Relationship between Nash Equilibria and Pareto Optimal Solutions for Games of Pure Coordination

Rohini Das
Dept. of Electronics and
Telecommunication Engineering
Jadavpur University
India, Kolkata 700032
Email: rohinidas.911@gmail.com

Sayan Goswami
Dept. of Electronics and
Telecommunication Engineering
Jadavpur University
India, Kolkata 700032
Email: sayan.goswami.106@gmail.com

Prof. Amit Konar
Dept. of Electronics and
Telecommunication Engineering
Jadavpur University
India, Kolkata 700032
Email: konaramit@yahoo.co.in

***Abstract*—Game theory is a study of strategic interaction between rational agents. Given the constraint on actions that the players of a game can take, and the individual interests of the players, there are reasonable solutions as to what the outcomes of the game might be. However, there has been no work to contradistinct between the quality of solutions obtained from the various game theoretic approaches like Nash equilibrium and Pareto optimality. Furthermore, no relationship has been established between the aforementioned solution concepts. Our work compares the best response strategy of the agents with their Pareto optimal strategy, for games of pure coordination. It is found that Pareto optimal strategies are a subset of Nash Equilibrium strategies, and the former give the maximum payoff to all agents. Calculation of Pareto Optimal solutions incurs a lesser computational cost than that for Nash Equilibrium solutions.**

Keywords — Game theory, Multi-agent systems, Strategic planning, Artificial intelligence, Economics

I. Introduction

Game Theory is a mathematical framework to model the strategic interactions between rational agents. A game is a pure coordination game if the players cooperate to reach a common goal. A game is purely competitive if there is nothing but conflicting interests between the agents.

A strategy is said to be Pareto efficient if it Pareto dominates all other strategies and a strategy is a Nash Equilibrium [1] if all the agents playing that strategy profile, are best responding to each other.

The solutions to the outcomes of a game may be reasonably modelled by Game Theoretic concepts given the set of permissible actions and the individual goals of the players. Von Neumann and Morgenstern's [2] classic work on game theory covers extensive analysis on zero sum games(or, pure competitive games), from which it can be easily deducted that all solutions are Pareto Optimal, and that, only a Nash Equilibrium in pure or mixed strategy, can give meaningful solutions, to such a game. The relationship between Pareto optimal strategies and Nash equilibrium strategies has however not been established for strategic interactions in pure coordination games. Our work sets out to derive a mathematical relationship between the two solution concepts for pure coordination games. We have considered normal form (bi-matrix) games to find out the optimal solution for the participating agents. The experiments support our proof that we establish in the paper.

The following work is divided into three primary sections. Section II defines the various concepts of Game Theory. Section III contains our proof that establishes the relationship between Pareto optimal and Nash equilibrium strategies. Section IV describes our experimental results obtained from a two robot coordination game of carrying a stick.

II. Related Work

In game theory and economics, coordination games are used to model situations in which players have a common goal and they are rewarded for agreeing on a common strategy. Recent works like [3], have studied the existence of Nash Equilibria which are Pareto Optimal with respect to rest of the Nash Equilibrium strategy profiles, for non cooperative games. The Pareto-optimal Nash equilibrium is one of the most important refinements of the Nash equilibrium that selects the Nash Equilibrium which is Pareto non-dominated with respect to the other Nash Equilibria of the game. In [4], an evolutionary method is described in order to correctly detect this equilibrium, for some two-player normal form games. Our paper studies the mathematical relationship between Pareto Optimal solutions and Nash Equilibrium solutions for games of pure coordination.

In the premise of reinforcement learning, the task at hand is to maximise the payoff obtained from the environment, given a set of actions, agents and states belonging to a state space. Similarly, in our approach, we have demonstrated that two agents exhibiting pure coordination, obtain their maximal payoff by following a Pareto Optimal strategy to choose the corresponding actions at each state of the combined state space of actions. Our work extends the concept of reward discounting to establish a relationship between Pareto Optimality and Nash Equilibrium.

III. Games and Solutions

Our proof adopts background concepts and terminology from the established field of game theory.

