Contents

Chapter 3

Repetition, Reputation, and Reciprocity

> *A young boy enters a barber shop and the barber whispers to his customer, "This is the dumbest kid in the world. Watch while I prove it to you." The barber puts a dollar bill in one hand and two quarters in the other, then calls the boy over and asks, "Which do you want, son?" The boy takes the quarters and leaves. "What did I tell you?" said the barber. "That kid never learns!"*
>
> *Later, when the customer leaves, he sees the same young boy coming out of the ice cream store. "Hey, son! May I ask you a question? Why did you take the quarters instead of the dollar bill?"*
>
> *The boy licked his cone and replied, "Because the day I take the dollar, the game is over!"*

3.1 Repeated Interactions

A primary rationale for politeness is that many interactions are bound to be repeated. We see our family members more than once in our lives, we speak with our colleagues daily, etc. Even in cases where there is little chance that we may see someone again, we may nonetheless behave as such for other reasons. This principle will figure in later chapters like **??**. The repetition of interaction can form the basis for a variety of relationships, and as such there are several notions from the body of workMailath and Samuelson [2006] on repeated games that we should consider:

- The difference between indefinite and infinite repetition
- The discount factor δ measuring the degree to which an agent cares about future interactions
- The probability of a future interaction
- An agent's reputation based on his past history
- The emergent relationship based on a history of interactions

· The impact of monitoring/ observation on repeated interaction

3.2 Actions, Histories, and Strategies

Within every interaction with others, we often have familiar choices based on the social context. If that interaction is ongoing, the repeated nature of the interaction might make us choose our actions differently. These choices might also be affected by our patience with our partner in the interaction or the likelihood that we will meet again, factors external to the circumstances governing the interaction itself, e.g. whether the interaction is antagonistic or cooperative. For instance, paying for public transit is a cooperative endeavor as everyone benefits despite the individual reasons against paying. One rationale for not paying the fee might be that the passenger is a tourist who will not benefit from living in that city later, a phenomenon more generally known as *free-riding*[Mailath and Samuelson, 2006]. [1] This gives us a frame for thinking about stage games within repeated games.

We consider familiar extensive and normal form games within the larger context of a repeated interaction. A *stage game* is one of these familiar examples. In the case of public transit, we might think of a Prisoner's Dilemma as the stage game that would be repeated among passengers of the service. The repeated game is, intuitively, the combinations of choices and payoffs available to players who play according to the stage game repeatedly. We now give a brief introduction, based on the more technical work in Mailath and Samuelson [2006] and course notes in Ratliff [1997].

We would like to distinguish between *actions* available in the stage game and *strategies* available in the repeated game. If we first consider that a stage game has a set of actions available to a player, like {C,D} in the Prisoner's Dilemma, then at a given stage t in the repeated game, we have an *action profile* $a^t = (a_i^t, a_j^t)$ for players i and j. From these we can build a history of the game.

Histories

As an example, a repeated instance of the Prisoner's Dilemma might look like

$$< (C, C), (D, C), (D, D), (C, D), \ldots >$$

The sequence of action profiles in the first three rounds is what we call the *history* $h_3 =< (C, C), (D, C), (D, D) >$. More formally, assuming action space A with actions $a \in A$ and each player i's action in round n as a_i^n, we define histories:

Definition 1. *For players X and Y, we define the **history** of a repeated game to be the sequence of action profiles $h =< (a_X^1, a_Y^1), (a_X^2, a_Y^2), (a_X^3, a_Y^3), \ldots >$.*

We can also abbreviate the first t action profiles to give the truncated *history* as: $h_t = (a^0, a^1, a^2, a^3, \ldots, a^{t-1})$. From each history, a player has a *strategy* which determines his subsequent move given a history, seen in Figure 3.1.

[1] This type of failure to cooperate most likely inspired the term.

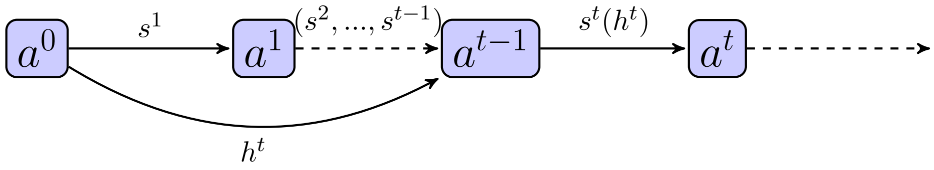

Figure 3.1: Diagram depicting (infinite) repetition of the game G, where each nth action profile is given by a^n. Note that the strategies map histories in the previous rounds to actions in the next round.

Strategies

In each stage of a repeated game, players make choices as to their respective actions, assigning a pair of strategies to each round of play. These can be based on previous encounters, and we will often take that into consideration. The strategies need not be chosen on a rational basis, but we will also appeal to some other notions that will allow for a strategic analysis of repetition. More formally,

Definition 2. *A **strategy** in a repeated game is a function s_i^t such that $s_i^t(h^t) = a_i^t$*

With this definition of strategy in place, notice now that we can think of a game's strategy profile in two ways: according to the players or to the time index. At a given time index t in a game with T rounds in total, we have $s^t = (s_i^t, s_j^t)$, and for a given player i, we have $s_i = (s_i^0, \dots, s_i^T)$. This gives us now two nice expressions for a repeated game's strategic profile s:

$$
\begin{aligned}
s =& (s_i, s_j) \\
s =& (s^0, s^1, \dots, s^T)
\end{aligned}
$$

These notions of histories and strategies will allow us to compute the respective utilities for each player in the repeated game. As we progress, we will have to be careful as to how we consider the mechanisms of repetition and how the players evaluate their utility within the game.

3.3 Utility in Repeated Games

As we repeat a game, notice that in each round we have a set of values for the utility of each player. Let us continue with the repeated Prisoner's Dilemma and an example of one its most famous strategies. Consider a game where one player always defects, and the other cooperates until defected on, and then defects from then out. The second player is playing the *Grim Trigger Strategy.* Such a history would look like this :

$$< (C, D), (D, D), (D, D), (D, D), \dots >$$

An initial round of (C, D) would have subsequent rounds of (D, D). [2] To

[2] We can also write this as $< CD, DD, DD, DD, \dots >$ for brevity's sake.

these profiles, we then can compute the utility in each round, given by the stage game in Table 3.1.

	C	D
C	2,2	0,3
D	3,0	1,1

Table 3.1: Version of the Prisoner's Dilemma.

This repeated game would have utilities of $<(0,3),(1,1),(1,1),(1,1),\ldots>$, and for players A and B be given by

$$U_X = 0+1+1+\ldots = 0+\sum_{j=0}^{\infty} 1$$
$$U_Y = 3+1+1+\ldots = 3+\sum_{j=0}^{\infty} 1$$

This is unfortunate, as both X and Y have infinite utilities, and yet it seems intuitive that Y has come out ahead in the strategic sense. One remedy is to adjust the ways in which we value future interactions. Either it needs to be the case that the game will not necessarily go on forever (indefinite repetition) or that our future rewards don't count as much as our current ones (infinite repetition with discounting). We begin with the notation:

Definition 3. *We denote the utility of a player X in the n^{th} stage of a repeated game as U^n. Thus, $U_X = \sum_{j=0} U_X^j$.*

This notation should not be confused with an exponent; rather it is merely a label. For instance, in our example above we see that $U_X^1 = 0$ and $U_X^2 = 1$ whereas $U_Y^1 = 3$ and $U_Y^2 = 1$. It should also follow that $EU_X = \sum_{j=0} EU_X^j$ for expected utilities. Notice one problem with this, however. This definition would give us an unbounded sum for our utilities and thus make it difficult to compare repeated interactions to one another or to one-shot games.

A primary motivation in this dissertation is finding ways in which the cooperative dilemmas in politeness give way to more amenable coordination problems and therefore linguistic conventions. As a first example, we will show how a repeated Prisoner's Dilemma can evolve into a Stag Hunt through *the shadow of the future* and two strategies. Before we get there however, we require some more terminology and mathematical background.

Indefinite vs. Infinite Repetition

We should make clear the difference between indefinite repetition and infinite repetition of games. *Indefinite repetition* occurs when the participants do not know whether the interaction (game) will occur again. They assign a probability $0 < p < 1$ to this occurrence. *Infinite repetition* occurs when the participants do know the interaction (game) will occur again and normalize their utilities based on a discount factor $\delta \in [0, 1]$ to which they value these future interactions. They therefore assign a probability $p = 1$ to this occurrence.

