function is the MLE.

Now, assume that players can tell lies. Each player may introduce some fraud at step k: $u_j[k] = x_j[k] - v_j[k]$, but, at the same time, these fraudulent statements undermine trust in the system, and, hence, they introduce disutility $\mathbf{1}^T \text{var}[u]$, where $\text{var}[u] = \sum_i t_i\,(u_i[k])^2 - (\sum_i t_i u_i[k])^2$ (Note that this disutility metric is a scalar, and it can be interpreted as the 1-norm of the variance of $u_i[k]$). After K steps of playing, the overall disutility due to fraud is given by $\mathbf{1}^T \sum_{k=1}^{K} \text{var}\,[u[k]]$. Each player makes a trade-off between $\mathbf{1}^T \sum_{k=1}^{K} \text{var}\,[u[k]]$ and $d_i(\mathbb{E}[v[K]])$ by solving the minimization problem

$$\arg \min_{u_i[k],k=1,2,\cdots} p_i \cdot \mathbf{1}^T \sum_{k=1}^{K} \text{var}\,[u[k]] - d_i^T \mathbb{E}[v[K]] \tag{3.1}$$

Lemma 1. *In the coalitional game with optinion exchange (which follows the system dyanmics* (2.3)*), there holds*

$$\mathbb{E}[v[K]] = \theta\mathbb{E}[u[K-1]] + \mathbb{E}[v[K-1]] = \sum_k \theta\mathbb{E}[u[k]] + \mathbb{E}[v[0]]$$

Proof. Let $T = [t_i]$, $W = [w_{ij}]$, $V[k] = [v_i[k]]$, $X[k] = [x_i[k]]$ and $U[k] = [u_i[k]]$. By definition

$$\mathbb{E}[v[k+1]] = T^{\top} V[k+1]$$

Substitute (2.3) into the above equation. We obtain

$$\mathbb{E}[v[k+1]] = T^{\top} \left(\theta W X[k] + (1-\theta) V[k]\right)$$

By the definition of u_i in the last subsection, it holds that

$$\begin{aligned}\mathbb{E}[v[k+1]] &= T^\top\left(\theta W(V[k]+U[k])+(1-\theta)V[k]\right)\\ &= T^\top\theta WV[k]+T^\top\theta WU[k]+(1-\theta)T^\top V[k]\end{aligned}$$

According to the definition of t_i, the influence among players satisfies $\lim_{k\to\infty} W^k = 1_n \left[\begin{array}{cccc} t_1 & t_2 & \cdots & t_n\end{array}\right]$, i.e., $T^\top W = T^\top$. Therefore

$$\begin{aligned}\mathbb{E}[v[k+1]] &= T^\top\theta V[k]+T^\top\theta U[k]+(1-\theta)T^\top V[k]\\ &= T^\top V[k]+\theta T^\top U[k]\\ &= \sum_k \theta\mathbb{E}[u[k]]+\mathbb{E}[v[0]]\end{aligned}$$

□

From Lemma 1, we obtain

$$\begin{aligned}&\arg\min_{u_i[k],k=1,2,\cdots} p_i\cdot\mathbf{1}^T\sum_{k=1}^{K}\text{var}\left[u[k]\right]-d_i^T\mathbb{E}[v[K]]\\ =&\arg\min_{u_i[k],k=1,2,\cdots}\sum_{k=1}^{K}\left(p_i\cdot\mathbf{1}^T\text{var}\left[u[k]\right]-\theta d_i^T\mathbb{E}[u[k]]\right)\end{aligned}$$

That says that the optimal strategy is indeed a myopic strategy. Therefore, one can seek to find the optimal strategy step-by-step. In step k, it holds that

$$
\begin{aligned}
&\arg\min_{u_i[k]} p_i \cdot \mathbf{1}^T \text{var}\left[u[k]\right] - \theta d_i^T \mathbb{E}[u[k]] \\
&= \arg\min_{u_i[k]} p_i \cdot \mathbf{1}^T \left[\sum_j t_j \left(u_j[k]\right)^2 - \left(\sum_j t_j u_j[k] \right)^2 \right] \\
&- \theta d_i^T \sum_j t_j \cdot u_j[k]
\end{aligned}
$$

Set the first derivative with respect to u_i to zero

$$
2p_i \left(t_i u_i[k] - t_i \left(\sum_j t_j u_j[k] \right) \right) = t_i d_i \theta \tag{3.2}
$$

The above equation defines the best strategy of player i given the actions of the other players. The linear equations above can be used to solve for pure strategy Nash equilibrium. Note that the coefficient matrix of the above linear equations has the rank of $n-1$, so there are multiple Nash equilibria.

