Introduction to Game Theory

June 8, 2021

In this chapter, we review some of the central concepts of Game Theory. These include the celebrated theorem of John Nash which gives us an insight into equilibrium behaviour of interactions between multiple individuals. We will also develop some of the mathematical machinery which will prove useful in our understanding of learning algorithms. Finally, we will discuss why learning is a useful concept and the real world successes that multi-agent learning has achieved in the past few years.

We provide many of the important proofs in the Appendix for those readers who are interested. However, these are by no means required to understand the subsequent chapters.

1 What is a Game?

What is the first image that comes to mind when you think of a game? For most people, it would be something like chess, hide and seek or perhaps even a video game. In fact, all of these can be studied under Game Theoretical terms, but we can extend our domain of interest much further. We do this by understanding what all of our above examples have in common

- I There are multiple players
- II All players want to win the game
- III There are rules which dictate who wins the game
- IV Each player's behaviour will depend on the behaviour of the others

With these in place, we can extend our analysis to almost any interaction between multiple players. Let us look at some examples, noticing in each, that components of a game that we discussed above show up in each of the following.

Example 1 (The Ultimatum Game [3])**.** *Let us consider that there are two players, the first of which is called the* Proposer*. The Proposer is given a sum of money, say £10 and is required to offer some of it to the second player, called the* Responder*. The Responder will then either accept or reject this offer. If accepted, both players will receive the money as the per the Proposer's suggestion. If it is rejected, both players receive nothing. What would be the likely outcome of this game? From the Proposer's perspective, the selfish choice would be* *to offer as little as possible, say £9.99. The Responder's rational choice would be to accept whatever is offered, since they would otherwise end up with nothing.*

This game has been studied extensively by experimental economists as a way of understanding the degree to which people act 'fairly'. The interested reader is pointed to [9] for an exposition into the incredibly interesting results of these experiments.

Example 2 (The El Farol Bar problem)**.** *In this example, there are a group of N (let us say 100) people who must decide whether or not to go to a bar. The space is limited and nobody wants to go to a crowded bar. In particular let us say that the bar is 'crowded' if there are more than 60 people in attendance. Now, since there is no way to tell in advance who will go and who will not, each agent must make decide whether or not to go based on if they* expect *fewer than 60 people to be in attendance.*

This is example is a particular case of the family of Minority Games*, as each agent wants to be in the minority group. Another, more timely example of this would include travel plans upon the lifting of lockdown restrictions - if I anticipate that many people will want to go on holiday immediately, then I would prefer not to travel immediately.*

Minority Games are the object of intense study from the point of view of dynamics and statistical physics (c.f. [2]) as we will see when we revisit them later in these lectures.

Example 3 (The Prisoner's Dilemma)**.** *Consider the following. Two criminals are arrested due to their connection with a serious crime. They have the options to: choose to confess to the crime or accept the punishment for some less serious crime that they have also been charged with. The judge offers them the following deal*

- *If they both confess, they will receive a reduced sentenced of five years for their crime.*
- *If only one confesses, whilst the other keeps shut, then the former will walk free, acquitted of both charges, whilst the other will receive the full penalty for their crime: a total of 20 years in prison.*

Of course, if neither of them confess, then they can only be sentenced for their less serious crime and so will receive a sentence of a single year.

Now notice, that the cost that each player incurs (i.e. their sentence) is dependent, not only on their own actions, but also those of their opponent (namely the other prisoner). A common way to model this situation is through a Payoff Matrix*, which is a table that tells us the utility of each player's action, depending on the action of the other player. For the Prisoner's Dilemma Game, we can write the payoff matrix as*

$$\begin{pmatrix} (-5,-5) & (0,-20) \\ (-20,0) & (-1,-1) \end{pmatrix} \tag{1}$$

Where, in each element of the matrix, the tuple (u^A, u^B) denotes the payoffs received by player one and two respectively. This idea is also depicted in Figure 1[1]*. Each agent will then try to maximise their payoff, which in this context means minimising their sentence.*

[1] Picture Credits: Vecteezy

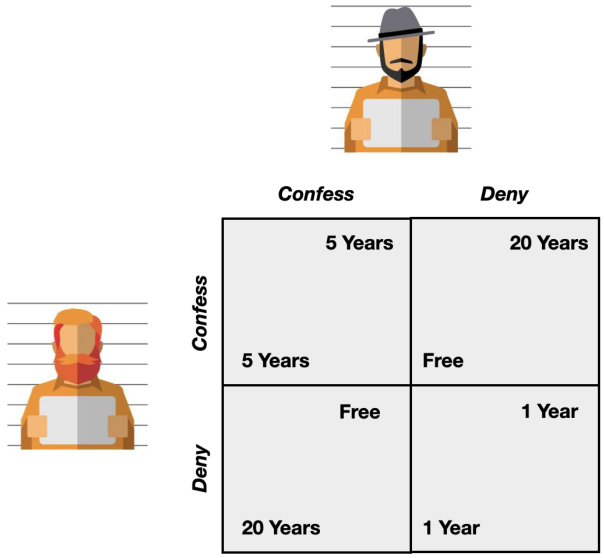


Figure 1: The Prisoner's Dilemma Game in Example 3

Example 4 (Rock-Paper-Scissors)**.** *In our final example, we consider the classic Rock-Paper-Scissors game*[2]*. For the uninitiated, the Rock-Paper-Scissors game is a two player game where each player must select either Rock, Paper, or Scissors as their action. In this format, Rock defeats Scissors, Scissors defeats Paper and Paper defeats Rock. To formalise the notion of 'winning' and 'losing' in this context, we say that the player with the winning action receives a payoff of* 1*, whilst the losing player receives payoff* -1*. If the players pick the same action, the game results in a draw and both players receive* 0 *payoff. So, for instance, if the first player plays rock and the second plays scissors, then the payoff received by both agents is given by the tuple* $(u^A, u^B) = (1, -1)$*, indicating that the first player has won the game. The matrix is then given as*