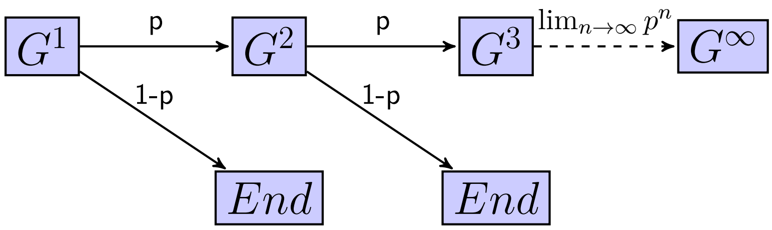

Figure 3.2: Markov Chain depicting indefinite repetition of the game G, where each nth round of the game is given by G^n

An interesting connection between the two exists. We can observe that the external circumstances of indefinite repetition and the internal preferences of patience, or discounting, can give us strategically equivalent games. If p is the probability of future interaction and δ is the discount factor, we can combine them.

In the case of indefinite repetition, we have players who value each potential round as much as the next, so $\delta = 1, p < 1$. In infinite repetition, we have a game that will continue to repeat with no end in sight; thus we have players that discount the future interactions $p = 1, \delta < 1$. Consider the ratio then of a player's expected utility from one round to another, r. We can think of

$$r = p \times \delta$$

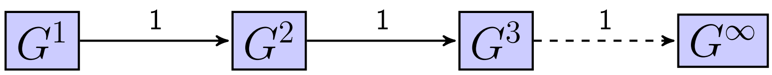

Figure 3.3: Markov Chain depicting infinite repetition of the game G, where each nth round of the game is given by G^n. The probability of repetition is $p = 1$ but each successive round of play is multiplied by a discount δ for $0 < \delta < 1$.

This gives us a way to compare the calculations of the two utilities. Per the derivation and examples in the next few pages, we will see that these types of games differ only by a constant in their utilities, and thus they are strategically equivalent.

We can now calculate utility in the repeated game using infinite series. Consider a player's future interactions within a game, where he evaluates his future return at a lower rate than his present. He may have a sequence according to the utility function V as $v_i^\infty = v_i^0, v_i^1, v_i^2, \ldots$. We first begin with the following observation:

$$0 < r < 1 \Rightarrow 1 + r + \ldots = \frac{1}{1-r}$$

To see why, consider the product $(r-1) \times (1 + r + \ldots + r^n)$:

$$\begin{aligned}
r(1 + r + \ldots + r^n) &= \qquad r + r^2 + \ldots + r^n + r^{n+1} \\
-1(1 + r + \ldots + r^n) &= -1 - r - r^2 - \ldots - r^n \\
\Rightarrow (r-1)(1 + r + \ldots + r^n) &= r^{n+1} - 1
\end{aligned}$$

Thus we have the following result:

$$\sum_{k=0}^{n} r^k = 1 + r + \ldots + r^n = \frac{r^{n+1} - 1}{r - 1}$$

Notice two things: first, we would like to use this to calculate expected utilities in repeated games, and as such, we would like to think of r as a probability, patience level, or combination of both. In general, we can simply think of r as the *ratio* of stage-game utility in the next round of the repeated game to the utility in the current round. If r is a probability (for instance the likelihood that a game will be repeated), we have the following:

$$0 < r < 1 \Rightarrow \lim_{n \to \infty} r^n = 0$$

With that said, for $0 < r < 1$ we now have

$$\sum_{j=0}^{\infty} r^j = 1 + r + r^2 + \ldots = \lim_{n \to \infty} \frac{r^{n+1} - 1}{r - 1} = \frac{-1}{r - 1} = \frac{1}{1 - r}$$

For example, if we are given the series $.7 + .7^2 + .7^3 + \ldots$, we can rewrite this as $.7(1 + .7 + \ldots) = .7\frac{1}{1-.7} = \frac{.7}{.3} = \frac{7}{3}$. With this background in mind, we can construct a utility function v for indefinite repetition. Given a probability of future interaction p and an action profile (a_i^t, a_j^t) in the *t-th* round of play, a rational player should seek to maximize

$$S_i = \sum_{t=0}^{\infty} p^t v_i(a_i^t, a_j^t) \tag{3.1}$$

Discounting Future Interactions

Another way in which we can evaluate our expected return on future actions is known as *discounting*, the notion that future rewards are not as valuable as current ones. When we first began this section, we noted that the sample utilities seen in an infinite Prisoner's Dilemma for *Always Defect* vs. *Grim Trigger* could resemble:

$$U_X = 0 + 1 + 1 + \ldots = 0 + \sum_{j=0}^{\infty} 1$$
$$U_Y = 3 + 1 + 1 + \ldots = 3 + \sum_{j=0}^{\infty} 1$$

At this point, the agents value each interaction as much as the next, and thus they arrive at infinite utilities, something highly inconvenient for analysis and equally unrealistic for application [Camerer, 2003]. We thus want to highlight two critical threads in the literature on repeated games: the means by which agents evaluate future interactions and the resulting relationships derived from those means. To do this, we will explore discounting and the subsequent folk theorem seen in 3.4.

Definition 4. *A* ***discount*** $0 < \delta < 1$ *value is the weight attached to future interactions.*

Discounting is a measure of a player's patience. For two discount values δ_1 and δ_2, we say that a player discounting at a rate of δ_2 is *more patient* if $\delta_2 > \delta_1$. To see this in more detail, we will consider two discount values of $\delta_1 = .6$ and $\delta_2 = .8$ in the table seen in 3.2. As we can see, the higher discount value means that the utilities of each round of play will also be higher. I.e. the agent with the higher discount parameter will value the exchanges of later rounds at more than one with a lower discount parameter. For discount parameters sufficiently large, we can also attain equilibria that favor strategies other than the NE of the original game. This requires delving deeper into the interaction between discounting and repeated games.

n	1	2	3	$\ldots$
δ_1^n	$.6^1$	$.6^2$	$.6^3$	
$U_A^n(\delta_1)$	$.6^1(0)$	$.6^2(1)$	$.6^3(1)$	
δ_2^n	$.8^1$	$.8^2$	$.8^3$	
$U_A^n(\delta_2)$	$.8^1(0)$	$.8^2(1)$	$.8^3(1)$	

Table 3.2: Stage game utilities with $\delta_1 = .6$ and $\delta_2 = .8$ in repeated Prisoner's Dilemma.

To understand how discounting plays in, recall the distinction between stage games and repeated games. A *stage game* is a step in a repeated game that is strategically equivalent to a one-shot game under the assumption of no repetition. Truncations of repeated games where we remove previous histories are known as *subgames*. For instance, we could truncate the first two rounds of play from $h =< (a_X^1, a_Y^1), (a_X^2, a_Y^2), (a_X^3, a_Y^3), \ldots >$ and have:

$$< (a_X^3, a_Y^3), (a_X^4, a_Y^4), (a_X^5, a_Y^5), \ldots >$$

Given these concepts and definitions, there are several points to make. First, notice that the utilities listed above where each agent values the subsequent utilities equally corresponds to a discount value $\delta = 1$. We are concerned with the case where $0 < \delta < 1$. A rational player in a repeated game should seek to maximize the *normalized sum S of discounted payoffs:*

$$S_i = (1 - \delta) \sum_{t=0}^{\infty} \delta^t v_i(a_i^t, a_j^t) \tag{3.2}$$

Again some unfortunate notation, as we should distinguish between what happens in round t from raising a number to the t power. Notice also that there is the potential for variation in discounting as [Rubinstein, 2003]suggests, although we will limit our discussion here to the simpler case where each player has a uniform and identical discount parameter, as opposed to varying and variable discount functions. Last, we can look back at Equation 3.3 and see that this utility function is identical except for the constant $(1 - \delta)$ in front, giving us the desired strategic equivalence.

To see why the definition of a normalized payoff makes sense, observe first that as $0 < \delta < 1$, we have

$$\sum_{t=0}^{\infty} \delta^t = \frac{1}{1-\delta} \quad \Rightarrow \quad (1-\delta) \sum_{t=0}^{\infty} \delta^t = 1$$

If for instance, we had a history of $< CC, CC, CC, \ldots >$ where players cooperated each round, then the players would score 2. Thus, for both agents, their normalized payoff for any $0 < \delta < 1$ would be $(1 - \delta) \sum_{t=0}^{\infty} \delta^t 2 = 1(2) = 2$. Thus we have a way of comparing the payoffs in a repeated game to those of the stage game.