Definition 1. *A finite n-person* ***normal form game*** *[5], is a tuple, $\tau \equiv (N, A, u)$, where:*

1) *N is a finite set of n players, indexed by i;*
2) *$A = A_1 \times ... \times A_n$, where A_i is a finite set of actions available to player i. Each vector $a = (a_1, ..., a_n) \in A$ is called an action profile;*
3) *$u = (u_1, ..., u_n)$ where $u_i : A \to \mathbb{R}$ is a real-valued utility function, that associates with each profile of actions, a payoff, $u_i(a)$, for every player i.*

In normal form games, players choose their strategies simultaneously in the game, meaning, no player observes the strategies played by other players before playing his or her own strategy, which is the case for *extensive form games*.

Definition 2. *A* ***pure coordination game*** *[5] is a game in which for all action profiles $a = (a_1, ..., a_n) \in A_1 \times ... \times A_n$ and any pair of agents (i, j), it is the case that $u_i(a) = u_j(a)$.*

Pure coordination games are also called *common-payoff games* or *team games*, in which, for every action profile, all players have same payoff. Here, the agents have no conflicting interests at all; their sole ambition is to coordinate on an action that gives equally maximal payoff to all the players in the game.

Definition 3. *A finite* ***zero-sum game of two players*** *is defined as $N = \{1, 2\}$, $A = A_1 \times A_2$ and $u = (u_1, u_2)$ with the constraint that for all $(a_1, a_2) \in A$, we have*

$$u_1(a_1, a_2) + u_2(a_1, a_2) = 0$$

Zero-sum games are also know as *pure competition games*. The two player zero-sum games hold an important role in game theory because of a multitude of reasons. They were the first set of games to be theoretically analyzed by Von-Neumann and Morgenstern [2] when they came up with the theory of games. Also, the zero-sum games are pervasive - examples include any real game where one player's loss is another player's gain.

Definition 4. *An action profile $(a_1^*, ..., a_n^*)$ in a strategic form game $\tau \equiv (N, A, u)$ is a* ***Nash equilibrium*** *of τ if for all $i \in N$, we have*

$$u_i(a_i^*, a_{-i}^*) \geq u_i(a_i, a_{-i}^*) \forall a_i \in A_i$$

The definition above requires that, given, actions of other players a_{-i}^*, which are the equilibrium actions for other players, a unilateral deviation by Player i is not profitable. The idea of a Nash equilibrium is that of a stable state, where each player is responding optimally given the actions of the other players. An alternate way to put out the definition is using the idea of best response. For any $a_{-i} \in A_{-i}$, we define $B_i(a_{-i})$ to be the set of player i's best action response, given a_{-i}, as follows:

$$B_i(a_{-i}) = \{a_i \in A_i : u_i(a_i, a_{-i}) \geq u_i(a_i', a_{-i}) \forall a_i' \in A_i\}$$

We call the set-valued function B_i, the **best response** function of player i.

Now, an action profile $(a_1^*, ..., a_n^*)$ is a **Nash equilibrium** if for all $i \in N$,

$$a_i^* \in B_i(a_{-i})$$

The above definition describes a *pure strategy Nash Equilibrium*. If the concept of action set, A_i, is extended to strategy set, S_i, for every player i, which is the set of all probability distributions over A_i, the aforementioned interpretation of Nash Equilibrium, will give us the *mixed strategy Nash Equilibrium*.

Definition 5. *Strategy profile $\tilde{s}$* ***Pareto dominates*** *strategy profile s if for all $i \in N$, $u_i(\tilde{s}) \geq u_i(s)$, and there exists some $j \in N$ for which $u_j(\tilde{s}) > u_j(s)$.*

In other words, in a Pareto-dominated strategy profile, a player can be made better off without making any other player worse off. Here, strategy profiles are considered as outcomes of a game.

Definition 6. *Strategy profile $\tilde{s}$ is* ***Pareto optimal****, or strictly Pareto efficient, if there does not exist another strategy profile $s \in S$ that Pareto dominates $\tilde{s}$.*

IV. Nash-Pareto commonality in pure Coordination games

In the following section, we describe our proof that establishes that solutions obtained through the conditions of Pareto optimality are a subset of the solutions obtained through the conditions of Nash Equilibrium.

Theorem 1. *Pareto optimal (PO) solutions are a subset of the Nash equilibrium (NE) solutions for games of pure coordination.*

Proof. We have considered a two-player pure coordination game.

