Notes on Game Theory

Debasis Mishra

October 31, 2017

1 Games in Strategic Form

A game in **strategic form** or **normal form** is a triple $\Gamma \equiv (N, \{S_i\}_{i \in N}, \{u_i\}_{i \in N})$ in which

- $N = \{1, 2, \ldots, n\}$ is a finite set of players,
- S_i is the set of strategies of player i, for every player $i \in N$ - the set of strategy profiles is denoted as $S \equiv S_1 \times \ldots \times S_n$,
- $u_i : S \to \mathbb{R}$ is a utility function that associates with each profile of strategies $s \equiv (s_1, \ldots, s_n)$, a payoff $u_i(s)$ for every player $i \in N$.

Here, the set of strategies can be finite or infinite. The assumption is that players choose these strategies simultaneously in the game, i.e., no player observes the strategies played by other players before playing his own strategy. Here, simultaneous only means they choose their strategies independently without observing each others strategies - one can think of a situation where each player writes down the possible course of action for every possible contingencies in the future and submit it to the game. A strategy profile of all the players will be denoted as $s \equiv (s_1, \ldots, s_n) \in S$. A strategy profile of all the players excluding a Player i will be denoted by s_{-i}. The set of all strategy profiles of players other than a Player i will be denoted by S_{-i}.

We give two examples to illustrate games in strategic form.

1. The first game is the game of Prisoner's Dilemma. Suppose $N = \{1, 2\}$. These players are prisoners. Because of lack of evidence, they have been questioned in separate rooms and made to confess their crimes. If they both confess, then they each achieve a payoff of 1. If both of them do not confess, then they can achieve higher payoffs of 2 each. However, if one of them confesses, but the other one does not confess, then the confessed player gets a payoff of 3 but the player who does not confess gets a payoff of 0.

   What are the strategies in this game? For both the players, the set of strategies is {Confess (C), Do not confess (D)}. The payoffs from the four strategy profiles can be written in a matrix form. It is shown in Table 1.

2. Two shops are competing to locate themselves on a street - represented by the compact interval $[0, 1]$. Suppose consumers are uniformly located on the street. Once shops are located, the consumers go the nearest shop - with ties broken using a equal probability. The utility of a shop is the *measure* of consumers he gets. Here the set of strategies

	c	d
C	$(1,1)$	$(3,0)$
D	$(0,3)$	$(2,2)$

Table 1: The Prisoner's Dilemma

are the points in $[0,1]$ - an infinite set. If location of shop 1 is x_1 and shop 2 is x_2, then the payoff of shop 1 is

$$u_1(x_1, x_2) = x_1 + \frac{x_2 - x_1}{2} \qquad \text{if } x_1 \leq x_2$$
$$u_1(x_1, x_2) = (1 - x_1) + \frac{x_1 - x_2}{2} \qquad \text{if } x_1 > x_2$$
$$u_2(x_1, x_2) = 1 - u_1(x_1, x_2).$$

The strategy of a game is a powerful tool for representation. It can potentially represent many situations. It provides a complete description of actions that need to be taken in all possible contingencies. As an example, suppose two individuals work every day together on some project for 2 days. Based on the effort put by the individuals on these days, they realize payoffs at the end of two days. Here, a strategy is an effort level in Day 1 and an effort level in Day 2. Players choose such strategies (a combination of effort levels for two days) and that results in payoffs. Later, we will show that many strategic interactions can be reduced to such strategic form by specifying the strategies appropriately. As we go along in the course, we will see that strategies have different meaning and definitions in different types of interactions agents can have. But, in all such cases, the common thread that will run is: *a strategy will describe what an agent must do in all possible contingencies.* One way to interpret this is that the agent has written down his strategy in an envelope (or, written down a computer program) that needs to describe his actions in all possible situations that can arise.

2 Beliefs of Players

The objective of game theory is to provide predictions of games. To arrive at reasonable predictions for normal form games, let us think how agents will behave in these games. One plausible idea is each agent forms a belief about how other agents will play the game and play his own strategy accordingly. For instance, in the Prisoner's Dilemma game in Table 1, Player 1 may believe that Player 2 will play c with probability $\frac{3}{4}$ and play d with probability $\frac{1}{4}$. In that case, he can compute his payoffs (using expected utility) from both the strategies:

- from playing C: $\frac{3}{4}1 + \frac{1}{4}3 = \frac{6}{4}$,
- from playing D: $\frac{3}{4}0 + \frac{1}{4}2 = \frac{2}{4}$.

Clearly, playing C is better under this belief. Hence, Player 1 will play D given his belief.

Note. From now on, unless stated otherwise, we will assume S_i for all i to be finite sets. Many results, with the help of extra notations and mathematics, extend to the case where strategy sets are not finite.