$$\begin{pmatrix} (0,0),(-1,1),(1,-1) \\ (1,-1),(0,0),(-1,1) \\ (-1,1),(1,-1),(0,0) \end{pmatrix} \tag{2}$$

With the above motivation, we can now write a definition of a game[3].

Definition (Game)**.** An N-player game is the tuple $\Gamma = (\mathcal{N}, (S^\mu, u^\mu)_{\mu\in\mathcal{N}})$, where $\mathcal{N} = \{1, \ldots, N\}$ is the set of all players (also called 'agents'), S^μ is the set of actions available to agent μ and $u^\mu : \times_\mu S^\mu \to \mathbb{R}$ is the payoff function associated to agent μ. Note that this payoff function is dependent on the actions of all players.

[2]We choose to forego the much more complicated Rock-Paper-Scissors-Lizard-Spock game as an avenue for future work

[3]This is actually the definition of a *normal form* game, which is the game that we will mostly be studying. There are other, more complex, games including *extensive form* games and *Bayesian games*, which are interesting in their own right. However, for our purposes we can use the word 'game' to refer exclusively to normal form games

Note that the number of actions (also called strategies) available to each agent μ is given by the cardinality[4] of their associated action set. We denote this as $|S^{\mu}|$

A special type of game which we will be considering is that of a *bimatrix* game, in which there are two players whose payoff functions u^{μ} can be represented through a payoff matrix.

Definition (Bimatrix Game)**.** Let $n := |S^1|$, $m := |S^2|$. Then, a bimatrix game is the tuple (A, B), $A, B \in M_{n\times m}(\mathbb{R})$. Here, $(A)_{ij}$ (resp. $(B)_{ij}$) denotes the payoff received by player one (resp. player two) when player one plays action i and player two plays action j.

Most of Mathematical Game Theory, including the Dynamics of Games, concerns itself with the study of bimatrix games in which $n = m$, although these assumptions are not always necessary.

The final two of our above examples, namely the Prisoner's Dilemma and Rock-Paper-Scissors game, are common objects of study in Game Theory. Both are $n\times n$ bimatrix games with, and we will revisit them often.

2 The Nash Equilibrium

Let us look at the Prisoner's Dilemma in a little more depth. What does game theory predict will be the outcome of this game?

- If the first prisoner thinks that the other is going to confess, then his best option is to confess as well. That way he will only receive 5 years in prison. However, if he thinks the other player is not going to deny then his best option is still to confess, since he will be able to walk free. Either way, he should confess.

- Now what about the second prisoner? She goes through the same thought process and realises that, regardless of player one's action, her best option is to confess. That way she will get the lowest sentence.

- Both players realise that their best option is to confess and so this is the outcome of the game.

This outcome is what is known as a *pure Nash Equilibrium* (NE), after the famous John Forbes Nash. Informally a pure NE is the choice which both agents realise is their best option, and so have no reason to deviate from it. Formally, we have the following definition of a pure NE in a bimatrix game.

Definition (Pure Nash Equilibrium)**.** In a bimatrix game, a pure Nash Equilibrium is a joint strategy $(\bar{s}^1, \bar{s}^2) \in S^1 \times S^2$ such that

$$u^1(s^1, \bar{s}^2) \le u^1(\bar{s}^1, \bar{s}^2) \qquad \forall s^1 \in S^1 \tag{3}$$
$$u^2(\bar{s}^1, s^2) \le u^2(\bar{s}^1, \bar{s}^2) \qquad \forall s^2 \in S^2 \tag{4}$$

[4] number of elements in the set

What is particularly interesting about the Prisoner's Dilemma is that, even though both players would have been better off if they had both denied the charges, since they made their decision as independent, rational agents, they ended up choosing the worse strategy in which they both confess. This situation is an example of the *tragedy of the commons* [6] which, loosely stated is a situation in which, if all agents act according to selfish pursuit of their own interests, the overall result is worse than if they had acted in accordance with the overall social welfare.

2.1 Mixed Nash Equilibria

Is it always the case that a pure Nash Equilibrium exists? The answer to this is a clear and resounding no and can be illustrated through the last of our examples: the Rock-Paper-Scissors game. In this instance if the first agent is to play, for example, rock, then the second agent's optimal choice is to play paper. Similarly, if I play paper, your best response is to play scissors and, of course, we can continue this idea for scissors. As such, there is no single joint strategy of player one and player two such that neither has any incentive to deviate.