Let (a_1^*, a_2^*), $a_1^* \in A_1$, $a_2^* \in A_2$ be a Nash equilibrium strategy. Then,

$$u_1(a_1^*, a_2^*) \geq u_1(a_1, a_2^*) \forall a_1 \in A_1 \tag{3}$$

$$u_2(a_1^*, a_2^*) \geq u_2(a_1^*, a_2) \forall a_2 \in A_2 \tag{4}$$

Now, from the condition of a pure coordination game,

$$u_1(a_1, a_2) = u_2(a_1, a_2) \forall a_1 \in A_1 \setminus \{a_1^*\}, a_2 \in A_2 \setminus \{a_2^*\}$$

and,

$$u_1(a_1^*, a_2^*) = u_2(a_1^*, a_2^*), \forall a_1^* \in A_1, a_2^* \in A_2$$

Now, let $(\tilde{a}_1, \tilde{a}_2)$ be the Pareto optimal outcome or strategy. Then,

$$\forall \tilde{a}_1 \in A_1, \tilde{a}_2 \in A_2, a_1 \in A_1 \setminus \{\tilde{a}_1\}, a_2 \in A_2 \setminus \{\tilde{a}_1\}$$

$$\begin{aligned} u_1(\tilde{a}_1, \tilde{a}_2) &\geq u_1(a_1, a_2) \\ u_1(\tilde{a}_1, \tilde{a}_2) &\geq u_1(a_1, \tilde{a}_2) \qquad (5) \\ u_1(\tilde{a}_1, \tilde{a}_2) &\geq u_1(\tilde{a}_1, a_2) \end{aligned}$$

And,

$$\begin{aligned} u_2(\tilde{a}_1, \tilde{a}_2) &> u_2(a_1, a_2) \qquad (1) \\ u_2(\tilde{a}_1, \tilde{a}_2) &> u_2(a_1, \tilde{a}_2) \\ u_2(\tilde{a}_1, \tilde{a}_2) &> u_2(\tilde{a}_1, a_2) \qquad (6) \end{aligned}$$

Now, from the condition of a pure coordination game,

$$u_1(a_1, a_2) = u_2(a_1, a_2) \forall a_1 \in A_1, a_2 \in A_2 \qquad (7)$$

Thus,

$$\begin{gathered} u_1(\tilde{a}_1, \tilde{a}_2) = u_2(\tilde{a}_1, \tilde{a}_2) \qquad (2) \\ \forall \tilde{a}_1 \in A_1, \tilde{a}_2 \in A_2 \end{gathered}$$

Now from equation (1), we have,

$$u_2(\tilde{a}_1, \tilde{a}_2) > u_2(a_1, a_2)$$

.

Extending from equation (2) and equation (7),

$$u_1(\tilde{a}_1, \tilde{a}_2) > u_1(a_1, a_2)$$

Now, let $(\tilde{a}_1, \tilde{a}_2)$ be a Nash equilibrium strategy.
Let, $\tilde{a}_1 = a_1^*$ and, $\tilde{a}_2 = a_2^*$.
Thus, equation (3) and equation (4) are now valid.

Now, the inequalities below,

$$\begin{aligned} u_1(\tilde{a}_1, \tilde{a}_2) &\geq u_1(a_1, \tilde{a}_2) \\ u_2(\tilde{a}_1, \tilde{a}_2) &\geq u_2(\tilde{a}_1, a_2) \end{aligned}$$

are also true from equation (5) and equation (6), which are the conditions for Pareto optimality.

Thus, our assumption that $(\tilde{a}_1, \tilde{a}_2)$ is also a Nash equilibrium strategy is correct. Since, $(\tilde{a}_1, \tilde{a}_2)$ is an arbitrary strategy, it can be concluded that, all Pareto optimal solutions are Nash equilibrium solutions too.

$$\therefore PO \subseteq NE \quad \square$$

V. Experiments and Results

The following section describes our experiment using a game of pure coordination. It affirms our proof that Pareto optimal strategies are a subset of Nash equilibrium strategies for games of pure coordination.