Example: Indefinite Repetition and the Prisoner's Dilemma

With the definitions and result seen above, we have one mathematical piece of the puzzle. Our next extension will be applying the above computations to a repeated game. To unravel how Skyrms transforms the Prisoner's Dilemma into a Stag Hunt [Skyrms, 2004], we show the method by which he invokes indefinite repetition as the *shadow of the future.*

	C	D
C	3,3	1,4
D	4,1	2,2

	S	H
S	9,9	0,8
H	8,0	7,7

Table 3.3: Prisoner's Dilemma and the Stag Hunt from Aumann [1990].

Consider the Prisoner's Dilemma with payoff matrix below in Table 3.3:

	C	D
C	3,3	1,4
D	4,1	2,2

	GT	AD
GT	7.5,7.5	4,7
AD	7,4	5,5

Table 3.4: Sample tables of the Prisoner's Dilemma and Stag Hunt as given by $p = .6$ of repeating and strategies of *Grim Trigger* and *Always Defect* with sample payoffs and formalized parameters on costs and benefits.

The setup, taken from Hume and Hobbes, is as follows. There is a probability $p = .6$ that the game will be repeated. If the game ends, both players receive 0. The new table will reflect the EU of each strategy against the other assuming an infinite horizon and no discounting. There are two strategies in the repeated version:

· **Grim Trigger:** Cooperate until Opponent Defects; Then Defect

· **Foole:** Always Defect

With Grim Trigger against itself, game-play will be mutual cooperation ad infinitum. With Grim Trigger against a Foole, the respective histories will be $(CD, DD, DD, \ldots)$. With the Foole against itself, game-play will be mutual defection $(DD, DD, DD, \ldots)$. As we have a probability $p = .6$ that the game will be repeated, we can first begin with the observation that

$$\sum_{j=1}^{\infty} .6^j = .6 \sum_{k=0}^{\infty} .6^k = .6\frac{1}{1-.6} = \frac{.6}{.4} = 1.5$$

We will need this to calculate the expected utilities in the repeated game:

$$\begin{aligned}
EU(GT|GT) &= 3 + 3\sum_{j=1}^{\infty} .6^j = 3 + 1.5(3) = 7.5 \\
EU(GT|AD) &= 1 + 2\sum_{j=1}^{\infty} .6^j = 1 + 1.5(2) = 4 \\
EU(AD|GT) &= 4 + 2\sum_{j=1}^{\infty} .6^j = 4 + 1.5(2) = 7 \\
EU(AD|AD) &= 2 + 2\sum_{j=1}^{\infty} .6^j = 2 + 1.5(2) = 5
\end{aligned}$$

This gives us the payoff matrix transition seen here:

$$\begin{bmatrix} 3 & 1 \\ 4 & 2 \end{bmatrix} \Rightarrow \begin{bmatrix} 3 + 1.5(3) & 1 + 1.5(2) \\ 4 + 1.5(2) & 2 + 1.5(2) \end{bmatrix} = \begin{bmatrix} 7.5 & 4 \\ 7 & 5 \end{bmatrix}$$

This gives us a game strategically equivalent to the Stag Hunt when put into normal form (Table 3.3). This then allows for meaningful communication to arise. Such analysis will form the basis for our later chapters on symmetric trust games. As these games contain strategic restrictions [3] equivalent to the Prisoner's Dilemma, we expect similar results. We will sketch out the strategic automata in place for such repetition and extend this analysis to the 4×4 symmetric trust game in **??**.

Example: Infinite Repetition Prisoner's Dilemma

Taking our example above, consider what would happen to players A and B. In this case, let us assume a discount value of $\delta = .6$ and make a table of their respective utilities in each round.

n	1	2	3	...
δ^n	$.6^1$	$.6^2$	$.6^3$	
U_A^n	$.6^1(0)$	$.6^2(1)$	$.6^3(1)$	
U_B^n	$.6^1(3)$	$.6^2(1)$	$.6^3(1)$	

Table 3.5: Stage utilities for two players in a repeated Prisoner's Dilemma with discount value $\delta = .6$

Based on Table 3.5, we can compute the following:

$$U_A = .6^1(0) + .6^2(1) + .6^3(1) + \ldots = 0 + .6^2 \sum_{j=0}^{\infty} .6^j$$

$$U_B = .6^1(3) + .6^2(1) + .6^3(1) + \ldots = .6(3) + .6^2 \sum_{j=0}^{\infty} .6^j$$

Let us return to the slightly different payoff matrix seen in Table 3.3, only this time, we will assume infinite repetition with a discount factor of $\delta = .6$. Just as before, let us assume there are two strategies playing against each other: *Grim Trigger* (GT) and *Always Defect* (AD). Just as before, we can compute the normalized payoffs over respective repeated plays:

[3] Subsets of their action space created by eliminating actions

$$EU(GT|GT) = (1 - .6)(3 + 3\sum_{j=1}^{\infty} .6^j) = .4(3 + 1.5(3)) = 3$$

$$EU(GT|AD) = (1 - .6)(1 + 2\sum_{j=1}^{\infty} .6^j) = .4(1 + 1.5(2)) = 1.6$$

$$EU(AD|GT) = (1 - .6)(4 + 2\sum_{j=1}^{\infty} .6^j) = .4(4 + 1.5(2)) = 2.8$$

$$EU(AD|AD) = (1 - .6)(2 + 2\sum_{j=1}^{\infty} .6^j) = .4(3 + 1.5(3)) = 2$$

	C	D
C	3,3	1,4
D	4,1	2,2

	GT	AD
GT	3,3	1.6,2.8
AD	2.8,1.6	2,2

Table 3.6: Sample tables of the normalized payoffs Prisoner's Dilemma and Stag Hunt as given by $p = .6$ of repeating and strategies of *Grim Trigger* and *Always Defect* with sample payoffs and formalized parameters on costs and benefits.

We see in the second part of Table 3.6 that a game strategically equivalent to the Stag Hunt appears, as we can see it has two equilibria, where one is payoff-dominant (GT, GT) and the other risk-dominant (AD, AD). We see this as $3 > 2$ and $(2 - 1.6)^2 > (3 - 2.8)^2$. Notice also that multiplying the table by 2.5 gives us the utility table from Table 3.4. As mentioned before, notice that as $\delta = .6$, $1 - \delta = .4 \Rightarrow \frac{1}{1-\delta} = 2.5$. This is exactly the case we made earlier, that the factor $(1 - \delta)$ is what distinguishes the indefinite game's utility from the infinite game, thus giving us strategic equivalence between the two.

3.4 The Folk Theorem

A major part of the literature on repeated games is the **Folk Theorem**, or set of folk theorems [Mailath and Samuelson, 2006]. Their name derives from the fact that they were believed to be true before formal proofs were discovered. The general idea is that there are strategic equilibria that are achievable in the repeated game that can differ from those in the stage game. These can outperform the equilibria in the stage game in terms of efficiency, and thus these theorems motivate a further look into long-run behavior.

Although there are many versions of the *Folk Theorem*, the primary application is that, given payoffs in the range of the game's original utilities and sufficient conditions (repetition, discounting, punishment, etc.), a set of equilibria exist that provide those payoffs. This gives us conditions for establishing long-run relationships. As one of the conditions upon which the folk theorem is based is the *discount factor*, or the degree to which an agent values future interaction, we can use the theorem's results to identify the minimum level of discounting required to sustain a given strategy. Notice that we can consider the discount to be an analog for patience or interest in the future of the relationship. For instance, certain games like the Prisoner's Dilemma might present more opportunities for cheating than others, or certain strategies, like punishment, might hold cooperative behavior in place without highly patient players.

Since we are using a normalized payoff function that makes the repeated game profiles comparable to the stage game, the folk theorem gives us two basic conditions on the outcomes that are possible equilibria:

- *Individually Rational:* Equilibria in the repeated game should be more efficient than the worst-case scenario in the stage game.
- *Socially Feasible:* The payoff profiles in the repeated game will only have utilities that are a combination of the utilities in the stage game. That is, each player's utilities will fall somewhere between their minimum and maximum payoffs in the original game.

The first point above means that players in the repeated game will accept an outcome in equilibrium only if it outperforms the lowest value they could be forced to receive in the stage game, also called the *minmax payoff.* This has to be true for both players; otherwise, there is no reason for a different strategy. The second point means that the repeated game outcomes could never be less than or greater than the payoffs in the original. In the end, the repeated game's payoffs are simply a weighted average of the stage game's payoffs, whose distribution is governed by how often the game lands in each outcome. More technically, the equilibria via the folk theorem are the ones that in the *convex hull* of the payoff matrix and that strictly dominate the minimax profile of the stage game [Mailath and Samuelson, 2006, Ratliff, 1997]. To see this in action, we will see the example of the Prisoner's Dilemma in the next section, as it has many of the characteristics of the repeated trust game in simpler form. An observation to make here is that this set of possible equilibria is infinite in cardinality as we can use mixed strategies.