Formally, each player i forms a belief $\mu_i \in \Delta S_{-i}$, where ΔS_{-i} is the set of all probability distributions over S_{-i}. Given these beliefs, it computes his utility given his beliefs as:

$$\mathcal{U}_i(s_i, \mu_i) := \sum_{s_{-i} \in S_{-i}} u_i(s_i, s_{-i}) \mu_i(s_{-i}) \ \forall \ s_i \in S_i.$$

Then it chooses a strategy s_i^* such that $\mathcal{U}_i(s_i^*, \mu_i) \geq \mathcal{U}_i(s_i, \mu_i)$ for all $s_i \in S_i$.

There are two reasons why this may not work. First, beliefs may not be formed, i.e., where do beliefs come from? Second, beliefs may be incorrect. Even if agent i believes certain strategies will be played by others, other agents may not play them. In game theory, there are two kinds of solution concepts to tackle these issues: (a) solution concepts that work independent of beliefs and (b) solution concepts that assume correct beliefs. The former is sometimes referred to as a *non-equilibrium* solution concept, while the latter is referred to as an *equilibrium* solution concept.

3 Domination

The idea of domination is probably the strongest possible prediction of a game. Dominance is a concept that uses strategies whose performance is good irrespective of the beliefs.

Definition **1** *A strategy $s_i \in S_i$ for Player i is* **strictly dominant** *if for every $s_{-i} \in S_{-i}$, we have*

$$u_i(s_i, s_{-i}) > u_i(s_i', s_{-i}) \ \forall \ s_i' \in S_i \setminus \{s_i\}.$$

Similarly, a strategy $s_i \in S_i$ for Player i is **weakly dominant** *if for every $s_{-i} \in S_{-i}$, we have*

$$u_i(s_i, s_{-i}) \geq u_i(s_i', s_{-i}) \ \forall \ s_i' \in S_i \setminus \{s_i\}.$$

It is fairly clear that the idea of domination requires a strategy to be optimal for a player irrespective of what he believes other players are doing. The following lemma formalizes it.

Lemma 1 *A strategy s_i for Player i is strictly dominant if and only if for* all *beliefs μ_i*

$$\mathcal{U}_i(s_i, \mu_i) > \mathcal{U}_i(s'_i, \mu_i) \ \forall \ s'_i \in S_i \setminus \{s_i\}.$$

A strategy s_i for Player i is weakly dominant if and only if for all *beliefs μ_i*

$$\mathcal{U}_i(s_i, \mu_i) \geq \mathcal{U}_i(s'_i, \mu_i) \ \forall \ s'_i \in S_i \setminus \{s_i\}.$$

Proof: We do the proof for strictly dominant strategies - the weak dominance part follows similarly. Suppose s_i is a strictly dominant strategy for Player i. Fix a belief μ_i. Now, note the following:

$$\begin{aligned}
\mathcal{U}_i(s_i, \mu_i) &= \sum_{s_{-i}} u_i(s_i, s_{-i})\mu_i(s_{-i}) \\
&> \sum_{s_{-i}} u_i(s'_i, s_{-i})\mu_i(s_{-i}) \qquad \text{(By definition of strict dominance)} \\
&= \mathcal{U}_i(s'_i, \mu_i).
\end{aligned}$$

For the other direction, suppose s_i is an optimal strategy for Player i for all beliefs μ_i. Now, choose some s_{-i} and consider the belief that $\mu_i(s_{-i}) = 1$. Then, it follows that

$$u_i(s_i, s_{-i}) = \mathcal{U}_i(s_i, \mu_i) > \mathcal{U}_i(s'_i, \mu_i) = u_i(s'_i, s_{-i}).$$

■

In the Prisoner's Dilemma game in Table 1, the strategy C (or c) is a strictly dominant strategy for each player.

If we assume a modest amount of *rationality* in players, we must believe that players must play strictly dominant strategies (whenever they exist). Here, rationality requires that every player plays a strategy that maximizes his utility given his belief about other players' strategies. However, many games do not have a strictly dominant strategy for both the players. For instance, in the game in Table 2, there is no strictly dominant strategy for either of the players.

However, irrespective of the strategy played by Player 2, Player 1 always gets less payoff in B than in M. In such a case, we will say that Strategy B is strictly dominated.

Definition 2 *A strategy $s_i \in S_i$ for Player i is* **strictly dominanted** *if there exists $s'_i \in S_i$ such that for every $s_{-i} \in S_{-i}$, we have*

$$u_i(s_i, s_{-i}) < u_i(s'_i, s_{-i}).$$

In this case, we say that s'_i strictly dominates s_i.

	L	C	R
T	$(2,2)$	$(6,1)$	$(1,1)$
M	$(1,3)$	$(5,5)$	$(9,2)$
B	$(0,0)$	$(4,2)$	$(8,8)$

Table 2: Domination

A belief based definition is also possible: irrespective of beliefs of Player i, playing s_i is worse than playing s'_i.

Another assumption of rationality is that a rational player will never play a strictly dominated strategy. But does that imply we can forget about a strictly dominated strategy? The main issue is removing a strategy of Player i influences the support of the belief of other players. So, unless we assume something about the knowledge level of other players, it is not clear whether we can remove a strategy from Player i. Note that belief of a player about others' strategies influences his choice of optimal strategy.