A. Stick Carrying Game

The stick carrying game [6] is a game of pure coordination. In a two agent stick carrying system, the task is to carry a fixed length stick from the starting location to the goal location without violating the physical constraints of the game world. For a two agent stick carrying game with a stick length of 1 unit, the agents start from a pair of cells (S_1, S_2) and reach the target (G_1, G_2) without dropping the stick. The physical constraint of maintaining the constant stick length is enforced my maintaining a Manhattan distance of unity at each state of the game between the agents' states.

Definition 7. *The **Manhattan distance** between two points (x_1, y_1) and (x_2, y_2) is given as*

$$d_{Manhattan} = |x_2 - x_1| + |y_2 - y_1|$$

where (x_1, y_1) and (x_2, y_2) are co-ordinates on the Cartesian plane.

The experiment is conducted on a square grid world of dimension 10 cells by 10 cells. Fig 1. represent the path taken by a two agent system when starting at the joint state (S_1, S_2) and reaching the goal state (G_1, G_2). Obstacles are placed in the environment at states 9, 27, 40, 46, 52, 54, 58, 61, 63, 67, 82 and 85. These obstacles are represented by the unconnected, greyed out nodes in the map.

Each agent can choose one out of up to four possible actions at every time step, they are Up (U), Right (R), Down (D) and Left (L). The optimal strategies for the agents are obtained using the discounted returns [7] for the agents.

Definition 8. *The discounted return G_t is defined as*

$$G_t \doteq R_{t+1} + \gamma R_{t+2} + \gamma^2 R_{t+3} + ... = \sum_{k=0}^{\infty} \gamma^k R_{t+k+1}$$

where γ is a parameter, $0 \leq \gamma \leq 1$, called the discount rate.

Hence, the reward obtained by the i^{th} agent at time step t, when it is d distance away from it's goal state is given as:

$$G^i{}_t = \gamma^d \times r_{max}$$

where, γ is the discount rate, considered 0.9 in our case and r_{max} is the maximum immediate reward, 100 in our environment. Any violation of the constraints is penalised with reward of $r = -1$. The agents are thus incentivized to reach their goal state as early as possible to maximise their expected return.

For the path traversed in Fig 1, the Nash Equilibrium and Pareto Optimal strategies for each time step are obtained, the results are shown in Table I, which supports our claim that the strategies obtained using the condition for Pareto optimality are indeed a subset of the strategies obtained using Nash equilibrium. The step wise calculation of the the discounted returns, Pareto optimal strategies, Nash equilibria for the states are shown in Table II.

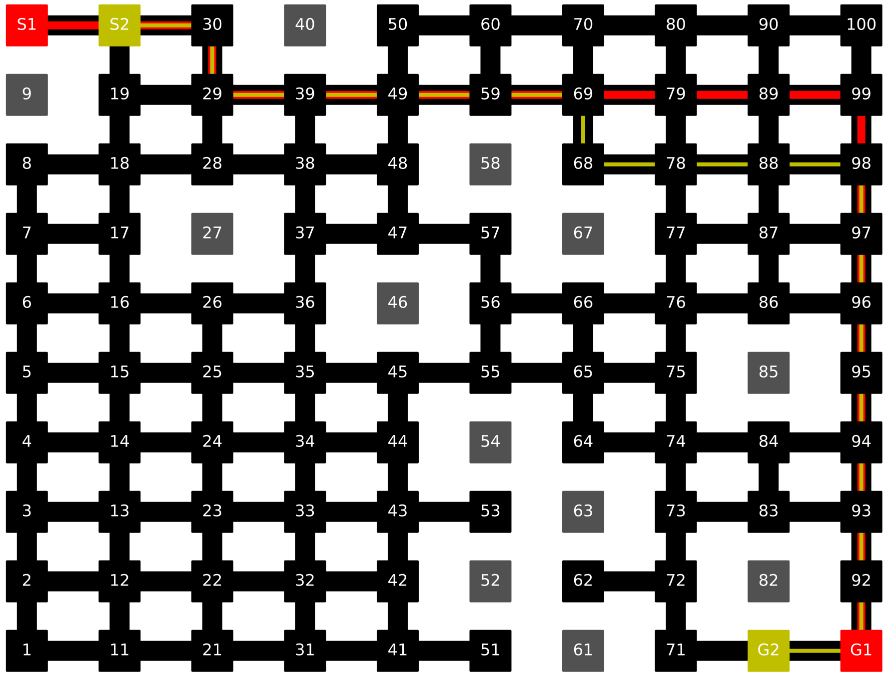

Fig. 1: Path taken by a two agent system in the stick carrying game to reach from (S_1, S_2) to (G_1, G_2). Red colour denotes the path taken by Agent 1 whereas yellow colour denotes the path taken by Agent 2.