Feasibility in the Prisoner's Dilemma

There are two primary conditions on the payoffs amenable to the **Folk Theorem.** First, they must be *socially feasible* as an ordered pair that is a weighted average of the original payoffs from the stage game. Geometrically, this means they will be within the convex hull of the stage game's payoffs. Second, each component of the ordered pair of repeated game payoffs must be *individually rational.* This means it will be strictly better than the payoffs in the joint minmax profile from the stage game. To have a better look at this, let us consider the payoffs in the Prisoner's Dilemma as seen below.

We will consider two players X(Row) and Y(Column), with utilities U_X and U_Y. Figures in the following pages will outline our technique for constructing the feasible region of individually rational outcomes graphically:

- Plot stage game outcomes as ordered pairs in the Cartesian plane as (U_X, U_Y). In Figure 3.4 and Figure 3.4 we have payoffs of $(3,3)$, $(4,0)$, $(0,4)$, and $(1,1)$.

- Construct the set of socially feasible outcomes. Geometrically this is the *convex hull* of the outcomes, i.e. the smallest convex region that contains the previously plotted points, seen in Figure 3.4. [4]

- Constrain the region by eliminating outcomes that are not individually rational. Geometrically, this means constructing boundaries lines $\{x > m_X, y > m_Y\}$, where m_X and m_Y are the minmax payoffs. These are the minmax payoffs, the lowest outcomes that each player could expect in equilibrium. (Figure 3.5)

NOTE TO SELF Match to table?

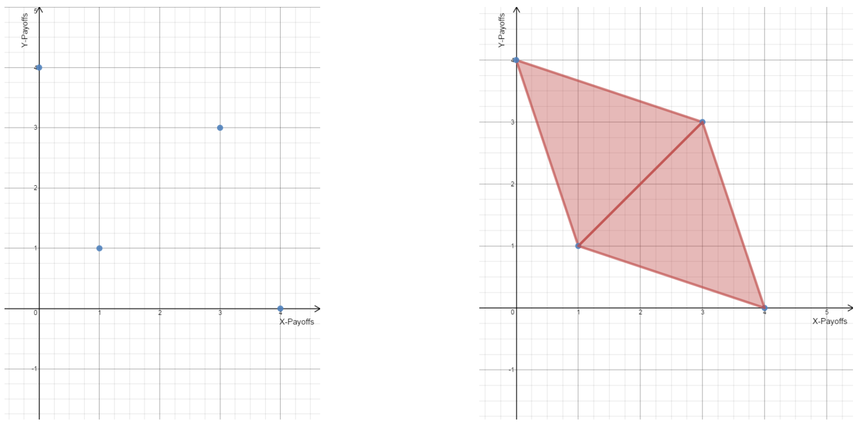

Figure 3.4: Outcomes in the stage game of the Prisoner's Dilemma (left) and the convex hull that determines the region of socially feasible outcomes in the repeated game. The central line connects the symmetric outcomes.

With the feasible region of equilibria under the folk theorem identified, we should make some points on some of the subtleties of the theorem. First, outcomes in this entire region are only possible for very patient players. If players are not sufficiently patient, they might default towards the minmax outcomes. For a small amount of patience, the next outcome selected will be the Pareto-optimal outcome.

Notice also that at the moment we only have an existential statement, not a constructive one. That is, we only have that we can have such outcomes given sufficient patience on the part of the players. We at present do not have a mechanism for arriving at them. One such mechanism is the threat of punishment if one's partner does not act in a way towards the mutual benefit of the group. We will see this when we investigate more the repeated Prisoner's Dilemma and the *Grim Trigger* strategy. This strategy gives us an *equilibrium path,* a way towards sustaining an equilibrium that outperforms the minmax profile.

[4]Intuitively, this is like stretching a rubber band around pegs placed at the various points given by the outcomes.

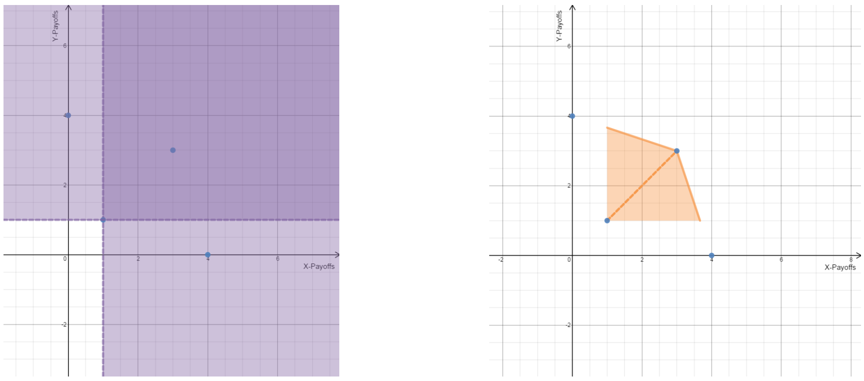

Figure 3.5: Boundaries of individually rational outcomes in the Prisoner's Dilemma (left) and *feasible region* of equilibria admissible under the folk theorem (right). The central line connects the symmetric outcomes. The region does not include points that perform equally well to the minmax outcome. In this case we have $U_X > 1$ and $U_Y > 1$.

Strategies in the Repeated Prisoner's Dilemma

We will now proceed to analyze strategies in the repeated Prisoner's Dilemma under the situation of an unknown discount factor. The two strategies we will investigate will be (*Grim Trigger*), mentioned briefly before in section 3.2, and (*Cooperative Tit-for-Tat*). These operate as follows:

- **Grim Trigger:** Cooperate until opponent Defects; then Always Defect.

- **Cooperative Tit-for-Tat:** Cooperate in first round; then imitate opponent's last move.

While both of these strategies punish defectors, the trigger strategy **GT** punishes them indefinitely, while Tit-for-Tat **TfT** would cooperate against an opponent who switched from defection to cooperation.

	C	D
C	6,6	0,7
D	7,0	1,1

Table 3.7: The Prisoner's Dilemma matrix featured in Figure 3.6.

For the moment, we will use the table in Table 3.7. We first begin with the feasibility region given in Figure 3.6. Notice that any outcome's utility values must fall within the shaded region to satisfy the conditions of the folk theorem. Next we will construct the repeated game strategies through the use of *finite state automata* (FSAs) that give us responses to a given action from one's opponent, as in Figure 3.7

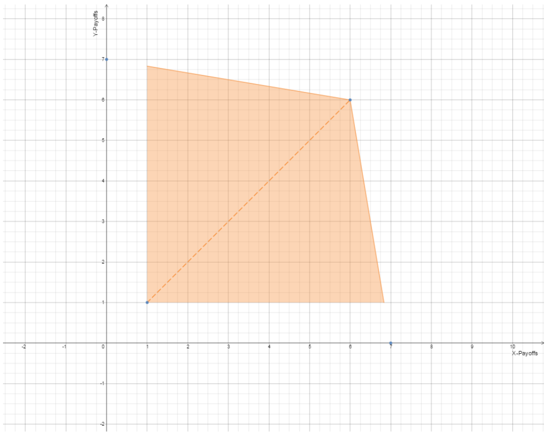

Figure 3.6: Feasible region for Table 3.7. The payoffs for agent X are along the $x-axis$; likewise with Y. We do **NOT** include $x = 1$ and $y = 1$.