To see this, consider the example in Table 2. Strategy B is strictly dominated by Strategy M for Player 1. Hence, if Player 1 is rational, then he will not play B. If Player 2 does not know that Player 1 is rational, then he cannot eliminate B from the support of his belief of Player 1's strategies. Suppose **Player 2 knows that Player 1 is rational**. Then, he can conclude that Player 1 will not play B ever. As a result, his belief on what Player 1 can play must put probability zero on B. In that case, his Strategy R is strictly dominated by Strategy L. So, he will not play R. Now, if **Player 1 knows that Player 2 is rational and Player 1 knows that Player 2 knows that Player 1 is rational**, then he will not play M because it is now strictly dominated by T. Continuing in this manner, we will get that Player 2 does not play C. Hence, the only strategy profile surviving such elimination is (T, L).

The process we just described is called *iterated elimination of strictly dominated strategies*. It requires more than rationality. We do not provide a formal treatment of this topic. Loosely, a *fact* is **common knowledge** among players in a game if for any finite chain of player $(i_1, \ldots, i_k)$ the following holds: Player i_1 knows that Player i_2 knows that Player i_3 knows that ... Player i_k knows the fact. Iterated elimination of strictly dominated strategies require the following assumption. **Common Knowledge of Rationality (CKR)**: The fact that all players are rational is common knowledge.

Let us consider another example in Table 3. Strategy R is strictly dominated by Strategy M for Player 2. If Player 2 is rational, he does not play R. If Player 1 knows that Player 2 is rational and he himself is rational, then he will assume that R is not played, and T strictly dominates B after removing R. So, he will not play B. If Player 2 knows that Player 1 is rational and Player 2 knows that Player 1 knows Player 2 is rational, then he will not play L. So, iteratively deleting all strictly dominated strategies lead to a unique prediction of (T, M).

	L	M	R
T	$(1,0)$	$(1,2)$	$(0,1)$
B	$(0,3)$	$(0,1)$	$(2,0)$

Table 3: Domination

In many games, iterated elimination of strictly dominated strategies lead to a unique outcome of the game. In those cases, we call it a **solution** of the game. However, absence of strictly dominated strategies will imply that no strategies can be eliminated. In such a case, iterated elimination of strictly dominated strategies result in no solution. However, the order in which we eliminate strictly dominated strategies does not matter. A formal proof of this fact will be presented later.

In some games, there may not exist any strictly dominated strategy. In such a case, the following weaker notion of weak domination is considered.

Definition **3** *Strategy s_i of Player i is* **weakly dominated** *if there exists another strategy t_i of Player i such that for all $s_{-i} \in S_{-i}$, we have*

$$u_i(s_i, s_{-i}) \leq u_i(t_i, s_{-i}),$$

with strict inequality holding for at least one $s_{-i} \in S_{-i}$. In this case, we say that t_i weakly dominates s_i.

There is no foundation for eliminating (iteratively or otherwise) weakly dominated strategies. Indeed, if we remove weakly dominated strategies iteratively, then the order of elimination matters. This is illustrated in the following example in Table 4.

	L	C	R
T	$(1,2)$	$(2,3)$	$(0,3)$
M	$(2,2)$	$(2,1)$	$(3,2)$
B	$(2,1)$	$(0,0)$	$(1,0)$

Table 4: Order of elimination of weakly dominated strategies

The game in Table 4, there are two weakly dominated strategies for Player 1: $\{T, B\}$. Suppose Player 1 eliminates T first. Then, strategies in $\{C, R\}$ are weakly dominated for

Player 2. Suppose Player 2 eliminates R. Then, Player 1 eliminates the weakly dominated strategy B. Finally, Player 2 eliminates Strategy C to leave us with (M, L).

Now, suppose Player 1 eliminates B first. Then, both L and C are weakly dominated. Suppose Player 2 eliminates L first. Then, T is weakly dominated for Player 1. Eliminating T, we see that C is weakly dominated for Player 2. So, we are left with (M, R).

3.1 An Auction Example

In some games, weakly dominant strategies give striking prediction. One such example is given below.

The Vickrey Auction. An indivisible object is being sold. There are n buyers (players). Each buyer i has a value v_i for the object, which is completely known to the buyer. Each buyer is asked to report or bid a non-negative real number - denote the bid of buyer i as b_i. The highest bidder wins the object but asked to pay an amount equal to the second highest bid. In case of a tie, all the highest bidders get the object with equal probability and pay the second highest bid, which is also their bid amount in this case. Any buyer who does not win the object pays zero. If a buyer i wins the object and pays a price p_i, then his utility is $v_i - p_i$.

Lemma **2** *In the Vickrey auction, it is a weakly dominant strategy for every buyer to bid his value.*

Proof: Suppose for all $j \in N \setminus \{i\}$, buyer j bids an amount b_j. If buyer i bids v_i, then there are two cases to consider.