However, what if we were to allow agents to randomise their strategy? Consider the following situation

- Player one plays each action with probability $P_1(Rock) = p_1$, $P_1(Paper) = p_2$, $P_1(Scissors) = p_3$.
- Player two plays each action as $P_2(Rock) = q_1$, $P_2(Paper) = q_2$, $P_2(Scissors) = q_3$.

The probability vector $p = (p_1, p_2, p_3)^T$ (resp. q) is called player one's (resp. two's) *mixed strategy* and, of course, must satisfy $p_1 + p_2 + p_3 = 1$. These probability vectors take their values in the unit-simplex, given (for player 1) by $\Delta_1 = \{\mathbf{p} \in \mathbb{R}^n | \sum_i p_i = 1\} \subset \mathbb{R}^n$. The unit simplices in $\mathbb{R}^2$ and $\mathbb{R}^3$ are depicted in Figure 2. Notice that they form $(n-1)-$ dimensional surfaces and so we often only show the projections (a line for $n = 2$ and a triangle for $n = 3$).

Now, we can consider asking what the *expected payoff* received by player two is. This is given by

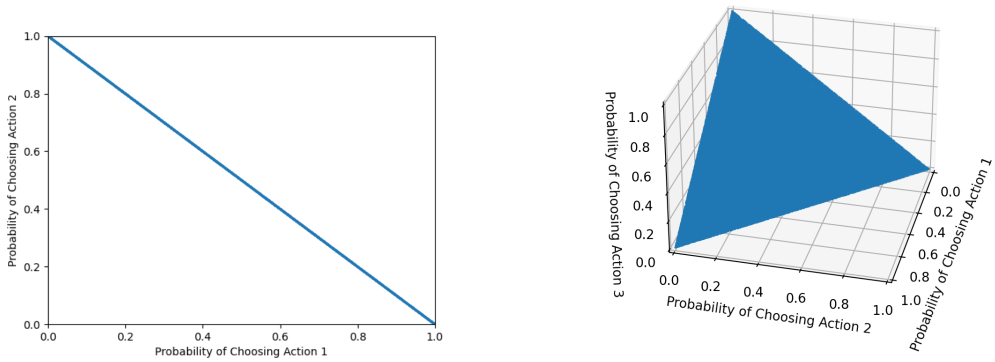
$$\mathbb{E}[u^2(i,j)]_{(p,q)} = \mathbb{E}[(B)_{ij}]_{(p,q)} = \sum_{ij} (B)_{ij} p_i q_j = \mathbf{p} \cdot \mathbf{B}\mathbf{q} \tag{5}$$

Now, a *mixed Nash Equilibrium* is a joint mixed strategy which maximises the agent's expected payoff. Formally we have the following definition.

Definition (Mixed Nash Equilibrium)**.** Consider the bimatrix game $(\mathbf{A}, \mathbf{B})$. A mixed Nash Equilibrium is a joint mixed strategy $(\bar{\mathbf{p}}, \bar{\mathbf{q}}) \in \Delta_1 \times \Delta_2$ such that

$$\mathbf{p} \cdot \mathbf{A}\bar{\mathbf{q}} \leq \bar{\mathbf{p}} \cdot \mathbf{A}\bar{\mathbf{q}} \qquad \forall \mathbf{p} \in \Delta_1 \tag{6}$$
$$\bar{\mathbf{p}} \cdot \mathbf{B}\mathbf{q} \leq \bar{\mathbf{p}} \cdot \mathbf{B}\bar{\mathbf{q}} \qquad \forall \mathbf{q} \in \Delta_2 \tag{7}$$

A mixed NE is called *strict* if the inequality is strict



(a) *The unit simplex in* $\mathbb{R}^2$ (b) *The unit simplex in* $\mathbb{R}^3$

Figure 2: The unit simplices in (a) $\mathbb{R}^2$ *and (b)* $\mathbb{R}^3$*. Since probability vectors have one constraint, they have* $n-1$ *degrees of freedom and so the simplex forms an* $(n-1)$ *dimensional surface.*

Some simple computation, which we leave as an exercise to the reader, allows us to verify that the joint strategy $\bar{\mathbf{p}} = (1/3, 1/3, 1/3)^T$, $\bar{\mathbf{q}} = (1/3, 1/3, 1/3)^T$ is a mixed NE for the Rock-Paper-Scissors game.

We note that a pure strategy i corresponds to the mixed strategy which has 1 in the i'th component and 0 everywhere else. In other words this is the unit vector e_i.

Now we can ask whether a mixed Nash Equilibrium always exists. The following theorem by J.F. Nash shows that this is indeed the case and is the reason why this equilibrium is named in his honour.

Theorem 1 (Nash [5])**.** *Every game with a finite number of players and actions admits at least one mixed Nash equilibrium.*

What is interesting about this theorem is that it's proof is a 'simple' application of a fixed point theorem (specifically Brouwer's fixed point theorem) [1]. We report this proof, along with the prerequisite fixed point theorem, in the Appendix.

Finally, we introduce the following Lemma, which will later be useful in giving the Nash Equilibrium[5] a more dynamical interpretation.