Example: Grim Trigger

One method of constructing a strategy in a repeated game can involve an automaton that generates paths between possible actions in the game. This automaton will consider the actions of one player as states, and actions of the other player will lead to transitions between the states. An equilibrium strategy profile in the repeated game will imply that each subsequent path also is an equilbrium strategy profile. This is known as having a *subgame perfect equilibrium*. We can see the transitions between states of a single player playing the Grim Trigger strategy in the diagram below(Figure 3.7) :

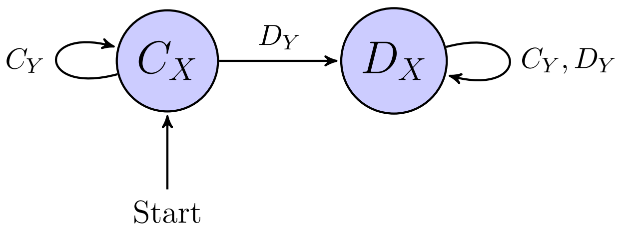

Figure 3.7: FSA depicting Grim Trigger strategies

We can also think of the space of strategies in a larger repeated context as Figure 3.8. We will now use the game's utility table and the guidelines from the strategy to build a discounted payoff matrix. We will do so with the strategies *Grim Trigger* (GT) and always defect (D).

$$M_X = \begin{array}{c c} & \begin{array}{cc} C & D \end{array} \\ \begin{array}{c} C \\ D \end{array} & \begin{bmatrix} 6 & 0 \\ 7 & 1 \end{bmatrix} \end{array}$$

Here we follow the initial strategy and snap to the last entry on the trigger:

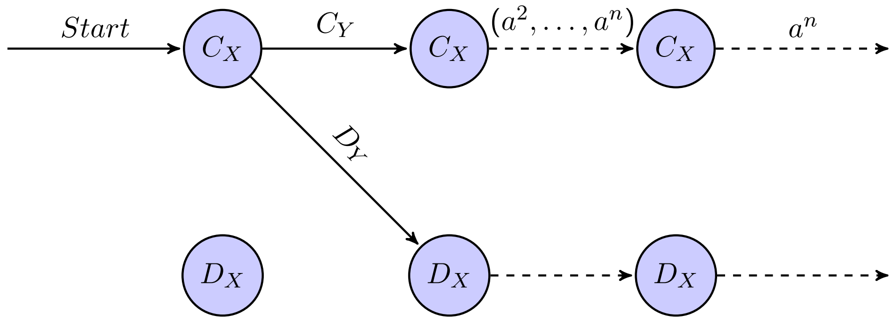

Figure 3.8: Infinite Markov Chain depicting Grim Trigger for a Player X against a Player Y

$$
\begin{aligned}
U(GT|GT) =& (1-\delta)(6+\delta\sum_{t=0}^{\infty}\delta^t(6)) = 6 \\
U(GT|D) =& (1-\delta)(0+\delta\sum_{t=0}^{\infty}\delta^t(1)) = 1\delta \\
U(D|GT) =& (1-\delta)(7+\delta\sum_{t=0}^{\infty}\delta^t(1)) = 6-7\delta \\
U(D|D) =& (1-\delta)(1+\delta\sum_{t=0}^{\infty}\delta^t(1)) = 1
\end{aligned}
$$

This gives us the redefined matrix:

$$
M_X = \begin{array}{c} \\ {\scriptstyle GT} \\ {\scriptstyle D} \end{array}
\begin{array}{c} \begin{array}{cc} {\scriptstyle GT} & {\scriptstyle D} \end{array} \\
\begin{bmatrix} 6 & 1\delta \\ 6-7\delta & 1 \end{bmatrix} \end{array}
$$

Thus we have an advantage in playing the *Grim Trigger* when $6 > 6-7\delta$ or $\delta > 1/6$. When looking back at the dynamic graphics that generated Figure 3.5, we see that this discount measure is what gives a positive area to the feasible region. Note however that GT can never be a truly dominant strategy, as $1 > \delta$ for all values.

Example: Tit-for-Tat

We now turn to *Tit-for-Tat*, a strategy made famous in the Prisoner's Dilemma tournaments sponsored by Axelrod, as documented in Axelrod [1984]. This strategy is in some ways more forgiving than Grim Trigger, for if an opponent cooperated after a string of defection, this strategy would cooperate in return. We can see an automaton based on this strategy, where player X would prompt his strategy in the next round based on the play of player Y, seen in Figure 3.9.

We will now use the game's utility table to build a discounted payoff matrix. Notice this looks identical to our previous analysis if we do so with the strategies

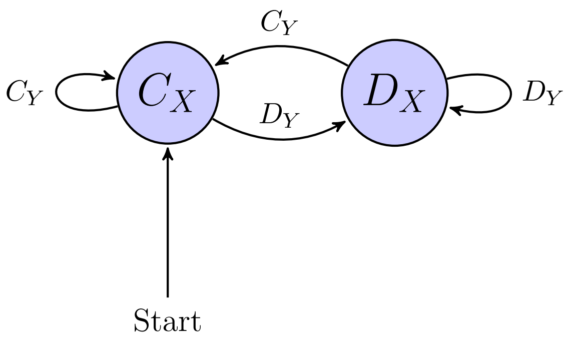

Figure 3.9: FSA depicting tit-for-tat strategies

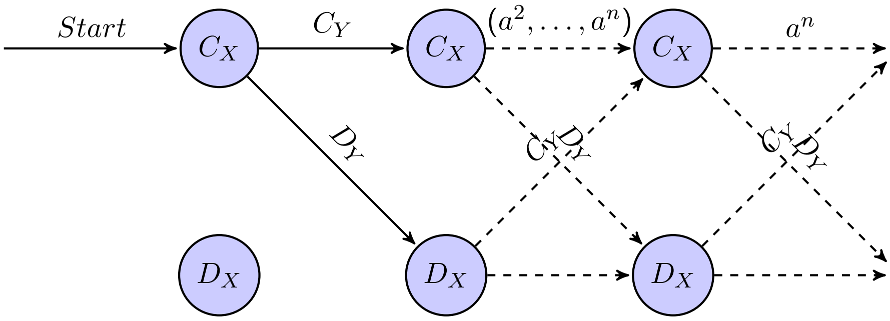

Figure 3.10: Infinite Markov Chain depicting Cooperative Tit-for-Tat for a Player X against a Player Y

Tit-for-Tat (TT) and always defect (D)

$$\begin{aligned}
U(TT|TT) =&(1-\delta)(6+\delta\sum_{t=0}^{\infty}\delta^t(6)) = 6 \\
U(TT|D) =&(1-\delta)(0+\delta\sum_{t=0}^{\infty}\delta^t(1)) = \delta \\
U(D|TT) =&(1-\delta)(7+\delta\sum_{t=0}^{\infty}\delta^t(1)) = 7-6\delta \\
U(D|D) =&(1-\delta)(1+\delta\sum_{t=0}^{\infty}\delta^t(1)) = 1
\end{aligned}$$

This gives us the redefined matrix:

$$M_X = \begin{array}{c} \\ TT \\ D \end{array} \begin{array}{c} \begin{array}{cc} TT & D \end{array} \\ \left[\begin{array}{cc} 6 & \delta \\ 7-6\delta & 1 \end{array}\right] \end{array}$$

Thus we have an advantage in playing the *Tit-for-Tat* when $6 > 7 - 6\delta$ or $\delta > 1/6$. One modification might be to see it play against more sophisticated strategies to see a richer discrepancy between the two. One choice might be to consider *Defecting Tit-for-Tat*, which defects on the first move.

n	0	1	2	3	...
δ^n	δ^0	δ^1	δ^2	δ^3	
DT	$\delta^1(7)$	$\delta^1(0)$	$\delta^2(7)$	$\delta^3(0)$	
CT	$\delta^1(0)$	$\delta^1(7)$	$\delta^2(0)$	$\delta^3(7)$	

Table 3.8: Stage utilities for two players in a repeated Prisoner's Dilemma with discount value δ among players using Cooperative Tit-for-Tat(CT) and Defective Tit-for-Tat (DT).

Consider how the rounds of the repeated Prisoner's Dilemma might play between the two strategies (Table 3.8). The symmetric strategies will play out the same as before, but the asymmetric ones CT vs. DT will follow the table Table 3.8, as every other round will repeat. Normalizing by $1-\delta$, we have

$$U(CT|DT) = (1-\delta)(0+\delta\sum_{t=0}^{\infty}\delta^{2t}(7)) = \frac{7\delta(1-\delta)}{1-\delta^2} = \frac{7\delta}{1+\delta};$$
$$U(DT|CT) = (1-\delta)(7+\delta^2\sum_{t=0}^{\infty}\delta^{2t}(7)) = (1-\delta)7+\frac{7\delta^2(1-\delta}{1-\delta^2} =$$
$$(1-\delta)7+\frac{7\delta^2}{1+\delta} = \frac{(1-\delta^2)7+7\delta^2}{1+\delta} = \frac{7}{1+\delta}$$

This gives us the discounted payoff matrix

$$M_X = \begin{array}{cc} & \begin{array}{cc} CT & DT \end{array} \\ \begin{array}{c} CT \\ DT \end{array} & \begin{bmatrix} 6 & \frac{7\delta}{1+\delta} \\ \frac{7}{1+\delta} & 1 \end{bmatrix} \end{array}$$

In the limit, as $\delta \to 1$, these two strategies will approach each other in their success rates, although DT will always outperform CT by a slim margin. However, in our analysis, we want the discount that leads the cooperative strategy to dominate, this occurs when $6 > \frac{7}{1+\delta}$, or $\delta > \frac{1}{6}$. Interestingly, we also have that when $\delta > \frac{1}{6}$, $\frac{7\delta}{1+\delta} > 1$, giving us that CT becomes a strictly dominant strategy, an encouraging result.