Suppose for now that the weight p_i of disutility is proportional to player i's influence t_i in the network, i.e. $p_i \propto t_i$. Further because of the property of the Shapley value, $\sum_j d_j = 0$, one can obtain a solution of (3.2) as

$$
u_i[k] = \frac{d_i \theta}{2p_i} \tag{3.3}
$$

The above solution is a pure strategy Nash-equilibrium, and it yields $\forall i, u_i \neq 0$, and $\sum_i t_i u_i[k] = 0$. In addition, at the equilibrium, $v_i[k]$ will converge, but not achieve consensus. The larger the value of p_i is, the smaller the opinion divergence is, and the more likely that $d_i(\hat{v}[\infty])$ is a stable coalition for all players.

Remark 1. *In practice, the assumption of $p_i \propto t_i$, i.e., the weight p_i of disutility* *is proportional to player i's influence t_i, implies that more responsible players are placed at more important positions in a network.*

Definition 1. *A coalitional game with information exchange is efficient if there exists a Nash equilibrium such that the average opinion is constant.*

Theorem 1. *In the fully rational risk-averse player scenario, i.e., opinion dynamics follows* (2.3) *and strategic players minimize the objective function* (3.1)*, the coalitional game with information exchange is efficient if $p_i \propto t_i$ over all players.*

Proof. Let $T = [t_i]$, $W = [w_{ij}]$, $V[k] = [v_i[k]]$, $X[k] = [x_i[k]]$ and $U[k] = [u_i[k]]$. By Lemma 1,

$$\hat{v}[k+1] = T^\top V[k] + \theta T^\top U[k]$$

Given $p_i \propto t_i$, the optimal strategy for each rational risk-averse player is given by (3.3). Because the solution (3.3) satisfies $\sum_i t_i u_i[k] = 0$, i.e. $T^\top U[k] = 0$, we find that

$$\hat{v}[k+1] = T^\top V[k] = \hat{v}[k]$$

is invariant over time steps k. □

3.3 Existence of stable coalition

This subsection discusses conditions for non-empty Bayesian core. Assuming that consensus is achieved, Proposition 4 gives a sufficient condition for which a stable coalition exists.

Proposition 4. *Suppose that consensus is achieved. There is a stable coalition under the consensused payoff function if all of the prior payoff functions $v_i[1]$ and all the reported payoff functions $x_i[k]$ are supermodular set functions.*

Proof. First, we want to show that $\forall k$, $v_i[k]$ is supermodular.

Let $k \geq 2$ be given. If $v_i[k-1]$ is supermodular, then $v_i[k] = \theta_i \sum_j w_{ij} x_j[k-1] + (1-\theta_i) v_i[k-1]$, as a positive weighted average of supermodular set functions, is also supermodular. In addition, because $v_i[1]$ are supermodular, by induction, we know that $\forall k$, $v_i[k]$ is supermodular.

Given that all payoff functions in step k are supermodular, and because the set of supermodular set functions is a closed set, and further because $v = \lim_{k\to\infty} v_i[k]$, the consensused payoff function v is supermodular. Furthermore, because the Shapley value of a supermodular payoff function is in the stable core, we reach the conclusion that there is a stable coalition under the consensused payoff function. □

In the above proposition, the assumption of achieved consensus may be too strong in practice. Therefore, in Theorem 2, the assumption of achieved consensus is removed, and a stable coalition is shown to exist when p_o is sufficiently large.

Theorem 2. *Assume that $p_i \propto t_i$ over all players, and define $p_o = p_i/t_i$. Then $v_i[\infty] = \lim_{k\to\infty} v_i[k]$ exists. In addition, if $v_i[0]$ is strictly supermodular for any given set of initial states $v_i[0], i \in N$, then $\exists p_o > 0$ s.t. the Bayesian core is non-empty with subcoalition payoff functions $v_i[\infty]$, that is, the Bayesian core is non-empty after the opinion consensus process.*

Proof. Because $\forall i \in N$, $v_i[0]$ is strictly supermodular, $\hat{v}[0]$, the weighted average of all $v_i[0]$, is also strictly supermodular. Further because $p_i \propto t_i$ over all players, the average opinion $\hat{v}[k]$ is invariant over time step k, hence $\hat{v}[k] = \hat{v}[0]$ is strictly supermodular.

Moreover, given the optimal strategy solution (3.3), one can rewrite the system dynamics of (2.3) as

$$v_i[k] = \theta \sum_j w_{ij} \left(v_j[k-1] + \frac{d_i \theta}{2p_o t_i} \right) + (1-\theta) v_i[k-1]$$