Lemma 1. *In a bimatrix game* $(\mathbf{A}, \mathbf{B})$ *a joint strategy* $\bar{\mathbf{p}}, \bar{\mathbf{q}}$ *is an NE if and only if* $(\mathbf{A}\bar{\mathbf{q}})_i =$ *const. for any* i *such that* $q_i \geq 0$ *and* $(\bar{\mathbf{p}}^T\mathbf{B})_j =$ *const. for any* j *such that* $p_j \geq 0$*. In particular* $(\mathbf{A}\bar{\mathbf{q}})_i = \bar{\mathbf{p}} \cdot \mathbf{A}\bar{\mathbf{q}}$ *and* $(\bar{\mathbf{p}}^T\mathbf{B})_j = \bar{\mathbf{p}} \cdot \mathbf{B}\bar{\mathbf{q}}$.

Proof. We begin by noticing that $(\mathbf{A}\bar{\mathbf{q}})_i = \mathbf{e}_i \cdot \mathbf{A}\bar{\mathbf{q}}$, where $\mathbf{e}_i$ is the unit vector with 1 in the i'th element and 0 elsewhere. In addition, due to the NE condition, we must have that, for all i, $\mathbf{e}_i \cdot \mathbf{A}\bar{\mathbf{q}} \leq \bar{\mathbf{p}} \cdot \mathbf{A}\bar{\mathbf{q}}$. Finally, we have that, since it is a probability vector, $\bar{\mathbf{p}}$ can be written as a convex combination of $\mathbf{e}_i$. In particular, $\bar{\mathbf{p}} = \sum_i \bar{\lambda}_i \mathbf{e}_i$, where $\bar{\lambda}_i \in [0, 1]$ and $\sum_i \bar{\lambda}_i = 1$. Then,

[5] at this point, we refer to mixed NE as the Nash Equilibrium since, as we mentioned, the pure case simply corresponds to a unit vector

$$\bar{\mathbf{p}} \cdot \mathbf{A}\bar{\mathbf{q}} = \left(\sum_i \bar{\lambda}_i \mathbf{e}_i\right) \cdot \mathbf{A}\bar{\mathbf{q}} = \left(\sum_i \bar{\lambda}_i \mathbf{e}_i \cdot \mathbf{A}\bar{\mathbf{q}}\right) \leq \left(\sum_i \bar{\lambda}_i \bar{\mathbf{p}} \cdot \mathbf{A}\bar{\mathbf{q}}\right) = \bar{\mathbf{p}} \cdot \mathbf{A}\bar{\mathbf{q}}$$

For this to be consistent, we cannot have the inequality being strict (since otherwise $\bar{\mathbf{p}} \cdot \mathbf{A}\bar{\mathbf{q}} < \bar{\mathbf{p}} \cdot \mathbf{A}\bar{\mathbf{q}}$ which does not make sense). Therefore, we must have that $\mathbf{e}_i \cdot \mathbf{A}\bar{\mathbf{q}} = \bar{\mathbf{p}} \cdot \mathbf{A}\bar{\mathbf{q}}$ for all i. The analagous argument also holds for $(\bar{\mathbf{p}}^T\mathbf{B})_j$.

□

3 Best Responses

An equivalent way in which to define the Nash equilibrium is through the so called *best response correspondence*. To motivate this, let us note that the definition of Nash Equilibrium can be written in the following manner:

$$\mathbf{p} \cdot \mathbf{A}\bar{\mathbf{q}} \leq \bar{\mathbf{p}} \cdot \mathbf{A}\bar{\mathbf{q}} \quad \forall \mathbf{p} \in \Delta_1 \implies \bar{\mathbf{p}} \in \arg\max_{\mathbf{p} \in \Delta_1} \mathbf{p} \cdot \mathbf{A}\bar{\mathbf{q}}$$
$$\bar{\mathbf{p}} \cdot \mathbf{B}\mathbf{q} \leq \bar{\mathbf{p}} \cdot \mathbf{B}\bar{\mathbf{q}} \quad \forall \mathbf{q} \in \Delta_2 \implies \bar{\mathbf{q}} \in \arg\max_{\mathbf{q} \in \Delta_2} \bar{\mathbf{p}} \cdot \mathbf{B}\mathbf{q}$$

With this in place, we define the best response correspondence as the set valued map which, for each player, looks at the strategy of the opponent and assigns the mixed strategy which maximises the player's payoff. More formally, we have the following definition

Definition (Best Response Correspondence)**.** Let $(\mathbf{A}, \mathbf{B})$ be a bimatrix game and let $\mathbf{P}(\Delta)$ denote the power set of Δ. Then $BR^A : \Delta_2 \to \mathbf{P}(\Delta_1)$, $BR^B : \Delta_1 \to \mathbf{P}(\Delta_2)$ are such that, for any $(\mathbf{p}, \mathbf{q}) \in \Delta_1 \times \Delta_2$

$$BR^A(q) := \arg\max_{\mathbf{p} \in \Delta_1} \mathbf{p} \cdot \mathbf{A}\mathbf{q} \qquad\qquad BR^B(p) \arg\max_{\mathbf{q} \in \Delta_2} \mathbf{p} \cdot \mathbf{B}\mathbf{q}$$

Example 5. *Consider the bimatrix game* $(\mathbf{A}, \mathbf{B})$ *defined by*

$$\mathbf{A} = \begin{pmatrix} 1 & 0 & 0.5 \\ 0.5 & 1 & 0 \\ 0 & 0.5 & 1 \end{pmatrix} \qquad \mathbf{B} = \begin{pmatrix} -0.5 & 1 & 0 \\ 0 & -0.5 & 1 \\ 1 & 0 & -0.5 \end{pmatrix} \tag{8}$$

This game is part of a family of bimatrix games known as the 'Shapley' family (after Nobel Prize winner Lloyd Shapley) and we will revisit it in our exposition of the Best Response/Fictitious Play dynamics.