Case 1. $v_i > \max_{j \neq i} b_j$. In this case, the payoff of buyer i from bidding v_i is $v_i - \max_{j \neq i} b_j > 0$. By bidding something else, if he is not the unique highest bidder (i.e., either he shares the object or loses the object), then he either does not get the object or he gets the object with lower probability and pays the same amount. In the first case, his payoff is zero and in the second case, his payoff is strictly less than $v_i - \max_{j \neq i} b_j$. Hence, bidding v_i is a weakly dominant strategy.

Case 2. $v_i \leq \max_{j \neq i} b_j$. In this case, the payoff of buyer i from bidding v_i is zero - this is because either he is not getting the object (in which case his payoff is zero) or he is sharing the object in which case he is paying $\max_{j \neq i} b_j = v_i$. If he bids an amount smaller than v_i, then he does not get the object and his payoff is zero. If he bids an amount larger than v_i, then he gets the object with probability one and pays $\max_{j\neq i} b_j$, and hence, his payoff is $v_i - \max_{j\neq i} b_j \leq 0$. Hence, bidding v_i is a weakly dominant strategy for buyer i. ■

4 Nash Equilibrium

One of the problems with the idea of domination is that often there are no dominated strategies. Hence, it fails to provide any prediction about many games. For instance, consider the game in Table 5. No pure strategy in this game is dominated.

	a	b
A	$(3,1)$	$(0,4)$
B	$(0,2)$	$(3,1)$

Table 5: No dominated strategies

We now revisit the strong requirement of domination that a strategy is best irrespective of the beliefs we have about what others are playing. In many cases, games are results of repeated outcomes. For instance, if two firms are interacting in a market, they have a good idea about each other's cost and technology. As a result, they can form accurate beliefs about what other player is playing. The idea of Nash equilibrium takes this accuracy to the limit - it assumes that each player has **correct** belief about what others are playing and responds optimally given his (correct) beliefs.

Definition 4 *A strategy profile $(s_1^*, \ldots, s_n^*)$ in a strategic form game $\Gamma \equiv (N, \{S_i\}_{i\in N}, \{u_i\}_{i\in N})$ is a* **Nash equilibrium** *of Γ if for all $i \in N$*

$$u_i(s_i^*, s_{-i}^*) \geq u_i(s_i, s_{-i}^*) \ \forall \ s_i \in S_i.$$

The game Γ in the above definition may be a finite or an infinite game. The definition above requires that given strategies of other players s_{-i}^*, a unilateral deviation by Player i is not profitable. A belief based definition is also possible. We will say that a strategy profile $(s_1^*, \ldots, s_n^*)$ is a Nash equilibrium if for all $i \in N$,

$$\begin{aligned} \mu_i(s_{-i}^*) &= 1 \\ \mathcal{U}_i(s_i^*; \mu_i) &\geq \mathcal{U}_i(s_i; \mu_i) \qquad \forall \ s_i \in S_i. \end{aligned}$$

The idea of a Nash equilibrium is that of a *steady state*, where each player is responding optimally given the strategies of the other players - no unilateral deviation is possible. It does not argue how this steady state is reached. It has a notion of stability - if a player finds certain unilateral deviation profitable, then such a steady state cannot be sustained.

An alternate definition using the idea of *best response* is often useful. A strategy s_i of Player i is a **best response** to the strategy s_{-i} of other players if

$$u_i(s_i, s_{-i}) \geq u_i(s'_i, s_{-i}) \ \forall \ s'_i \in S_i.$$

The set of all best response strategies of Player i given the strategies of other players is denoted by $B_i(s_{-i})$.

Now, a strategy profile $(s_1^*, \ldots, s_n^*)$ is a **Nash equilibrium** if for all $i \in N$,

$$s_i^* \in B_i(s_{-i}^*).$$

The following observation is immediate.

CLAIM **1** *If s_i^* is a strictly dominant strategy of Player i, then $\{s_i^*\} = B_i(s_{-i})$ for all $s_{-i} \in S_{-i}$. Hence, if $(s_1^*, \ldots, s_n^*)$ is a strictly dominant strategy equilibrium, it is a* unique *Nash equilibrium.*

It is extremely important to remember that Nash equilibrium assumes correct beliefs and best responding with respect to these correct beliefs of other players. There are other interpretations of Nash equilibrium. Consider a mediator who offers the players a strategy profile to play. A player agrees with the mediator if (a) he believes that others will agree with the mediator and (b) strategy proposed to him by the mediator is a best response to the strategy proposed to others.

4.1 Examples

We give various examples of games where a Nash equilibrium (in pure strategies) exist. In Table 6, we consider the Prisoner's Dilemma game. By Claim 1, (A, a) is a Nash equilibrium of this game since it is the outcome in strictly dominant strategies.