TABLE I: State-wise optimal strategies for the stick carrying game

Joint States	N.E. Strategies	P.O. Strategies	Comment
$(10, 20)$	$\{RR, RD\}$	$\{RR, RD\}$	$PO = NE$
$(20, 30)$	$\{DL, RD, DD\}$	$\{DL, RD, DD\}$	$PO = NE$
$(30, 29)$	$\{LL, LU, DR, DD\}$	$\{DR, DD\}$	$PO \subset NE$
$(29, 39)$	$\{RR, RD, DL\}$	$\{RR, RD\}$	$PO \subset NE$
$(39, 49)$	$\{LL, RR, RU, RD, DD\}$	$\{RR, RU, RD, DD\}$	$PO \subset NE$
$(49, 59)$	$\{RR, RU, DL\}$	$\{RR, RU, DL\}$	$PO = NE$
$(59, 69)$	$\{LL, RR, RU, RD, UL, UU\}$	$\{LL, RR, RU, RD, UL, UU\}$	$PO = NE$
$(69, 68)$	$\{RR, DR, RU\}$	$\{RR, DR, RU\}$	$PO = NE$
$(79, 78)$	$\{RR, DR, LU, UU, DD, DL\}$	$\{RR, DD, DR, DL\}$	$PO \subset NE$
$(89, 88)$	$\{LL, LU, RR, UU, DR, DD\}$	$\{RR, DR, DD\}$	$PO \subset NE$
$(99, 98)$	$\{LL, LU, UU, DD\}$	$\{DD\}$	$PO \subset NE$
$(98, 97)$	$\{LL, LU, UU, DD\}$	$\{DD\}$	$PO \subset NE$
$(97, 96)$	$\{LL, LU, UU, DD\}$	$\{DD\}$	$PO \subset NE$
$(96, 95)$	$\{LU, UU, DD\}$	$\{DD\}$	$PO \subset NE$
$(95, 94)$	$\{UU, DD\}$	$\{DD\}$	$PO \subset NE$
$(94, 93)$	$\{LL, LU, UU, DD\}$	$\{DD\}$	$PO \subset NE$
$(93, 92)$	$\{LU, UU, DD\}$	$\{DD\}$	$PO \subset NE$
$(92, 91)$	$\{UU, DL\}$	$\{DL\}$	$PO \subset NE$

VI. Conclusion

Pareto optimal solutions are a subset of the Nash equilibrium solutions for games of pure coordination. Evaluating Pareto optimal solutions is computationally less expensive since it requires $O(n)$ operations, which involves, finding the globally maximum payoffs. However, calculation of Nash equilibria requires $O(n^2)$ computations, that involves, finding the joint state, wherein all the agents play their dominant strategies. Pareto optimal strategies ensure maximum payoff to all agents of the game. Thus, Pareto optimal solutions provide the best outcome for games of pure coordination.

TABLE II: Calculation of Nash Equilibrium and Pareto Optimal Strategies for a few of the intermediate states of the stick carrying game are shown as follows.

(a) $(S_1, S_2) = (10, 20)$

A_1 \ A_2	R	L	D
R	(83, 83)	(70, 70)	(83, 83)

N.E. Strategies = $\{RR, RD\}$
P.O. Strategies = $\{RR, RD\}$

(b) $(S_1, S_2) = (20, 30)$

A_1 \ A_2	L	D
L	(82, 82)	(70, 70)
D	(84, 84)	(84, 84)
R	(70, 70)	(84, 84)

N.E. Strategies = $\{DL, RD, DD\}$
P.O. Strategies = $\{DL, RD, DD\}$

(c) $(S_1, S_2) = (30, 29)$

A_1 \ A_2	L	R	U	D
L	(83, 83)	(70, 70)	(83, 83)	(70, 70)
D	(70, 70)	(85, 85)	(70, 70)	(85, 85)