Let us assume that the second player is playing the mixed strategy $(0.5, 0.3, 0.2)^T$. *What would player one's best response be?*

$$
\begin{aligned}
BR^A((0.5,0.3,0.2)^T) &= \arg\max_{\mathbf{p}\in\Delta_1} \mathbf{p}\cdot\begin{pmatrix}1 & 0 & 0.5\\ 0.5 & 1 & 0\\ 0 & 0.5 & 1\end{pmatrix}\begin{pmatrix}0.5\\0.3\\0.2\end{pmatrix}\\
&= \arg\max_{\mathbf{p}\in\Delta_1} \mathbf{p}\cdot\begin{pmatrix}0.6\\0.55\\0.35\end{pmatrix}\\
&= \begin{pmatrix}1\\0\\0\end{pmatrix}\\
&= \mathbf{e}_1
\end{aligned}
$$

With this in mind, we can write an alternate definition of the Nash Equilibrium in terms of the best response correspondence.

Definition (Alternate Definition of Nash Equilibrium)**.** Let $(\mathbf{A}, \mathbf{B})$ be a bimatrix game. A joint mixed strategy $(\bar{\mathbf{p}}, \bar{\mathbf{q}}) \in \Delta_1 \times \Delta_2$ is a Nash equilibrium if

$$
\bar{\mathbf{p}} \in BR^A(\bar{\mathbf{q}}) \qquad\qquad \bar{\mathbf{q}} \in BR^B(\bar{\mathbf{p}}) \tag{9}
$$

The best response correspondences are almost everywhere single valued. In particular, it will almost always give the singleton $\{e_i\}$ where e_i is one of the unit vectors corresponding to a pure strategy. In the case that it takes on multiple values, it is the convex combination of a subset of all the unit vectors, which could include the simplex itself.

Example 6. *Let us reconsider the bimatrix game from Example 5. We already saw that the best response of player one to* $\mathbf{q} = (0.5, 0.3, 0.2)^T$ *is* $\mathbf{e}_1$*. Let us generalise this result*

$$
\mathbf{A}\mathbf{q} = \begin{pmatrix}1 & 0 & 0.5\\ 0.5 & 1 & 0\\ 0 & 0.5 & 1\end{pmatrix}\begin{pmatrix}q_1\\q_2\\q_3\end{pmatrix} = \begin{pmatrix}q_1 + 0.5q_3\\ 0.5q_1 + q_2\\ 0.5q_2 + q_3\end{pmatrix} \tag{10}
$$

So, then, if $q_1 > q_2$ *and* $q_1 > q_3$*, then player one's best response is* $\mathbf{e}_1$*. Similarly, if* q_2 *dominates* q_1 *and* q_3 *then the best response is* $\mathbf{e}_2$*. We can do a similar process for player two's strategy.*

$$
\mathbf{p}^T\mathbf{B} = \begin{pmatrix}p_1 & p_2 & p_3\end{pmatrix}\begin{pmatrix}-0.5 & 1 & 0\\ 0 & -0.5 & 1\\ 1 & 0 & -0.5\end{pmatrix} = \begin{pmatrix}p_2 - 0.5p_1\\ p_1 - 0.5p_2\\ p_2 - 0.5p_3\end{pmatrix} \tag{11}
$$

So that if $p_2 > p_1$ *and* $p_2 > p_3$ *then player two's best response is to play* $\mathbf{e}_3$ *and so on.*

We can plot the values of $\mathbf{p}$ *for which player two's best response is any* $\mathbf{e}_i$ *and similarly the values of* $\mathbf{q}$ *for which player one's best response is* $\mathbf{e}_j$*. This is shown in Figure*

Let us consider the case in which player two plays $\mathbf{q} = (1/3, 1/3, 1/3)^T$*. Then it can be seen that player one's has best response*