Consider now the game (called the *coordination game*) in Table 7. The game is called coordination game since if players do not coordinate in this game they both get zero payoff. If they coordinate, then they get the same payoff but (A, a) is worse than (B, b) for both the players. If Player 2 plays a, then $B_1(a) = \{A\}$ and if Player 1 plays A, then $B_2(A) = \{a\}$. So, (A, a) is a Nash equilibrium. Now, if Player 2 plays b, then $B_1(b) = \{B\}$ and if Player 1

	a	b
A	$(1,1)$	$(5,0)$
B	$(0,5)$	$(4,4)$

Table 6: Nash equilibrium in Prisoner's Dilemma

	a	b
A	$(1,1)$	$(0,0)$
B	$(0,0)$	$(3,3)$

Table 7: Nash equilibrium in the Coordination game

plays B, then $B_2(B) = \{b\}$. Hence, (B,b) is another Nash equilibrium. This example shows you that there may be more than one Nash equilibrium in a game.

Another game that has more than one Nash equilibrium is the *Battle of the sexes*. A man and a woman are deciding which movie to go between two movies $\{X,Y\}$. Man wants to see movie X and woman wants to see movie Y. However, if both of them go to separate movies, then they get zero payoff. Their preferences are reflected in Table 8. If Woman plays x, then Man's best response is $\{X\}$ and if Man plays X, then Woman's best response is $\{x\}$. Hence, (X,x) is a Nash equilibrium. Using a similar logic, we can compute (Y,y) to be a Nash equilibrium. These are the only Nash equilibria of the game.

	x	y
X	$(2,1)$	$(0,0)$
Y	$(0,0)$	$(1,2)$

Table 8: Nash equilibrium in the Battle of the Sexes game

Now, we discuss a game with infinite number of strategies. This game is called the the Cournot Duopoly game. Two firms $\{1,2\}$ produce the same product in a market where there is a common price for the product. They simultaneously decide how much to produce - denote by q_1 and q_2 respectively the quantities produced by firms 1 and 2. If the total quantity produced by both the firms is $q_1 + q_2$, then the product price is assumed to be $2 - q_1 - q_2$. Suppose the per unit cost of productions are: $c_1 > 0$ for firm 1 and $c_2 > 0$ for firm 2. We will assume that $q_1, q_2, c_1, c_2 \in [0,1]$. We will now compute the Nash equilibrium of this game.

This is a two player game. Each player's strategy is the quantity it produces. If firms 1 and 2 produce q_1 and q_2 respectively, then their payoffs are given by

$$u_1(q_1, q_2) = q_1(2 - q_1 - q_2) - c_1 q_1$$
$$u_2(q_1, q_2) = q_2(2 - q_1 - q_2) - c_2 q_2.$$

Given q_2, firm 1 can maximize its payoff my maximizing u_1 over all q_1. To do so, we take the first order condition for u_1 to get

$$2 - 2q_1 - q_2 - c_1 = 0.$$

This simplifies to

$$q_1 = \frac{1}{2}(2 - c_1 - q_2).$$

Similarly, we get

$$q_2 = \frac{1}{2}(2 - c_2 - q_1).$$

Solving these two equations we get

$$q_1^* = \frac{2 - 2c_1 + c_2}{3}, q_2^* = \frac{2 - 2c_2 + c_1}{3}.$$

These are necessary conditions for optimality. Since the utility functions are strictly concave (verify this!), these will be the unique optimal solutions. We can also directly verify that it is a Nash equilibrium. For this, first note that

$$u_1(q_1^*, q_2^*) = (q_1^*)^2$$
$$u_2(q_1^*, q_2^*) = (q_2^*)^2$$

Now, given firm 2 sets q_2^*, let us find the utility when firm 1 sets q_1:

$$\begin{aligned} u_1(q_1, q_2^*) &= \frac{q_1}{3}\left[4 + 2c_2 - 4c_1 - 3q_1\right]. \\ &= 2q_1 q_1^* - (q_1)^2 \\ &\leq (q_1^*)^2 \\ &= u_1(q_1^*, q_2^*). \end{aligned}$$

A similar calculation suggests

$$u_2(q_1^*, q_2) \leq u_2(q_1^*, q_2^*).$$

Hence, (q_1^*, q_2^*) is a Nash equilibrium. This is also a unique Nash equilibrium (why?).

We now consider an example of a two-player game where payoffs of both the players add up to zero. This particular game is called the *matching pennies*. Two players toss two coins. If they both turn Heads or Tails, then Player 1 is paid by Player 2 Rs. 1. Else, Player 1 pays Player 2 Rs. 1. The payoff of each player is the money he receives (or the negative of the money he pays). The payoffs are shown in Table 9. For the moment assume that, what turns up in the coin is in the control of the players - for instance, a player may choose to show Heads in his coin.

The Matching Pennies game has no Nash equilibrium. To see this, note that when Player 2 plays h, then the unique best response of Player 1 is H. But when Player 1 plays H, the unique best response of Player 2 is t. Also, when Player 2 plays t the unique best response of Player 1 is T, but when Player 1 plays T the unique best response of Player 2 is h.

	h	t
H	$(1, -1)$	$(-1, 1)$
T	$(-1, 1)$	$(1, -1)$

Table 9: The Matching Pennies game

5 Elimination of Dominated Strategies

We now formally introduce the notion of iterated elimination of strictly dominated strategies. The definition below formalizes our notion of iteratively eliminating dominated strategies.