Define $V[k] = [v_i[k]]$, $X[k] = [x_i[k]]$, $U[k] = \left[\frac{d_i\theta}{2p_o t_i}\right]$, $W = [w_{ij}]$ and $\overline{W} = \theta W + (1-\theta)I$. Because W is a stochastic matrix, the limits $\lim_{k\to\infty} W^k$ and $\lim_{k\to\infty} \overline{W}^k$ exist and are equal to each other. We define $T = \lim_{k\to\infty} W^k = \lim_{k\to\infty} \overline{W}^k$. A stochastic matrix has a eigenvalue equals 1 and all other eigenvalues inside the unit disk, so one can define an eigenvalue decomposition $\overline{W} = D^{-1}[S_1 + S_2] D$, where $S_1 = \begin{bmatrix} 1 & 0 & \cdots & 0 \\ 0 & 0 & \cdots & 0 \\ \vdots & \vdots & \ddots & \vdots \\ 0 & 0 & \cdots & 0 \end{bmatrix}$ and S_2 is a diagonal matrix with the first entry 0 and all other entries inside the unit disk. The system dynamics is given by

$$\begin{aligned} V[k] &= \theta W \left(V[k-1] + U \right) + (1-\theta) V[k-1] \\ &= \overline{W} V[k-1] + \theta W U \\ &= \overline{W}^k V[0] + \theta W \left(I + \overline{W} + \overline{W}^2 + \cdots + \overline{W}^{k-1} \right) U \end{aligned}$$

If we further consider the eigenvalue decomposition $\overline{W} = D^{-1}[S_1 + S_2] D$, we obtain

$$\begin{aligned} V[k] &= \overline{W}^k V[0] + \theta W D^{\mathsf{T}} \left[I + (S_1 + S_2) + (S_1 + S_2)^2 \right. \\ &\quad \left. + \cdots (S_1 + S_2)^{k-1} \right] D U \\ &= \overline{W}^k V[0] + \theta W D^{\mathsf{T}} \left[I + S_2 + S_2^2 + \cdots S_2^{k-1} \right] D U \\ &\quad + (k-1) \theta W T U \end{aligned}$$

Because the solution (3.3) satisfies $\sum_i t_i u_i = 0$, i.e. $TU = 0$, it holds that

$$V[k] = \overline{W}^k V[0] + \theta W D^\top \left[I + S_2 + S_2^2 + \cdots S_2^{k-1}\right] DU$$

Considering the fact that S_2 is a diagonal matrix with all entries in the unit disk, when $k \to \infty$, the series $I + S_2 + S_2^2 + \cdots S_2^{k-1}$ converges to $(1 - S_2)^{-1}$. Further because $\lim_{k\to\infty} W^k = T$, we obtain

$$\lim_{k\to\infty} V[k] = TV[0] + \theta W D^\top (1 - S_2)^{-1} DU$$

Recall that we have $U = \left[\frac{d_i \theta}{2p_o t_i}\right]$, hence,

$$\lim_{p_o\to\infty} \lim_{k\to\infty} V[k] = TV[0]$$

i.e.,

$$\lim_{p_o\to\infty} v_i[\infty] = \hat{v}[\infty]$$

Given that $\hat{v}[\infty]$ is strictly supermodular, the Shapley value $d(\hat{v}[\infty])$ is in the interior of the core of the coalitional game with subcoalition payoff functions $\hat{v}[\infty]$. For each player i, $\lim_{p_o\to\infty} v_i[\infty] = \hat{v}[\infty]$ and the mapping from $v_i[\infty]$ to the core is continuous. Hence, for sufficiently large p_o, $d(\hat{v}[\infty])$ is also in the core of the coalitional game with subcoalition payoff function $v_i[\infty]$, concluding the proof. □

Remark 2. *The above derivation can be interpreted as follows: 1. The more influential one is, the more risk-averse one should be to ensure zero-drift, and 2. Everyone's best strategy is to be a little overconfident, as pointed out in [3].*

Definition 2. *Suppose $f : 2^N \to \mathbb{R}$ is a supermodular set function. Define a func-*

tional variation

$$g : 2^N \to \mathbb{R},\ s.t. \sum_{\Omega \in 2^N} (g(\Omega))^2 = 1.$$

Then the degree of supermodularity R of the function f is defined as

$$R = \sup_r \{\forall g,\ f + r \cdot g \text{ is supermodular}\}.$$

Remark 3. *In the traditional coalitional game, a grand coalition exists as long as the value set function is supermodular. Moreover, in a coalitional game with imperfect information, the degree of this supermodularity R is related positively to* $\sup\{p_0 \text{ s.t. Baysian core is non-empty}\}$, *i.e., the more supermodular the value set function is, the more robust the grand coalition is when dealing with irresponsible players.*

3.4 Herding model

As pointed out in [19], people do not always follow the rational strategy. In a social context, people are usually affected by social pressures and reveals herding behavior. Considering this herding effect in our coalitional game with opinion exchange, we add a social pressure function to each player's utility function. Now suppose the social pressure of player j is given by $\sum_i (x_j - v_i)^2$.

To illustrate the outcome of this herding model, we look at a two-player example. Suppose player 1 get a unit in the end according to Shapley value and player 2 get $(1 - a)$. Similar to the above, p is the social pressure coefficient. Note a only exists when consensus is achieved.

Player 1:

$$\arg\max_{x_1[k]} a(\mathbf{v}_i[\infty]) - p\sum_{k=1}^{N}\mathbb{E}\left[\sum_{j=1}^{2}(x_1[k]-v_j[k])^2\right]$$

Player 2:

$$\arg\max_{x_2[k]} (1-a(\mathbf{v}_i[\infty])) - p\sum_{k=1}^{N}\mathbb{E}\left[\sum_{j=1}^{2}(x_1[k]-v_j[k])^2\right]$$

Assume at the beginning $\mathbf{v}_i$ are private i.i.d. observations of the true payoff function with standard normal noise.