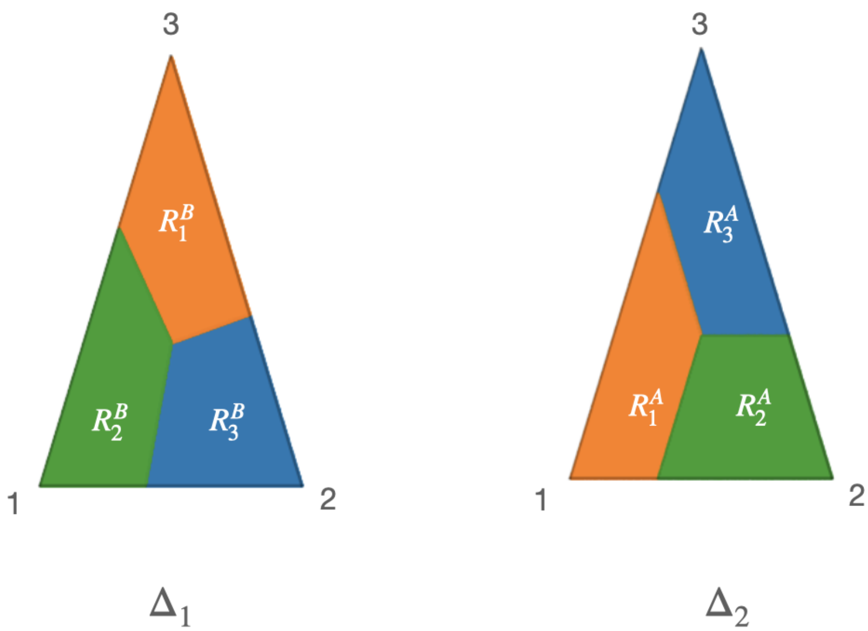


Figure 3: Division of the simplices Δ_1, Δ_2 for Example 6. In Δ_1, the blue portion corresponds to those values for $\mathbf{p}$ for which player two's best response is $\mathbf{e}_3$. The orange portion corresponds to those for which two's best response is $\mathbf{e}_1$ and finally if $\mathbf{p}$ lies in the green region then two's best response is $\mathbf{e}_2$. The same is true for the division of Δ_2.

$$\begin{pmatrix}1/3\\1/3\\1/3\end{pmatrix} \in \arg\max_{\mathbf{p}\in\Delta_2} \begin{pmatrix}0.5\\0.5\\0.5\end{pmatrix} \tag{12}$$

And similarly for player two if one plays $\mathbf{p} = (1/3, 1/3, 1/3)^T$. By Definition 3 the joint strategy $(\bar{\mathbf{p}}, \bar{\mathbf{q}}) = ((1/3, 1/3, 1/3)^T, (1/3, 1/3, 1/3)^T)$ is a Nash Equilibrium of $(\mathbf{A}, \mathbf{B})$.

As we saw from the above example, we can break up the simplices Δ_1 and Δ_2 according to the values that the best response map takes. In particular we define [7]

$$\begin{aligned} R_j^B &:= (BR^B)^{-1}(e_j) \subset \Delta_1 \qquad \text{for } j = 1, \ldots, n \\ R_i^A &:= (BR^A)^{-1}(e_i) \subset \Delta_2 \qquad \text{for } i = 1, \ldots, m. \end{aligned}$$

R_i^A denotes the *preference region* of i in Δ_2. If player two's mixed strategy takes its value in R_i^A, then player one's best response is e_i. Now, the intersections between two preference regions R_i^A and R_j^A correspond to the case in which the BR_A yields two pure strategies, e_i and e_j. This intersection is referred to as an *indifference plane* [7], due to the fact that the payoff received by player one is equal if they were to choose either e_i or e_j (or, indeed, any of their convex combinations). Formally it is given as

$$Z_{ij}^A := R_i^A \cap R_j^A = \{\mathbf{q} \in \Delta_2 | (\mathbf{A}\mathbf{q})_i = (\mathbf{A}\mathbf{q})_j \geq (\mathbf{A}\mathbf{q})_k \;\; \forall k = 1, \ldots, m\} \subset \Delta_2 \tag{13}$$

Similarly, for player two's indifference planes we have

$$Z_{ij}^B := R_i^B \cap R_j^B = \{\mathbf{p} \in \Delta_1 | (\mathbf{p}^T\mathbf{A})_i = (\mathbf{p}^T\mathbf{A})_j \geq (\mathbf{p}^T\mathbf{A})_k \ \forall k = 1, \ldots, n\} \subset \Delta_1 \tag{14}$$

Finally, using Lemma 1 we can state the following

Corollary 1. *In a bimatrix game* $(\mathbf{A}, \mathbf{B})$*, the point* $(\bar{\mathbf{p}}, \bar{\mathbf{q}}) \in int\Delta_1 \times \Delta_2$ *is an NE if and only if* $\bar{\mathbf{p}}$ *lies on the intersection of all preference regions in* Δ_1 *and* $\bar{\mathbf{q}}$ *lies on the intersection of all preference regions in* Δ_2*. Equivalently,* $(\bar{\mathbf{p}}, \bar{\mathbf{q}})$ *is an NE if and only if* $\bar{\mathbf{p}}$ *lies on the intersection of all indifference planes in* Δ_1 *and* $\bar{\mathbf{q}}$ *lies on the intersection of all indifference planes in* Δ_2*.*

4 The Case for Learning

Let us look back at the Prisoner's Dilemma and consider what it took for both parties to decide that they would confess. Both players looked at the possible actions that the other agent had available and evaluated, given either action of the other agent, what their payoff maximising strategy would be. This requires that, in the heat of the moment, an agent should have a complete understanding of how their actions, in conjunction with those of their neighbours affects their payoff. Further, it requires that the agent always makes a rational decision based on this information. In fact, in the appendix we discuss the celebrated *minimax theorem* of John von Neumann which, in some sense, illustrates how the agents view the game when playing an NE.