Definition **5** *The set $X \subseteq S$ of strategy profiles survives* **iterated elimination of strictly dominated strategies** *if $X \equiv \times_{j \in N} X_j$ and there is a collection $(\{X_j^t\}_{j \in N})_0^T$ of sets that satisfy for each $j \in N$ the following:*

- *$X_j^0 = S_j$ and $X_j^T = X_j$,*
- *$X_j^{t+1} \subseteq X_j^t$ for each $t < T$,*
- *for each $t < T$, every strategy in $X_j^t \setminus X_j^{t+1}$ is strictly dominated in the game $(N, \{X_i^t\}_i, \{u_i^t\}_i)$, where u_i^t is the restriction of u_i to strategy profiles in this game.*
- *No strategy in X_j^T is strictly dominated.*

Note that the definition does not require you to eliminate **all** the strictly dominated strategies in a stage of elimination.

The next result states that if we eliminate some strategies (dominated or not) of a player, then every Nash equilibrium of the original game that survived this elimination continues to be a Nash equilibrium of the new game.

LEMMA **3** *Let* Γ *be a finite game in strategic form and* Γ' *be a game derived from* Γ *by eliminating some of the strategies of each player. If* s^* *is a Nash equilibrium of* Γ *and* s^* *is available in* Γ'*, then* s^* *is a Nash equilibrium in* Γ'*.*

Proof: Let S'_i be the set of strategies remaining for each player i in Γ' and S_i be the set of original strategies in Γ for each player i. By definition,

$$u_i(s^*) \geq u_i(s_i, s^*_{-i}) \;\; \forall \; s_i \in S_i.$$

But $S'_i \subseteq S_i$ implies that $u_i(s^*) \geq u_i(s_i, s^*_{-i}) \;\; \forall \; s_i \in S'_i$. Hence, s^* is also a Nash equilibrium of Γ'. ■

Note that eliminating arbitrary strategies though will not eliminate original Nash equilibria, it may introduce new Nash equilibria. The following theorem shows that this is not possible if weakly dominated strategies are eliminated.

LEMMA **4** *Let* Γ *be a finite game in strategic form and* s_j *be a weakly dominated strategy for Player* j *in this game. Denote by* Γ' *the game derived by eliminating strategy* s_j *from* Γ*. Then, every Nash equilibrium of* Γ' *is also a Nash equilibrium of* Γ*.*

Proof: Let s^* be a Nash equilibrium of Γ'. Consider a player $i \neq j$. By definition, $u_i(s^*) = \max_{s_i \in S_i} u_i(s_i, s^*_{-i})$. Since the set of strategies of i is the same in both the games, this establishes that i cannot unilaterally deviate. For Player j, we note that s_j is weakly dominated, say by strategy t_j. Then,

$$u_j(s_j, s^*_{-j}) \leq u_j(t_j, s^*_{-j}) \leq \max_{s'_j \in S_j : s'_j \neq s_j} u_j(s'_j, s^*_{-j}) = u_j(s^*_j, s^*_{-j}),$$

where the last equality follows since s^* is a Nash equilibrium of Γ'. This shows that $u_j(s^*_j, s^*_{-j}) \geq u_j(s'_j, s^*_{-j})$ for all $s'_j \in S_j$. Hence, s^* is also a Nash equilibrium of Γ. ■

The above theorem implies that if we iteratively eliminate weakly dominated strategies and look at the Nash equilibria of the resulting game, they will also be Nash equilibria of the original game. However, we may lose some of the Nash equilibria of the original game. Consider the game in Table 10. Suppose Player 2 eliminates L and then Player 1 eliminates B. We are then left with (T, R). However, (B, L) is a Nash equilibrium of the original game. Note that (T, R) is also a Nash equilibrium of the original game (implied by Theorem 4).

However, this cannot happen if we eliminate strictly dominated strategies.

	L	R
T	$(0,0)$	$(2,1)$
B	$(3,2)$	$(1,2)$

Table 10: Elimination may lose equilibria

Theorem 1 *Let Γ be a finite game in strategic form and s_j be a strictly dominated strategy for Player j in this game. Denote by Γ' the game derived by eliminating strategy s_j from Γ. Then, the set of Nash equilibria in Γ and Γ' are the same.*

Proof: By Lemma 4, we need to show that if s^* is a Nash equilibrium of Γ, then s^* is also a Nash equilibrium of Γ'. Note that the strategy profile s^* is still available to all the agents in Γ' since only a strictly dominated strategy is eliminated for Player j. Formally, for Player j, there exists a strategy t_j such that $u_j(t_j, s^*_{-j}) > u_j(s_j, s^*_{-j})$. Hence, $u_j(s^*_j, s^*_{-j}) \geq u_j(t_j, s^*_{-j}) > u_j(s_j, s^*_{-j})$. So, $s^*_j \neq s_j$. Since s^* is available in Γ', by Lemma 3, s^* is a Nash equilibrium of game Γ'. ■

This theorem leads to some interesting corollaries. First, a strictly dominated strategy cannot be part of a Nash equilibrium. Second, if elimination of strictly dominated strategies lead to a unique outcome, then that outcome is the unique Nash equilibrium of the original game. In other words, to compute the Nash equilibrium or maxmin value, we can iteratively eliminate all strictly dominated strategies of the players.