If players only maximize the first term ($p=0$), then they should always report [1,0] and [0,1] respectively. In this case, consensus will not be achieved.

On the other hand, if they both only minimize the second term ($p \to \infty$), then at the first instance when $k=1$, they should report $\mathbf{x}_i[1] = \mathbf{v}_i[1]$, because at the beginning $\mathbb{E}[v_j] = v_i$ in player i's perspective. After that, when $k \geq 2$, they will choose $\mathbf{x}_1[k] = \mathbf{v}_2[k]$, $\mathbf{x}_2[k] = \mathbf{v}_1[k]$ and $\mathbf{v}_i[k] = \sum_j w_{ji}\mathbf{x}_j[k-1]$. Hence the system becomes

$$\begin{bmatrix} \mathbf{v}_1[k] \\ \mathbf{v}_2[k] \end{bmatrix} = \left(\Theta \begin{bmatrix} w_{11} & w_{12} \\ w_{21} & w_{22} \end{bmatrix}\begin{bmatrix} 0 & 1 \\ 1 & 0 \end{bmatrix} + (1-\Theta)\right)\begin{bmatrix} \mathbf{v}_1[k-1] \\ \mathbf{v}_2[k-1] \end{bmatrix}$$

So the system achieves consensus iff the eigenvalues of $\Theta \begin{bmatrix} w_{11} & w_{12} \\ w_{21} & w_{22} \end{bmatrix}\begin{bmatrix} 0 & 1 \\ 1 & 0 \end{bmatrix} + (1-\Theta)$ is strictly in a unit circle.

In the herding model there are two interesting implications: 1. judging on others with difference opinions case enforce consensus, and 2. when the degree of judgment p is large, two players tend to switch positions.

Chapter 4

Algorithmic playing

In the above analysis, we made the assumption that all of the players are rational. In the case in which all players are fully rational and risk-averse, an equilibrium exists and convergence can be achieved. However, from a single player's perspective, he or she does not have control over how the other players play. What should a player do if he or she is fully rational while others are not? In this chapter, we show that an R-learning algorithm [18] can provide such a player the best strategy.

4.1 R-Learning Formulation

At each step k, the reward of a rational player is given by:

$$r_i[k] = -p_i \cdot \mathbf{1}^T \text{var}\left[u[k]\right] + \theta d_i^T \mathbb{E}\left[u[k]\right]$$

Define $s[k] = \mathbb{E}\left[v[k]\right]$ as the state and u_i as the action of each player i. For convenience, let $i-$ denote the players other than player i. Furthermore, because all of the state variables and action variables are continuous, a model of environment

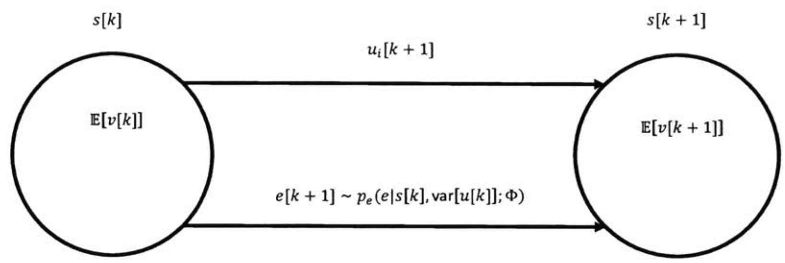

Figure 4-1: States transition of R-learning

(i.e., action pattern of players $i-$) with finite parameters must be defined prior to the learning process. Thus we define the environment $e = \begin{bmatrix} \mathbb{E}\left[u_{i-}[k]\right] \\ \text{var}\left[u_{i-}[k]\right] \end{bmatrix}$, and the environment model $p_e(e|s, \text{var}\left[u[k-1]\right]; \Phi)$ as the probability distribution of e given state s, parameterized by Φ. Note Φ is the set of finite environment parameters to be learned. In addition, the environment e is independent of current action u_i, but the rewards and next state are functions of environment e and the action u_i.

One may find it problematic that the states and the associated rewards are not observable for player i, hence the learning process cannot proceed unless $\text{var}\left[u[k]\right]$ and $\mathbb{E}\left[v[k]\right]$ are broadcast centrally. Furthermore, $\mathbb{E}\left[v[k]\right]$ cannot be obtained so even a central broadcast would be problematic. However, the rewards in each step depend only on the decision variable and environment, but not directly on any state variable; i.e., the impact of state variables only goes into the system via environment e. As a result, the choice of state variables in the R-learning process depends only on how $i-$ players are modeled, and it is possible to choose state variables other than $\mathbb{E}\left[v[k]\right]$, e.g. $\mathbb{E}\left[x[k]\right]$.