N.E. Strategies = $\{LU, LL, DR, DD\}$
P.O. Strategies = $\{DR, DD\}$

(d) $(S_1, S_2) = (29, 39)$

A_1 \ A_2	L	R	D
L	(84, 84)	(70, 70)	(70, 70)
R	(70, 70)	(86, 86)	(86, 86)
U	(84, 84)	(70, 70)	(70, 70)
D	(85, 85)	(70, 70)	(84, 84)

N.E. Strategies = $\{RR, RD, DL\}$
P.O. Strategies = $\{RR, RD\}$

(e) $(S_1, S_2) = (39, 49)$

A_1 \ A_2	L	R	U	D
L	(85, 85)	(70, 70)	(70, 70)	(70, 70)
D	(70, 70)	(86.8, 86.8)	(86.8, 86.8)	(86.8, 86.8)
R	(85, 85)	(70, 70)	(70, 70)	(86.8, 86.8)

N.E. Strategies = $\{LL, RR, RU, Rd, DD\}$
P.O. Strategies = $\{RR, RU, RD, DD\}$

(f) $(S_1, S_2) = (49, 59)$

A_1 \ A_2	L	R	U
L	(86, 86)	(70, 70)	(70, 70)
R	(70, 70)	(87.7, 87.7)	(87.7, 87.7)
U	(86, 86)	(70, 70)	(86, 86)
D	(87.7, 87.7)	(70, 70)	(70, 70)

N.E. Strategies = $\{RR, RU, DL\}$
P.O. Strategies = $\{RR, RU, DL\}$

(g) $(S_1, S_2) = (93, 92)$

A_1 \ A_2	U	D
L	(96, 96)	(70, 70)
U	(96, 96)	(70, 70)
D	(70, 70)	(98, 98)

N.E. Strategies = $\{LU, UU, DD\}$
P.O. Strategies = $\{DD\}$

(h) $(S_1, S_2) = (59, 69)$

A_1 \ A_2	L	R	U	D
L	(86.8, 86.8)	(70, 70)	(70, 70)	(70, 70)
R	(70, 70)	(86.8, 86.8)	(86.8, 86.8)	(86.8, 86.8)
U	(86.8, 86.8)	(70, 70)	(86.8, 86.8)	(70, 70)

N.E. Strategies = $\{LL, RR, RU, RD, DD\}$
P.O. Strategies = $\{RR, RU, RD, DD\}$

(i) $(S_1, S_2) = (92, 91)$

A_1 \ A_2	L	U
U	(70, 70)	(97, 97)
D	(99, 99)	(70, 70)

N.E. Strategies = $\{UU, DL\}$
P.O. Strategies = $\{DL\}$

TABLE III: Calculation of Nash Equilibrium and Pareto Optimal Strategies for a few of the intermediate states of the stick carrying game are shown as follows.

(a) $(S_1, S_2) = (69, 68)$

A_1 \ A_2	U	R
L	(87.7, 87.7)	(70, 70)
R	(89.5, 89.5)	(89.5, 89.5)
U	(87.7, 87.7)	(70, 70)
D	(70, 70)	(89.5, 89.5)

N.E. Strategies = $\{RR, DR, RU\}$
P.O. Strategies = $\{RR, DR, RU\}$

(b) $(S_1, S_2) = (79, 78)$

A_1 \ A_2	L	R	U	D
L	(88.6, 88.6)	(70, 70)	(88.6, 88.6)	(70, 70)
R	(70, 70)	(90.4, 90.4)	(88.6, 88.6)	(70, 70)
U	(70, 70)	(70, 70)	(88.6, 88.6)	(70, 70)
D	(90.4, 90.4)	(90.4, 90.4)	(70, 70)	(90.4, 90.4)

N.E. Strategies = $\{RR, DR, LU, UU, DD, DL\}$
P.O. Strategies = $\{RR, DD, DL, DR\}$

(c) $(S_1, S_2) = (89, 88)$

A_1 \ A_2	L	R	U	D
L	(89.5, 89.5)	(70, 70)	(89.5, 89.5)	(70, 70)
R	(70, 70)	(91.3, 91.3)	(87.7, 87.7)	(70, 70)
U	(70, 70)	(70, 70)	(89.5, 89.5)	(70, 70)
D	(89.5, 89.5)	(91.3, 91.3)	(70, 70)	(91.3, 91.3)