6 Existence of Nash Equilibrium

In many games Nash equilibria exist. The question that we investigate in next three sections is the following:

> *What are some sufficient conditions on the game that ensures existence of Nash equilibrium?*

We will discuss three classes of games where we will show that a Nash equilibrium exists. The first two are somewhat technical in nature - but general enough to be applied to a large variety of games. The last one is somewhat simpler in nature, and the existence in those class of games were proved by Nash himself - in fact, the this class of games is a subclass of the first class of games, but even then we discuss it because of other reasons. All these classes of games have one thing in common: the strategy sets of each player has a lot of structure (geometrical) and the utility functions are well-behaved over these strategy sets.

An existence result is a technical result and may not appeal to everyone. However, it has its own beauty and importance. First, it shows that in some class of games, we can begin to think of computing and describing Nash equilibria. Second, it illustrates that the game is consistent with *some* steady state solution - though the precise steady state(s) are not found by proving an existence result.

All the existence results rely on some kind of **fixed point** result. We elaborate on this a little bit before proceeding further. Let X be some non-empty set and $f : X \to X$. We say $x \in X$ is a **fixed point** of f if

$$x = f(x).$$

A fixed point theorem identifies conditions on X and f such that a fixed point exists. These versions of fixed point theorems are indirectly useful - we will see the exact usefulness later (two of our existence results, including the original Nash result was proved by such existence results).

However, a set-theoretic version (or, correspondence version) of the fixed point theorem is immediately useful. As before, fix a set X and let $f : X \to 2^X$. So, for every $x \in X$, the function value $f(x)$ gives a subset of X. Such a function f has a **fixed point** $x \in X$ if

$$x \in f(x).$$

A fixed point theorem here would identify conditions on X and f such that a fixed point exists.

The usefulness of correspondence version of fixed point theorems is somewhat direct. Fix a strategy profile $s \in S$. Remember that the best response of agent i for s_{-i} is $B_i(s_{-i})$ and it gives all the strategies that maximize agent i's payoff against s_{-i}. Define the function $B : S \to 2^S$ as follows: for every $s \in S$,

$$B(s) = B_i(s_{-1}) \times \ldots \times B_n(s_{-n}).$$

We refer to B as the **best response correspondence**.

Take the game in Table 11. Consider the strategy profile $s \equiv (s_1 = M, s_2 = L)$. Now, $B_1(s_2) = \{T\}$ and $B_2(s_1) = \{C, R\}$. Hence,

$$B(s_1, s_2) = \{T\} \times \{C, R\} = \{(T, C), (T, R)\}.$$

The following claim establishes that such fixed point theorems will be useful for showing existence of Nash equilibrium.

	L	C	R
T	$(3,3)$	$(0,0)$	$(0,2)$
M	$(0,0)$	$(3,3)$	$(0,3)$
B	$(2,2)$	$(2,2)$	$(2,0)$

Table 11: Best response maps

CLAIM **2** *A Nash equilibrium exists if and only if the best response correspondence has a fixed point.*

Proof: If a Nash equilibrium s exists, then $s_i \in B_i(s_{-i})$ for all $i \in N$. Hence, $s \in B(s)$ - so, a fixed point of B exists. If a fixed point s of B exists, then $s \in B(s)$, which in turn implies that $s_i \in B_i(s_{-i})$. Hence, s is a Nash equilibrium. ■

Claim 2 forms the foundation for proving most of the existence results about Nash equilibrium. We will see this in next few sections.

6.1 Convex Strategy Sets with Concave and Continuous Utility Functions

The first such existence theorem is in a class of infinite games. The strategy space is assumed to have some geometric structure and the utility functions are assumed to be well-behaved.

THEOREM **2** *Suppose $\Gamma \equiv (N, \{S_i\}_{i\in N}, \{u_i\}_{i\in N})$ is a game in strategic form such that for each $i \in N$*

1. *S_i is a compact and convex subset of $\mathbb{R}^{K_i}$ for some integer K_i.*
2. *$u_i(s_i, s_{-i})$ is continuous in s_{-i}.*
3. *$u_i(s_i, s_{-i})$ is continuous and concave in s_i.* [1]

Then, Γ has a pure strategy Nash equilibrium.

Proof: The proof of this theorem is done using Kakutani's fixed point theorem.