One-Shot Deviations and Generating Equilibria

It is not always the case that we can truly access every point in the feasible region of games like the repeated Prisoner's Dilemma via an equilibrium path governed by a strategy like Grim Trigger. The salient factor is that the discount rate δ must be high enough to sustain the larger space of equilibria. Otherwise, agents will default to the minmax outcomes or the Pareto-optimal ones.

To see this in more detail, we will examine a version of the Prisoner's Dilemma with the feasible region in and given by the payoff table in Table 3.9.

	C	D
C	b-c,b-c	1-c, b
D	b,1-c	1,1

	C	D
C	6,6	0,7
D	7,0	1,1

Table 3.9: The Prisoner's Dilemma matrix featured in Figure 3.11.

On a deeper level, what guarantees the existence of these infinite spaces of equilibria is the *one-shot deviation principle.* This principle says that strategy profiles that do not allow a player to deviate from them in one stage of a game will be equilibria in the infinitely repeated game. These equilibria are known as *subgame-perfect equilibria*, for these strategy profiles should also be equilibria for any truncated version of the game that omits part of the game's history.

What guarantees that certain actions will not be profitable one-shot deviations from a given strategy profile is the existence of contingency plans on the part of the agents. These *continuation promises* give us a way to enforce cooperation through the threat of punishment. Each agent with such a plan can maintain various levels of cooperation through the common knowledge of the existence of these plans. That is, common knowledge of the threat of punishment keeps the appropriate strategy profile in place. Note that this profile need not be the highest-performing, as the figures in Figure 3.11 depict for various levels of discounting. These regions give neighborhoods surrounding the payoffs of the four action profiles in the Prisoner's Dilemma.

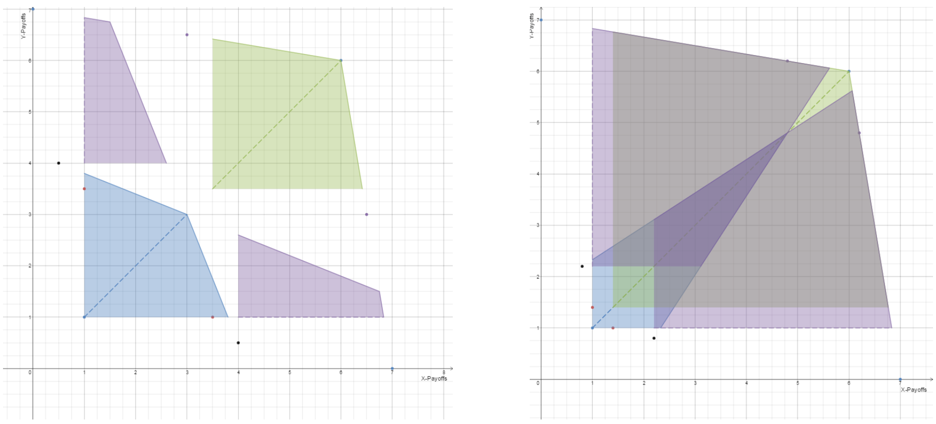

Figure 3.11: For varying levels of the discount δ, games like the repeated Prisoner's Dilemma admits different subsets of the feasible region. The two figures depict cases of $\delta = .5$ and $\delta = .8$. As we increase δ, we approach covering the entire feasible region of equilibria admissible under the folk theorem.

Consider now that in 2×2 symmetric games like the Prisoner's Dilemma we have two actions, $C$and D, and from these actions we derive strategies, denoted for instance by σ. These strategies lead to various payoffs, as seen in Table 3.10 for the generalized case. Now consider the case of a desirable (efficient) action profile CC and a strategy like Grim Trigger promoting adherence to σ but punishing indefinitely in the case of defection. A player deviating from σ to choose action

	C	D
C	P,P	Q,R
D	R,Q	S,S

Table 3.10: Schema for a symmetric game

D according to a strategy σ' should have the following payoffs in the repeated case:

$$U(\sigma'|\sigma) =(1-\delta)(R+\delta\sum_{t=0}^{\infty}\delta^t(S))$$

Note however that

$$\delta\sum_{t=0}^{\infty}\delta^t(S) = \frac{\delta}{1-\delta}$$
$$\Rightarrow (1-\delta)\delta\sum_{t=0}^{\infty}\delta^t(S) = \delta S$$

How does this play into the concept of *one-shot deviation* and *enforceable* profiles? According to Mailath and Samuelson [2006], this is a case of *perfect monitoring*, where each player knows the actions of the other in the previous round. In games like this, we have a way to describe certain action profiles as leading to *subgame-perfect equilibrium* payoffs. This is the case when there is an action profile for all of the players a^* and a function g where

$$(1-\delta)U_i(a^*) + \delta g_i(a^*) \geq (1-\delta)U_i(a_i, a_j^*) + \delta g_i(a_i, a_j^*) \tag{3.3}$$

I.e. these payoffs follow the one-shot deviation principle, as they outperform changes in strategy on the part of a single agent i. To explicate the notation in (3.3), what the function g_i does is regulate the behavior of the player i. If this function exists, we can rule out possible deviations from the given strategy profile. This function is tied to the discount values, and it will differ on the various pure-strategy profiles. What this gives us is that we can find ways to enforce a given action profile a^*, should our partner deviate. This should hold for both players i, j. In the case of the Prisoner's Dilemma we want that

$$(1-\delta)U_1(CC) + \delta g_1(CC) \geq (1-\delta)(U_1(DC) + \delta g_1(DC)$$

and

$$(1-\delta)U_2(CC) + \delta g_2(CC) \geq (1-\delta)(U_2(CD) + \delta g_2(CD)$$

In the case of the table seen in Table 3.11, we chose the values to provide a more detailed picture of the varying feasible regions for different discount values. First, let us consider the discount value δ that promotes the payoff profile $(b-c, b-c)$ of pure cooperation. As the game is symmetric, we will only consider the payoffs of one player. For these particular payoffs, we want that

$$(1-\delta)(b-c) + \delta g_1(CC) \geq (1-\delta)(b + \delta g_1(DC)$$

In this case, we want the Pareto-optimal case of promoting CC throughout, so we want that $\delta g_1(CC) = b - c$ and $\delta g_1(DC) = 1$. This gives us

$$\delta \geq \frac{c}{b-1} = \frac{1}{6}$$

However, we can also achieve lesser payoffs in the case of our prescribed values. Imagine a case where cooperation is not as profitable, and a group playing CC would receive a payoff profile of (k, k), where $k \leq b - c$. In the game seen in Figure 3.11, this might be a profile of $(4, 4)$. A viable strategy might be to keep the punishment for defection the same at $\delta g_1(DC) = 1$, but a different discount might emerge. This would give us

$$\begin{aligned} (1-\delta)(b-c) + \delta k &\geq (1-\delta)(b) + \delta g_1(DC) \\ (1-\delta)(b-c) + \delta k &\geq (1-\delta)(b) + \delta(1) \\ &\Rightarrow \delta \geq \frac{c}{c+k-1} \end{aligned}$$

We can note two things. One, in the event that $k = b - c$, we have an equivalent result to the previous case. In the event that $k < b - c$, we have that the discount value in this case is always greater. What does this mean? This gives us that sustaining sub-optimal payoffs that are distinct from the minmax profile will require more patience on the part of the players. The intuitive reason is that since we can sustain the Pareto-optimal payoff of $(b-c, b-c)$ with Grim Trigger, anything less than $b - c$ would require more patient players.

	C	D
C	b-c,b-c	1-c, b
D	b,1-c	1,1

	C	D
C	6,6	0,7
D	7,0	1,1

Table 3.11: The Prisoner's Dilemma matrix featured in Figure 3.11.