Remark 4. *In chapter 4, if everyone is rational and adopts the R-learning algorithm, then* (3.3) *is the optimal strategy. In this case, although chapter 4 and section 3.2* *have the same objective function and the same optimal strategy, some assumptions are different, i.e., section 3.2 assumes that all players know that all players are rational. However, chapter 4 does not have this assumption, but it requires that* $var[u[k]]$ *and* $\mathbb{E}[v[k]]$ *can be broadcast centrally.*

Remark 5. *The learning process justifies the need for gradual consensus of opinion, i.e., the participants learn each other's patterns during the consensus process.*

4.2 Simulations

As an illustrative example, first, we look at a two-player coalitional game with opinion exchange. Suppose player 2's expressed opinion is quasilinear in its true opinion and depends on the mean opinion, i.e. $x_2[k+1] = v_2[k+1] + f(\mathbb{E}[x[k]]) + w$ where w is white noise. Further, assume that player 1 is a risk-averse, rational player as defined in section 3.2, and uses an R-learning algorithm to learn the $f(\cdot)$ function during the opinion consensus process to maximize his or her own utility in the coalitional game.

From player 1's perspective, his or her optimal strategy is given by the solution of (3.2)

$$2p_1\left(u_1[k] - (t_1u_1[k] + t_2u_2[k])\right) = d_1\theta$$

$$u_1^*[k+1] = \frac{d_1\theta}{2p_1(1-t_1)} + \frac{t_2\bar{f}\left(\mathbb{E}[x[k]]\right)}{1-t_1}$$

where $\bar{f}(\cdot)$ is player 1's estimate of $f(\cdot)$.

In the simulation, assume that $v(\{1,2\}) = 1$, $v(\{\emptyset\}) = 0$, $v_i[k] = \begin{bmatrix} v(\{1\}) \\ v(\{2\}) \end{bmatrix}$. Let the initial conditions be $v_1[0] = \begin{bmatrix} 0.7 \\ 0.1 \end{bmatrix}$ and $v_2[0] = \begin{bmatrix} 0.3 \\ 0.5 \end{bmatrix}$. Furthermore, let

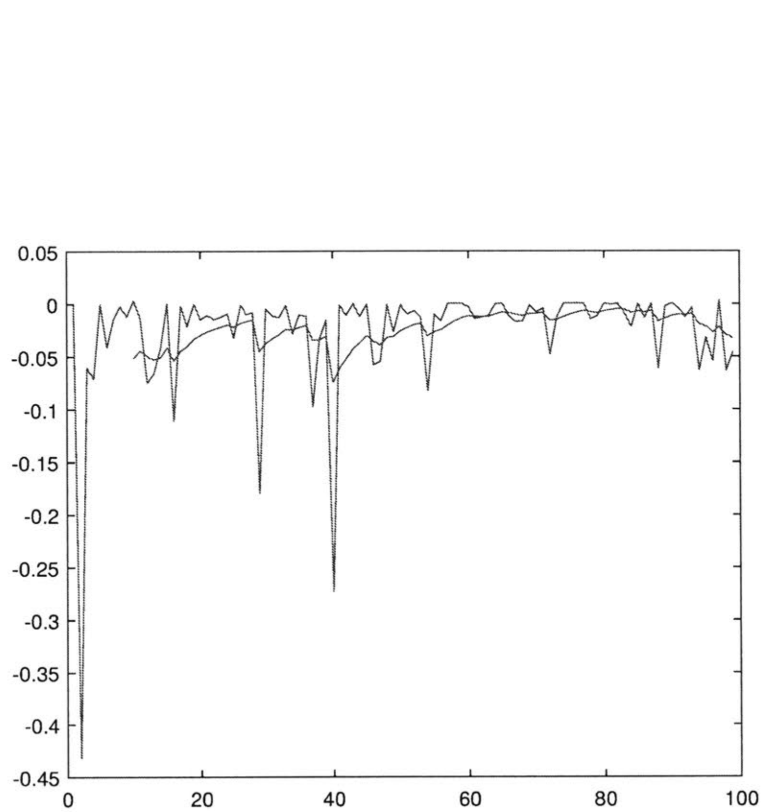

Figure 4-2: Performance $r[k]$ when $\gamma = 0.5$

the system parameters be $\theta = 0.1$, $W = \begin{bmatrix} 0.3 & 0.7 \\ 0.4 & 0.6 \end{bmatrix}$. Player 1 has the probability of $\gamma = 0.5$ of implementing the optimal strategy given its current estimate $\bar{f}(\cdot)$ (exploitation), and this player has the probability of $1-\gamma$ of carrying out exploration.

When player 1 has the probability of $\gamma = 0.5$ of implementing the optimal strategy (exploitation), and the probability of $1 - \gamma = 0.5$ of choosing a random action (exploration), the results are shown in Figures 4-2-4-4. Figure 4-2 shows the performance index $r[k]$ of player 1 in each time step k, Figure 4-3 shows player 2's strategy $f(\cdot)$ (solid line) and player 1's samples on player 2's strategy (circles). Furthermore, Figure 4-4 shows the opinion consensus and evolution process.