We do not need to look far to realise that, in the real world, individuals do not satisfy these requirements. Neither are we always aware of the actions available to us, nor do we always make the most rational decision. In many ways, then, the concept of the Nash Equilibria requires what Susan Wolf refers to as the *moral saint* - a character whose "every action is as morally good as possible, that is, who is as morally worthy as can be". This ideal is rarely achieved in practice and, therefore, it can be argued that the NE is not an accurate way in which to assess the behaviour of the economy, or of population ecology.

What about computers? Whilst we might say that artificial agents can be considered as purely rational agents, and that the payoffs of each agent are typically provided by the designers, it would still require that we compute the NE. Typically finding the NE is achieved by framing the problem as a linear program (see [6] for examples). However, the task of finding the NE in general N-player, n-action games is not always so straightforward. The interested reader should consider reading the classical book 'Algorithmic Game Theory' [6] in which the authors spend some time discussing the fact that computing the NE is *PPAD-complete* (a concept which we will not discuss here, c.f. [6]), a class of algorithms which are considered 'hard' to solve, although it is unclear whether it falls into the class of polynomial time algorithms or into the NP class.

So neither is the NE an easy quantity to compute, nor is it necessarily achieved in practice. As such, we must look for a new way to analyse multi-agent interactions, as well as design tractable computational solutions. For this, we turn to nature as our guide. Specifically, we know that people base their actions on prior experience, and sometimes a prospective estimation on the consequences of their actions. This motivates the need for learning algorithms.

Learning algorithms model the evolution of how agents choose their actions in their interactions. In particular, they show how, based on their memory of playing a game in the past, an agent will adapt the probabilities with which they favour certain actions over others. In particular, the learning algorithms we will consider are

I The Replicator Dynamic

II Best Response and Fictitious Play

III Reinforcement Learning

This is, by no means, an exhaustive list of all learning algorithms. Rather, these are chosen based on our own, limited, knowledge of the vast array of learning algorithms available. The interested reader should also consult [4, 8], which give an insight into various other learning algorithms.

We will focus on the techniques which are used to analyse these systems from the view of Dynamical Systems and Statistical Physics. Again, these are not the only lenses through which we can view learning, and we hope that the reader will be inspired to understand these same problems from other perspectives. We also hope that the reader will go on to bring their own expertise to bear on the many open problems in this incredibly exciting and active area of research.

5 Appendix

Existence of the Nash Equilibrium

As mentioned, the proof of the existence of the NE relies on Brouwer's fixed point theorem. This is a standard result from topology and so we do not go into the full details here.

Theorem 2 (Brouwer Fixed Point Theorem)**.** *Let $\mathcal{X} \subset \mathbb{R}^n$ be convex and compact. If $T : \mathcal{X} \to \mathcal{X}$ is continuous, then it admits a fixed point, i.e. a point $x \in \mathcal{X}$ such that $T(x) = x$.*

Proof. We prove this theorem for the one-dimensional case as it is a little simpler than the general case. Since our set $\mathcal{X}$ is in $\mathbb{R}$ and is compact, we can write it as a closed and bounded interval $[a, b]$ and $T : [a, b] \to [a, b]$. Our goal is to show that a fixed point of T exists in this interval. First, if either $T(a) = a$ or $T(b) = b$, then we are done. Now let us look at the case where the fixed point is neither at a or b. Then we can say that $T(a) > a$ and $T(b) < b$.

Consider the function $g(x) = T(x) - x$. Then $g(a) > 0$ and $g(b) < 0$ and g is continuous, since T is. Then, by the Intermediate Value Theorem, there is an $x \in [a, b]$ such that $g(x) = 0$, and so $T(x) = x$. □

Theorem 3 (Existence of the Nash Equilibrium)**.** *Every game with a finite number of players and actions admits at least one mixed Nash equilibrium.*

Proof. This proof works roughly as follows, first we define a continuous function which acts on the product space of all agents simplices (which is compact). Therefore, by Brouwer's Theorem, this function must have a fixed point. We then show that, at this fixed point, the NE condition is satisfied. Therefore, the game must admit an NE.

To recall our notation, a game consists of the set of players $\mathcal{N}$ and their set of actions (resp. payoff function) S^μ (resp. u^μ) $\forall \mu \in \mathcal{N}$. Let us denote the concatenation of all agents' mixed strategies x^μ as the vector $\mathbf{x} = (x^1, \ldots, x^N)$.

Then, for some $\mathbf{x} \in \Delta := \times_\mu \Delta_\mu$ and $s^\mu \in S^\mu$, define the function

$$G^\mu_{s^\mu}(\mathbf{x}) := \max\{u^\mu(s^\mu, x_{-\mu}) - u^\mu(\mathbf{x}), 0\}.$$

We can consider this to be the 'incentive' that agent μ has from deviating from the joint mixed strategy $\mathbf{x}$. In particular, if μ would receive a higher payoff by playing s^μ rather than x^μ, then $G^\mu_{s^\mu}(\mathbf{x}) = u^\mu(s^\mu, x_{-\mu}) - u^\mu(\mathbf{x}) > 0$.

Now define the function $f : \Delta \to \Delta$ through

$$f^\mu_{s^\mu}(\mathbf{x}) = \frac{x^\mu + G^\mu_{s^\mu}(\mathbf{x})}{1 + \sum_j G^\mu_j(\mathbf{x})} \tag{15}$$

This is a continuous function acting on a compact space Δ, so by Brouwer's Fixed Point Theorem, it must have a fixed point. Let us call this $\bar{\mathbf{x}}$.