We can see this as assuming that the discount leading to this outcome will lead to a contradiction.

$$\frac{c}{c+k-1} < \frac{c}{b-1} \Rightarrow b - c < k \Rightarrow \perp$$

Is this outcome reachable? This would mean that $\delta < 1$. Certainly, as this would give us

$$\frac{c}{c+k-1} < 1 \Rightarrow k > 1$$

This will be true as long as $k > 1$, the minmax payoff. [5] If not, we are violating the conditions of the Folk Theorem and our original intent of producing outcomes that perform better than the minmax profile. Thus we now have that for any symmetric outcome k between $U(CC)$ and $U(DD)$, we can find an appropriate discount that will guarantee its existence. As for outcomes near the payoffs from the asymmetric profiles, we refer the reader to Mailath and Samuelson [2006].

General Results on the Repeated Prisoner's Dilemma

In the previous few sections, we have experimented with variations on the Prisoner's Dilemma to see applications of the folk theorem in repeated games. We adopt here the general form of the Prisoner's Dilemma that will be a cornerstone of our analysis in our discussion of generalized trust games **??**, as the Prisoner's Dilemma is equivalent to a subgame of symmetric trust games. If we consider the Grim Trigger (GT) strategy again vs. Always Defect(D), we will follow the initial strategy and snap to the last entry on the trigger:

$$M_X = \begin{array}{c c} & \begin{array}{cc} C & D \end{array} \\ \begin{array}{c} C \\ D \end{array} & \begin{bmatrix} b-c & -c \\ b & 0 \end{bmatrix} \end{array}$$

$$\begin{aligned} U(GT|GT) =& (1-\delta)(b-c+\delta\sum_{t=0}^{\infty}\delta^t(b-c)) = b-c \\ U(GT|D) =& (1-\delta)(-c+\delta\sum_{t=0}^{\infty}\delta^t(0)) = -c(1-\delta) \\ U(D|GT) =& (1-\delta)(b+\delta\sum_{t=0}^{\infty}\delta^t(0)) = b(1-\delta) \\ U(D|D) =& (1-\delta)(0+\delta\sum_{t=0}^{\infty}\delta^t(0)) = 0 \end{aligned}$$

This gives us the redefined matrix:

$$M_X = \begin{array}{c c} & \begin{array}{cc} GT & D \end{array} \\ \begin{array}{c} GT \\ D \end{array} & \begin{bmatrix} b-c & -c(1-\delta) \\ b(1-\delta) & 0 \end{bmatrix} \end{array}$$

[5] While the feasible discount will always be less than one, the payoff needs to be greater than the minmax payoff, which in this example is also one.

Thus we have an advantage in playing the *Grim Trigger* when $b-c > b(1-\delta)$ or $\delta > \frac{c}{b}$. I.e. we need the discount parameter to be greater than the ratio of the cost of cooperation compared to the benefit of the other cooperating. Can we achieve a truly dominant strategy of Grim Trigger? In this case, no, as we will always have $-c(1-\delta) < 0$. So in this case, the repeated Prisoner's Dilemma should resemble the incentives found in the Stag Hunt as long as the outcome DD is risk-dominant. This will happen when

$$(b(1-\delta) - (b-c))^2 < (-c(1-\delta))^2$$

As the square root function is monotone on intervals greater than zero, we can compare the absolute values of these quantities, giving us

$$|b(1-\delta) - (b-c)| < |-c(1-\delta)|$$

Note also that if we take discount values where $b-c > b(1-\delta)$, as done above, we will have

$$\begin{aligned} b - c &< c(1-\delta) \\ \Rightarrow \delta &> \frac{2c}{b+c} \end{aligned}$$

In the end, provided we have that $\delta > max(\frac{c}{b}, \frac{2c}{b+c})$, the Stag Hunt will emerge. As we have that $c < b$, we now have that $\frac{c}{b} < \frac{2c}{b+c})$, and thus we only need that $\delta > \frac{2c}{b+c}$.

Example: Variable MinMax in Hawks & Doves

It could be the case that the minmax payoffs are not immediately clear. For example, in games like Hawks & Doves we find that the least efficient outcome is not the equilibrium and that the two equilbria arise from the asymmetric action profiles. One remedy to this ambiguity is to examine the minimum payoffs in a mixed strategy equilibrium. We next examine the case of a repeated game of Hawks & Doves. This will help use analyze in more detail the concept of the *MinMax* strategy.

	H	D
H	0;0	7;2
D	2;7	6;6

Table 3.12: Hawks & Doves with players A and B. Let$Pr(H_A) = p$ and $Pr(H_B) = q$.

We begin with the table seen in 3.12. We would like to consider the set of mixed strategies that could arise, where A plays a strategy M_1 and B plays M_2. From here, we take the probabilities that A plays H to be p and the probabilities that B plays H to be q. We now begin the calculation seen below. Observe that this game is symmetric, so the same calculation can be made for B.

$$
\begin{aligned}
EU_A &= p(0q + 7(1-q)) + (1-p)(2q + 6(1-q)) && (3.4)\\
&= 7p(1-q) + 2(1-p)q + 6(1-p)(1-q) && (3.5)\\
&= 7p - 7pq + 2q - 2pq + 6 - 6p - 6q + 6pq \quad = p - 4q - 3pq + 6 && (3.6)
\end{aligned}
$$

Given this utility function, let us now consider how it varies along changes in q. This will tell us how the agent B can minimize the score of A. As EU_A is a function of two varying parameters, we can compute:

$$\frac{\partial U_A(M_1, M_2)}{\partial q} = -4 - 3p$$

Note that since $p \geq 0$, this expression is negative for all q, and thus EU_A is decreasing as q gets larger. Thus, by setting $q = 1$, we get the strategy profile at which A's score would achieve its greatest loss. Thus for A to minmize this expected loss, A should set $p = 0$, achieving a utility of 2 in the strategy profile (D, H). The interesting thing is the symmetry involved here, as the reverse would be true for B with (H, D). What we see in the end is that the region of feasiblity amenable to the Folk Theorem in this game must therefore include payoffs $U_i > 2$ for all players i. This gives a tight constraint on the individually rational outcomes that outperform the mutual minmax, as seen in the feasible region of Figure 3.12.

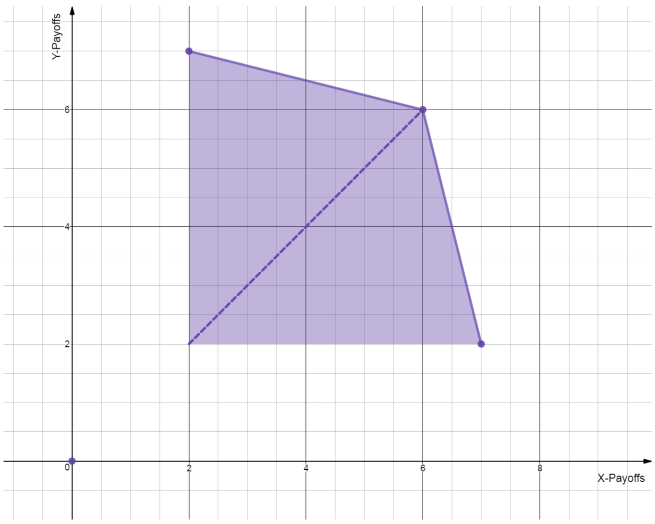

Figure 3.12: Feasible region for Hawks & Doves. We do **NOT** include the minmax boundaries $x = 2$ and $y = 2$. NE at $(7, 2)$ and $(2, 7)$.

Looking back at our game, it should not surprise us that 2 is the minmax payoff for both, as it is the lowest payoff present in equilibrium. One subtlety of the feasible region in Figure 3.12 is that although it resembles the Prisoner's

Dilemma in some respects, we see that in this case the minmax outcome is not one of the original payoff points and that the two asymmetric outcomes lie on the edges of the region, as opposed to being outside of it.

Monitoring and Punishment

Monitoring is a concept used to describe the ability of players to observe each others' actions. A strong (necessary) mechanism for enforcing optimal behavior in repeated games, monitoring ties in closely to an agent's preference for having a "good" reputation. There are several categories of monitoring, including *perfect* and *imperfect* monitoring. The distinction between the two is how reliably an agent can tell what the other did in the previous round. Note that this will factor in as we derive incentives for deception and flattery. In the previous cases, we have only considered perfect monitoring, as each agent knew the previous move of his opponent. We also want to distinguish between cases where agents are observed by those with whom they are interacting and cases where agents are observed by an outside party. We claim these cases also contribute to the interest in face and reputation.

When considering reputation and monitoring, the more easily one's actions can be monitored from move to move, the less likely one is to exploit one's opponent. If an agent is caught attempting to exploit his opponent, one option is for his opponent to punish him.