When player 1 has the probability of $\gamma = 0.9$ of implementing the optimal strategy (exploitation), and the probability of $1 - \gamma = 0.1$ of choosing a random action

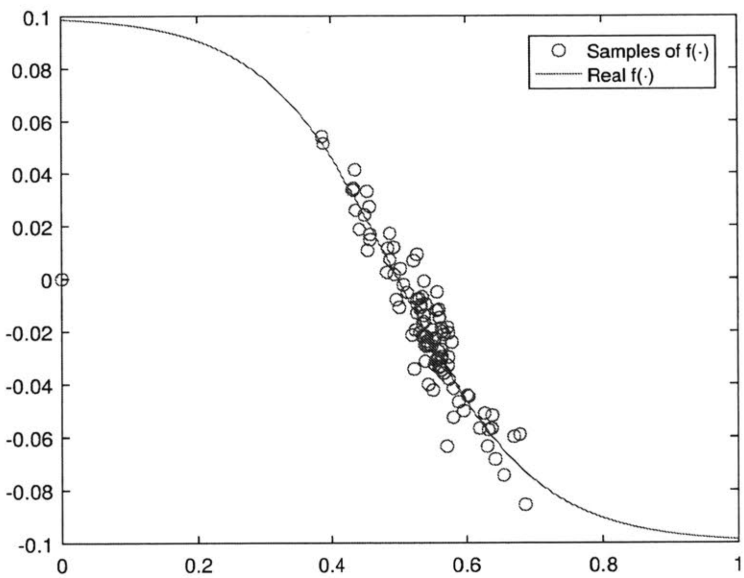

Figure 4-3: Player 2's strategy $f(\cdot)$ when $\gamma = 0.5$

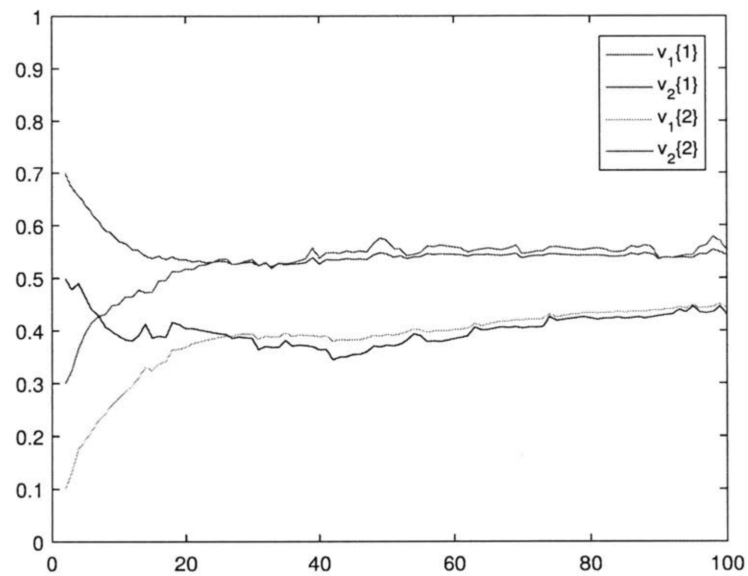

Figure 4-4: True opinion $v[k]$ when $\gamma = 0.5$

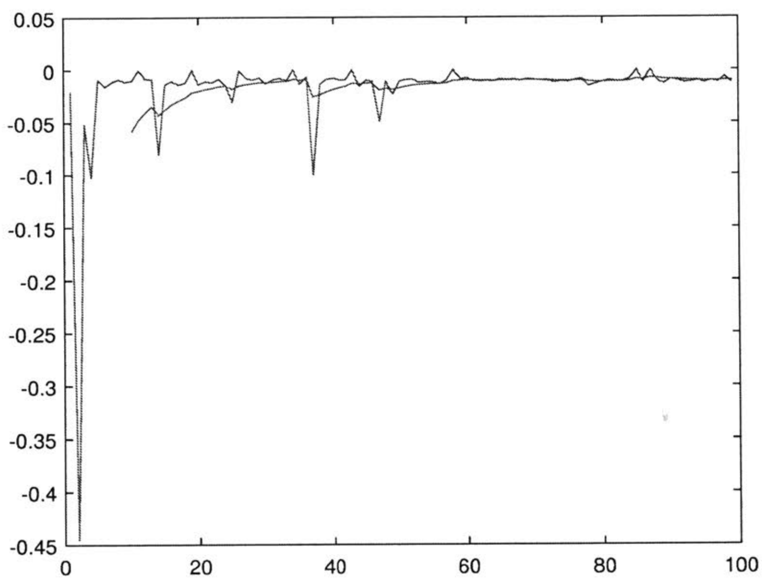

Figure 4-5: Performance $r[k]$ when $\gamma = 0.9$

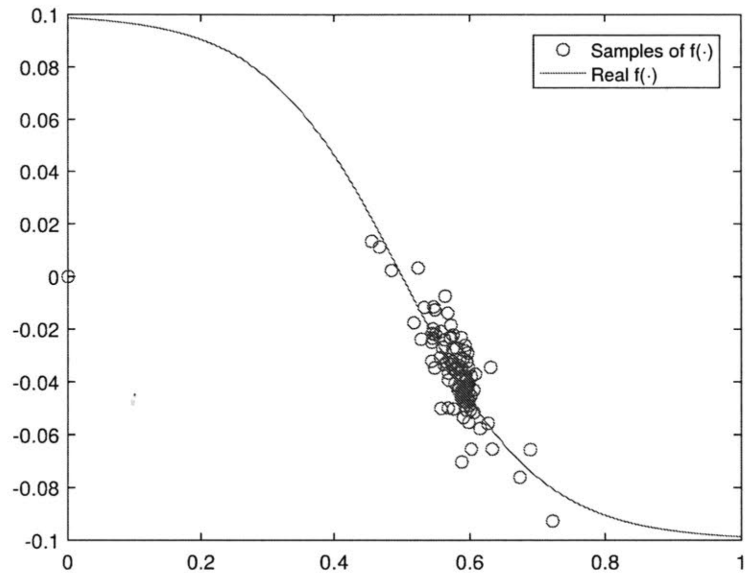

Figure 4-6: Player 2's strategy $f(\cdot)$ when $\gamma = 0.9$

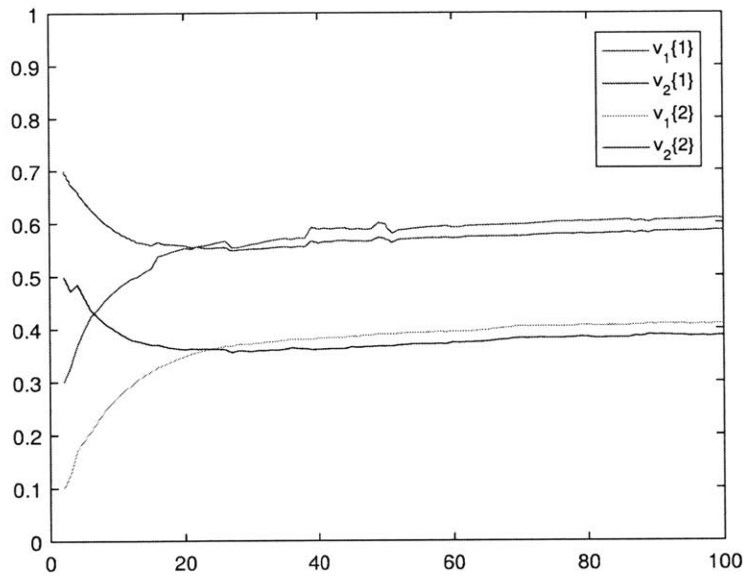

Figure 4-7: True opinion $v[k]$ when $\gamma = 0.9$

(exploration), the results are shown in Figures 4-5-4-7. Figure 4-5 shows the performance index $r[k]$ of player 1 in each time step k. In comparison to Figure 4-2, player 1 with $\gamma = 0.9$ is achieving a higher performance index. Furthermore, Figure 4-6 shows player 2's strategy $f(\cdot)$ (solid line) and player 1's samples on player 2's strategy (circles). Compared to Figure 4-3, samples in Figure 4-6 has a narrower spread. Furthermore, Figure 4-7 shows the opinion consensus and evolution process.

In the above scenario, player 1 is rational and learns the strategy of player 2, but player 2 is not fully rational. As a result, during the opinion consensus process, the average opinion drifts and the coalitional game with information exchange is not efficient. Now, assume that both players are rational and risk-averse, but do not know that their opponents are rational. The coalitional game with information exchange in this case will be efficient, i.e. $d(\hat{v}[k])$ is invarient over k, as shown in the example in Figure 4-8.

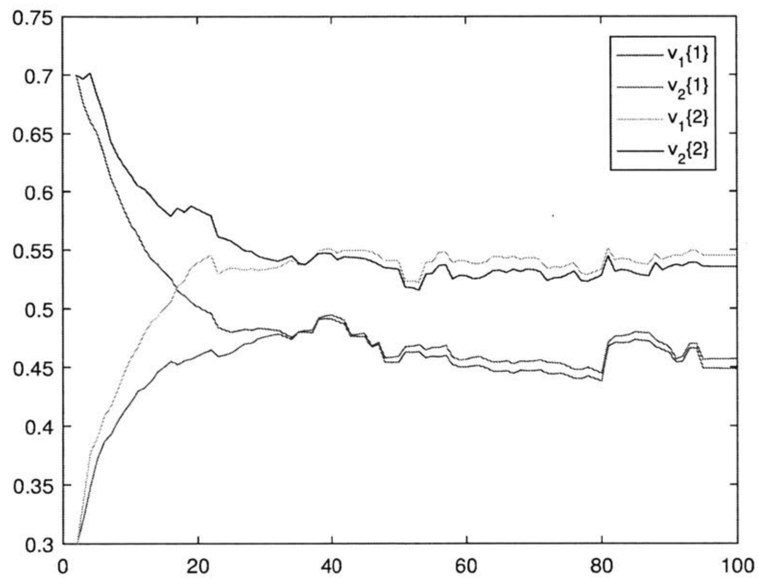

Figure 4-8: True opinion $v[k]$ when $\gamma = 0.8$. Here both players are rational and risk-averse. Both players are doing R-learning to learn the behavior of their opponents.