Now we claim that $\sum_j G^\mu_j(\mathbf{x}) = 0$. To show this, let us assume for contradiction that $\sum_j G^\mu_j(\mathbf{x}) > 0$ (it cannot be less than 0 due to the max operation). Now, if we multiply across the fraction in (15), we get

$$\begin{aligned} f^\mu_{s^\mu}(\bar{\mathbf{x}})(1 + \sum_j G^\mu_j(\bar{\mathbf{x}})) &= \bar{x^\mu} + G^\mu_{s^\mu}(\bar{\mathbf{x}}) \\ x^\mu(1 + \sum_j G^\mu_j(\bar{\mathbf{x}})) &= \bar{x^\mu} + G^\mu_{s^\mu}(\bar{\mathbf{x}}) \\ x^\mu \sum_j G^\mu_j(\bar{\mathbf{x}}) &= G^\mu_{s^\mu}(\bar{\mathbf{x}}) \end{aligned}$$

Now, this means that if $G^\mu_{s^\mu}(\bar{\mathbf{x}}) = 0$, then we must have that $x^\mu = 0$ (since we assumed $\sum_{t^\mu} G^\mu_j(\bar{\mathbf{x}}) > 0$). In particular, this means that we can define the set $\mathcal{I}_\mu := \{i : \bar{x^\mu_i} > 0\}$ and say that it is contained in the set $\{i : G^\mu_i(\bar{\mathbf{x}}) > 0\}$. As such, we have that $\sum_{i=1}^n \bar{x^\mu_i} = \sum_{i \in \mathcal{I}_\mu} \bar{x^\mu_i} = 1$, where n is the total number of actions for agent μ.

This leads to the result

$$\begin{aligned} G^\mu_i(\bar{\mathbf{x}}) > 0 &\implies u^\mu(s^\mu, \bar{x}_{-\mu}) > u^\mu(\bar{\mathbf{x}}) \\ &\implies \bar{x^\mu_i} u^\mu(s^\mu, \bar{x}_{-\mu}) > \bar{x^\mu_i} u^\mu(\bar{\mathbf{x}}) \\ &\implies u^\mu(\bar{\mathbf{x}}) = \sum_{i=1}^n \bar{x^\mu_i} u^\mu(s^\mu, \bar{x}_{-\mu}) \geq \sum_{i \in \mathcal{I}_\mu} \bar{x^\mu_i} u^\mu(s^\mu, \bar{x}_{-\mu}) > \sum_{i \in \mathcal{I}_\mu} \bar{x^\mu_i} u^\mu(\bar{\mathbf{x}}) = u^\mu(\bar{\mathbf{x}}) \end{aligned}$$

which is clearly a contradiction. This means our initial assumption was wrong and that we must have $\sum_j G^{\mu}_j(\mathbf{x}) = 0$. Again, due to the max operation, we cannot have negative values in $G^{\mu}_j(\mathbf{x})$, so we must have that, for all $j \in S^{\mu}$, $G^{\mu}_j(\mathbf{x}) = 0$ which is equivalent to the statement that

$$u^{\mu}(j, \bar{x}_{-\mu}) \leq u^{\mu}(\bar{\mathbf{x}}) \qquad \forall j. \tag{16}$$

This argument was independent of our choice of agent μ and so holds for all agents. This was precisely the definition of the Nash Equilibrium.

□

The Minimax Theorem

TO DO

References

[1] Border, K. C. *Fixed Point Theorems with Applications to Economics and Game Theory*. Cambridge University Press, 4 1985.

[2] Coolen, A. C. C. *The Mathematical Theory of Minority Games: Statistical Mechanics of Interacting Agents (Oxford Finance Series)*. Oxford University Press, Inc., USA, 2005.

[3] Güth, W., Kocher, M. G., Schiller, F., and Jena, U. More than thirty years of ultimatum bargaining experiments: Motives, variations, and a survey of the recent literature More than thirty years of ultimatum bargaining experiments: Motives, variations, and a survey of the recent literature*.

[4] Hofbauer, J., and Sigmund, K. EVOLUTIONARY GAME DYNAMICS. Tech. Rep. 4, 2003.

[5] Nash, J. F. Equilibrium points in n-person games. *Proceedings of the National Academy of Sciences 36*, 1 (1 1950), 48–49.

[6] Nisan, N., Roughgarden, T., Tardos, , and Vazirani, V. V. *Algorithmic game theory*, vol. 9780521872829. Cambridge University Press, 1 2007.

[7] Ostrovski, G. Topics arising from fictitious play dynamics.

[8] Pangallo, M., Sanders, J. B. T., Galla, T., and Farmer, J. D. A taxonomy of learning dynamics in 2 × 2 games. Tech. rep., 2020.

[9] Weber, B., Miyapuram, K. P., Kahnt, T., Vavra, P., Chang, L. J., and Sanfey, A. G. Expectations in the Ultimatum Game: Distinct Effects of Mean and Variance of Expected Offers. *Frontiers in Psychology — www.frontiersin.org 1* (2018), 992.