Definition 5. *We say that an agent i **punishes** j in a repeated game if i takes action that minimizes the score of j conditional on some previous play.*

Punishment amounts to minimizing the score of your opponent. In the case of the repeated Prisoner's Dilemma, defecting against someone who has defected on you minimizes their score. Examples like the *Grim Trigger* strategy punish more harshly than *Tit-for-Tat.* Players may choose to *forgive* another player, i.e. play the non-punishing strategy with non-zero probability, in which case we call them *generous.* One modification to punishing strategies is making them asymmetric in payoffs, and we hope this might provide a suitable analog to power. E.g. it is less costly for a powerful person to punish a defector.

3.5 Repetition and Reputation

When considering the previous examples from the sections on the Prisoner's Dilemma, we can imagine that there must be ways to improve the overall outcome for each player. One we we have seen is repetition. When agents have not directly observed each other, another is to introduce reputation. [6] Both of these notions in some ways go hand-in-hand, as our strategies for a repeated game might change if our partner could develop a belief about our future behavior.

Within a one-shot game or considering one-shot strategies, a reputation could be the probability of an action within that game. For example, consider the

[6]The formal idea of a mechanism may be left to the appendix

Prisoner's Dilemma and the case where X believes $Pr(C_Y) = \frac{3}{4}$. If we consider the two types in the game to be *cooperators* and *defectors*, we could say that the player Y has a *reputation* for playing C 75% of the time. This in some ways is an unintelligent reputation, as it says nothing about adapting to new strategies.

It does not always have to be the case that reputations are so simple. A player may have a reputation conditioned on another strategy. E.g. in the repeated Prisoner's Dilemma we might have that a player Y has the following reputation under the beliefs of X:

$$Pr(C_Y|C_X) = 1$$
$$Pr(C_Y|D_X) = .2$$

In this previous case we see that the player Y has a reputation for playing something similar to *Generous Tit-for-Tat*. This is the simple case that now we have a reputation conditioned on the action of a previous player. We could also have a reputation conditioned on the history between two players.

$$Pr(C_Y|D_X \notin H_X) = 1$$
$$Pr(C_Y|D_X \in H_X) = 0$$

This belief depicted above shows a player Y to have a reputation for playing a *Grim Trigger* strategy that is not just conditioned on the previous actions but an entire history. This history is based on whether the player X played D in any of the previous rounds . One avenue therefore for exploring whether an agent is maintaining face is how they play according to a strategy like Tit-for-Tat that shows *strong reciprocity*. Notice that we have advanced these beliefs on reputation from actions an agent might choose in a one-shot game to strategies an agent might adopt in the repeated game. We could even adapt this to more general classes of games, where we think of reputation as a belief over an agent's *type* or set of preferences. For instance, if one agent has a reputation for having sympathetic preferences (seen in Chapter 6), then his opponent might attempt to exploit that preference.

Typology of Reputation

There are several notions of reputation of interest to our discussion. We will build our catalog and notation from Mui et al. [2002] and Mailath and Samuelson [2006]. Before we begin the exposition, we should make some caveats on the nature and variety of notions for reputation. In contrast to the economic literature, we keep the discussion person-centered and steer away from firm-based notions.

The first point to make is that reputation depends on context. This context may be socially based or interaction based. The second point to mention is that reputation can be individual or group-centric. Individual reputation can be direct or indirect. Further, direct reputation can be observed through interaction with others or encountered by an agent himself. Indirect reputation could be a prior probability, propagated by word of mouth (network dissemination) or a result of being in a group.

To highlight the various forms of reputation, we will adapt the typology of reputation from Mui et al. [2001].

1. Group Reputation
2. Individual Reputations
   (a) Direct
      i. Encounter-Based
      ii. Observation-Based
   (b) Indirect
      i. Prior
      ii. Group
      iii. Propagated

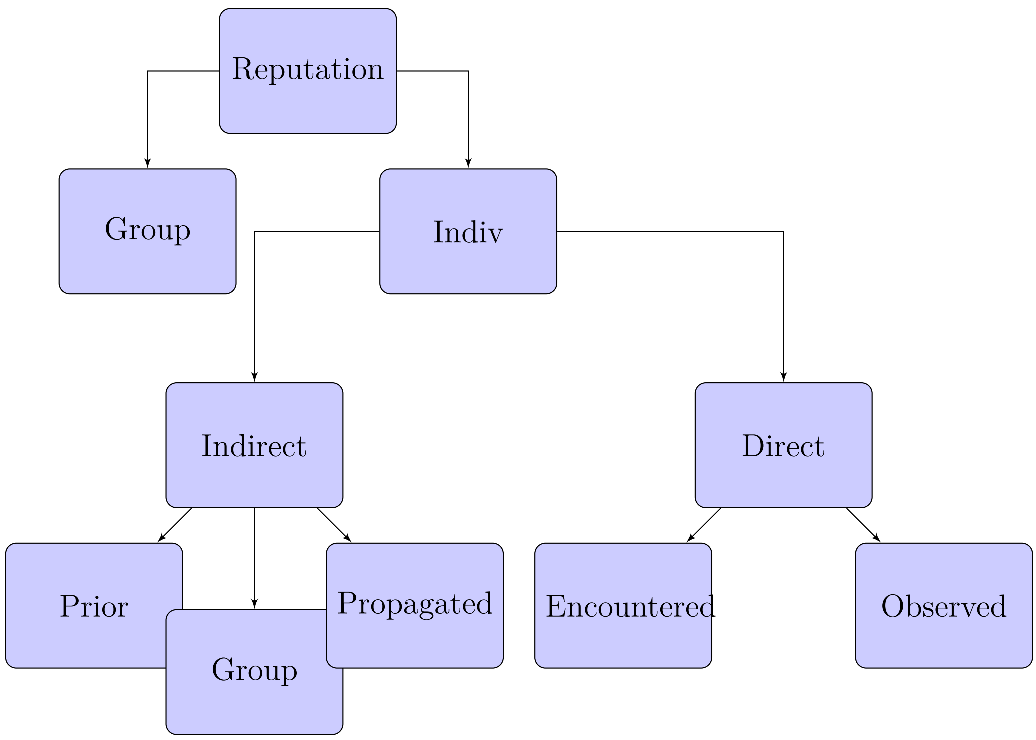

Figure 3.13: A typology of reputation

This typology of reputation will feature in **??** as we examine what notions of reputation are important for the maintenance of linguistically oriented cooperation.

3.6 Conclusion

In this chapter we have introduced several concepts and laid the groundwork for our later development of trust games. To highlight a few points

- We would like to claim that polite language can conventionalize as a rational and payoff-dominant norm founded on repeated interaction.
- Repeated instances of the Prisoner's Dilemma show how cooperation can emerge through conditions on repeated interaction like reciprocity, punishment, and discounting.
- We saw the strategic relationship between indefinite repetition (something controlled by external circumstances) and discounting of repeated interaction with infinite horizon (a preference for present utility over future utility).
- We derived and constructed strategies in repeated games by which cooperation in a single encounter gives way to coordination norms, for instance when we derived the discount conditions necessary to transform a repeated Prisoner's Dilemma into a Stag Hunt.

Bibliography

Robert J. Aumann. *Nash Equilibria Are Not Self-Enforcing*, chapter 34, pages 667–677. Elsevier, 1990.

Robert Axelrod. *The Evolution of Cooperation.* Perseus Books Group, Cambridge, 1984.

Colin Camerer. *Behavioral Game Theory: Experiments in Strategic Interaction.* Princeton University Press, Princeton, 2003.

George J. Mailath and Larry Samuelson. *Repeated games and reputations: long-run relationships.* Oxford University Press, Oxford, 2006.

L. Mui, M. Mohtashemi, C. Ang, P. Szolovits, and A. Halberstadt. Ratings in distributed systems: A bayesian approach. In *Proceedings of the Workshop on Information Technologies and Systems (WITS)*, pages 1–7, 2001.

Lik Mui, Mojdeh Mohtashemi, and Ari Halberstadt. Notions of reputation in multi-agents systems: a review. In *Proceedings of the first international joint conference on Autonomous agents and multiagent systems: part 1*, pages 280–287. ACM, 2002.

Jim Ratliff. Graduate-level course in game theory, 1997.

Ariel Rubinstein. "economics and psychology"? the case of hyperbolic discounting*. *International Economic Review*, 44(4):1207–1216, 2003.

Brian Skyrms. *The Stag Hunt and the Evolution of Social Structure.* Cambridge University Press, Cambridge, 2004.