N.E. Strategies = $\{LL, LU, RR, UU, DR, DD\}$
P.O. Strategies = $\{RR, DD, DR\}$

(d) $(S_1, S_2) = (99, 98)$

A_1 \ A_2	L	U	D
L	(90.4, 90.4)	(90.4, 90.4)	(70, 70)
U	(70, 70)	(90.4, 90.4)	(70, 70)
D	(90.4, 90.4)	(70, 70)	(92.2, 92.2)

N.E. Strategies = $\{LL, LU, UU, DD\}$
P.O. Strategies = $\{DD\}$

(e) $(S_1, S_2) = (98, 97)$

A_1 \ A_2	L	U	D
L	(91.3, 91.3)	(91.3, 91.3)	(70, 70)
U	(70, 70)	(91.3, 91.3)	(70, 70)
D	(91.3, 91.3)	(70, 70)	(93.2, 93.2)

N.E. Strategies = $\{LL, LU, UU, DD\}$
P.O. Strategies = $\{DD\}$

(f) $(S_1, S_2) = (97, 96)$

A_1 \ A_2	L	U	D
L	(92.2, 92.2)	(92.2, 92.2)	(70, 70)
U	(70, 70)	(92.2, 92.2)	(70, 70)
D	(92.2, 92.2)	(70, 70)	(94.1, 94.1)

N.E. Strategies = $\{LL, LU, UU, DD\}$
P.O. Strategies = $\{DD\}$

(g) $(S_1, S_2) = (96, 95)$

A_1 \ A_2	U	D
L	(93.2, 93.2)	(70, 70)
U	(93.2, 93.2)	(70, 70)
D	(70, 70)	(95.1, 95.1)

N.E. Strategies = $\{LU, UU, DD\}$
P.O. Strategies = $\{DD\}$

(h) $(S_1, S_2) = (95, 94)$

A_1 \ A_2	L	U	D
U	(70, 70)	(94.1, 94.1)	(70, 70)
D	(94.1, 94.1)	(70, 70)	(96, 96)

N.E. Strategies = $\{UU, DD\}$
P.O. Strategies = $\{DD\}$

(i) $(S_1, S_2) = (94, 93)$

A_1 \ A_2	L	U	D
L	(95, 95)	(95, 95)	(70, 70)
U	(70, 70)	(95, 95)	(70, 70)
D	(95, 95)	(70, 70)	(97, 97)

N.E. Strategies = $\{LL, LU, UU, DD\}$
P.O. Strategies = $\{DD\}$

References

[1] J. F. Nash, “Equilibrium points in n-person games,” *Proceedings of the National Academy of Sciences*, vol. 36, no. 1, pp. 48–49, 1950. [Online]. Available: https://www.pnas.org/content/36/1/48

[2] J. von Neumann and O. Morgenstern, *Theory of Games and Economic Behavior (60th-Anniversary Edition)*. Princeton University Press, 2007. [Online]. Available: http://press.princeton.edu/titles/7802.html

[3] V. I. Zhukovskiy and K. N. Kudryavtsev, “Pareto-optimal nash equilibrium: Sufficient conditions and existence in mixed strategies,” *Automation and Remote Control*, vol. 77, no. 8, pp. 1500–1510, 2016. [Online]. Available: https://doi.org/10.1134/S0005117916080154

[4] N. Gask, M. Suciu, R. Ioana Lung, and D. Dumitrescu, “Pareto-optimal nash equilibrium detection using an evolutionary approach,” *Acta Universitatis Sapientiae. Informatica*, vol. 4, 01 2012.

[5] Y. Shoham and K. Leyton-Brown, *Multiagent Systems - Algorithmic, Game-Theoretic, and Logical Foundations*. Cambridge University Press, 2009.

[6] A. Kumar Sadhu and A. Konar, “Improving the speed of convergence of multi-agent q-learning for cooperative task-planning by a robot-team,” *Robotics and Autonomous Systems*, vol. 92, 03 2017.

[7] R. S. Sutton and A. G. Barto, “Reinforcement learning: An introduction,” *IEEE Trans. Neural Networks*, vol. 9, no. 5, pp. 1054–1054, 1998. [Online]. Available: https://doi.org/10.1109/TNN.1998.712192