Chapter 5

Conclusions and Future work

In this thesis, a new framework for coalitional games is presented with an unrealized subset payoff function and information exchange among players. The framework creates an interplay between the traditional model of the coalitional game and the opinion consensus model. Many interesting implications arise from the new framework, including the sufficient condition of non-stable core and the sufficient condition of efficient information exchange. Furthermore, the case of algorithmic learning players was studied, and the results were compared and connected to the case of pure rational players.

In the future, the dependency of equilibrium on the topology of the opinion consensus network may be considered. It is clear that different communication topologies will result in different steady states. From the perspective of an investor in the business scenario, there is a need to design a communication topology and rule (mechanism) that ensures truth telling. From the perspective of the participants, questions may arise concerning with whom and in what order should issues be addressed to ensure favorable outcomes. Additional future work that is needed is related to algorithmic learning. In this approach, quantizing the rewards associated with the possible "exit" action of each player also could be considered.

Bibliography

[1] Giacomo Albi, Lorenzo Pareschi, and Mattia Zanella. Boltzmann-type control of opinion consensus through leaders. *Philosophical Transactions of the Royal Society of London A: Mathematical, Physical and Engineering Sciences*, 372(2028):20140138, 2014.

[2] Colin F Camerer, Teck-Hua Ho, and Juin-Kuan Chong. A cognitive hierarchy model of games. *The Quarterly Journal of Economics*, pages 861–898, 2004.

[3] Christopher Chabris and Daniel Simons. *The invisible gorilla: And other ways our intuitions deceive us.* Harmony Books, 2011.

[4] Georgios Chalkiadakis and Craig Boutilier. Bayesian reinforcement learning for coalition formation under uncertainty. In *Proceedings of the Third International Joint Conference on Autonomous Agents and Multiagent Systems-Volume 3*, pages 1090–1097. IEEE Computer Society, 2004.

[5] Georgios Chalkiadakis and Craig Boutilier. Sequentially optimal repeated coalition formation under uncertainty. *Autonomous Agents and Multi-Agent Systems*, 24(3):441–484, 2012.

[6] Georgios Chalkiadakis, Evangelos Markakis, and Craig Boutilier. Coalition formation under uncertainty: Bargaining equilibria and the bayesian core stability concept. In *Proceedings of the 6th international joint conference on Autonomous agents and multiagent systems*, page 64. ACM, 2007.

[7] James Chapman. The power of propaganda. *Journal of Contemporary History*, 35(4):679–688, 2000.

[8] Joseph Farrell. Cheap talk, coordination, and entry. *The RAND Journal of Economics*, pages 34–39, 1987.

[9] ANTHONY HA. Foundrs.com: Find co-founders, don't get screwed, Oct 2010.

[10] Daylond J Hooper. Coalition formation under uncertainty. Technical report, DTIC Document, 2010.

[11] Samuel Ieong and Yoav Shoham. Bayesian coalitional games. In *AAAI*, pages 95–100, 2008.

[12] Don Rose and Cam Patterson. *Research to Revenue: A Practical Guide to University Start-ups.* UNC Press Books, 2016.

[13] Walid Saad, Zhu Han, Are Hjorungnes, Dusit Niyato, and Ekram Hossain. Coalition formation games for distributed cooperation among roadside units in vehicular networks. *IEEE Journal on Selected Areas in Communications*, 29(1):48–60, 2011.

[14] Walid Saad, Zhu Han, H Vincent Poor, and Tamer Basar. Game-theoretic methods for the smart grid: An overview of microgrid systems, demand-side management, and smart grid communications. *IEEE Signal Processing Magazine*, 29(5):86–105, 2012.

[15] Lloyd S Shapley. A value for n-person games. *The Shapley value*, pages 31–40, 1988.

[16] Onn Shehory and Sarit Kraus. Methods for task allocation via agent coalition formation. *Artificial Intelligence*, 101(1):165–200, 1998.

[17] Joel Spolsky. Where twitter and facebook went wrong — a fair way to divide up ownership of any new company, Apr 2011.

[18] Richard S Sutton and Andrew G Barto. *Reinforcement learning: An introduction*, volume 1. MIT press Cambridge, 1998.

[19] Daniele Vilone, José J Ramasco, Angel Sánchez, and Maxi San Miguel. Social and strategic imitation: the way to consensus. *Scientific reports*, 2, 2012.