THEOREM **3 (Kakutani's Fixed Point Theorem)** *Let A be a non-empty subset of a finite dimensional Euclidean space. Let $f : A \to 2^A$ be a map which satisfies the following properties.*

[1] A concave function is continuous in the interior of the domain. Requiring continuity here makes it continuous even at the boundary points.

1. *A is compact and convex.*

2. *$f(x)$ is a non-empty subset of A for each $x \in A$.*

3. *$f(x)$ is a convex subset of A for each $x \in A$.*

4. *$f(x)$ has a closed graph for each $x \in A$, i.e., if $\{x^k, y^k\} \to \{x, y\}$ with $y^k \in f(x^k)$ for each k, then $y \in f(x)$.*

Then, there exists $x \in A$ such that $x \in f(x)$.

We use Theorem 3 in a straightforward manner to establish existence of Nash equilibrium. For every strategy profile s, we know by Claim 2 that s is a Nash equilibrium if and only if s is a fixed point of the best response correspondence B. We show that B satisfies all the conditions of Theorem 3, and we will be done.

1. Since each S_i is compact and convex, the set of strategy profiles $S_1 \times \ldots \times S_n$ is also compact and convex.

2. For every s and for every $i \in N$,

$$B_i(s_{-i}) = \{s'_i \in S_i : u_i(s'_i, s_{-i}) = \max_{s''_i \in S_i} u_i(s''_i, s_{-i})\}.$$

   This set is non-empty because of u_i is continuous in s''_i and S_i is compact - so, by Weirstrass theorem a maximum of the function exists. As a result $B(s)$ is also non-empty.

3. Next, we show that $B(s)$ is convex. Pick, $t, t' \in B(s)$ and $\lambda \in (0, 1)$. Define $t'' \equiv \lambda t + (1 - \lambda)t'$. We show that for every $i \in N$, $t''_i \in B_i(s_{-i})$. Since $t_i, t'_i \in B_i(s_{-i})$, we get

$$u_i(t_i, s_{-i}) = u_i(t'_i, s_{-i}) = \max_{s'_i} u_i(s'_i, s_{-i}).$$

   But then concavity of u_i implies that

$$u_i(t''_i, s_{-i}) \geq \lambda u_i(t_i, s_{-i}) + (1 - \lambda)u_i(t'_i, s_{-i}) = \max_{s'_i} u_i(s'_i, s_{-i}).$$

   Hence, $t''_i \in B_i(s_{-i})$, and this implies that $B(s)$ is convex.

4. Finally, we show that B has a closed graph. To see this, assume for contradiction that B does not have a closed graph. Then, for some sequence $\{t^k, \bar{t}^k\} \to \{t, \bar{t}\}$ with $\bar{t}^k \in B(t^k)$, we have $\bar{t} \notin B(t)$. This means, for some $i \in N$, $\bar{t}_i \notin B_i(t_{-i})$. This implies that $u_i(\bar{s}_i, t_{-i}) > u_i(\bar{t}_i, t_{-i})$ for some $\bar{s}_i \in S_i$. The argument then follows from continuity of u_i in both his own strategy and the strategy of others. The precise argument is given below. But roughly, continuity ensures that we can find some point in the sequence t^k such that t^k is arbitrarily close to t and $u_i(\bar{s}_i, t_{-i})$ is arbitrarily close to $u_i(\bar{s}_i, t^k_{-i})$ and $u_i(\bar{t}_i, t_{-i})$ is arbitrarily close to $u_i(\bar{t}_i, t^k_{-i})$ - it is close enough such that $u_i(\bar{s}_i, t^k_{-i}) > u_i(\bar{t}_i, t^k_{-i})$ is maintained, and this is ensured by continuity with respect to other players' strategies. See Figure 1 for an illustration.

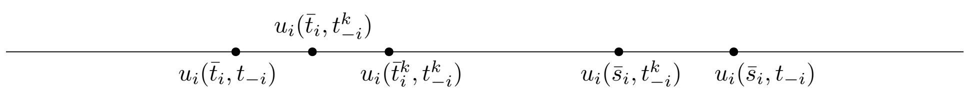

Figure 1: Illustration of proof of closed graph property

Now, we use continuity of u_i in i's strategy. Since $\bar{t}^k_i$ is arbitrarily close to $\bar{t}_i$, we conclude that $u_i(\bar{t}^k_i, t^k_{-i})$ is arbitrarily close to $u_i(\bar{t}_i, t_{-i})$ - it sufficiently close to maintain the relationship that

$$u_i(\bar{s}_i, t^k_{-i}) > u_i(\bar{t}^k_i, t^k_{-i}).$$

See Figure 1 for an illustration. But this contradicts the fact that $\bar{t}^k_i \in B_i(t^k_{-i})$.

Now, we turn to a somewhat precise argument. There exists some $\bar{s}_i \in S_i$ and $\epsilon > 0$ such that

$$u_i(\bar{s}_i, t_{-i}) > u_i(\bar{t}_i, t_{-i}) + \epsilon.$$

By the continuity of u_i, we get that there is some k such that t^k and t are close enough such that

$$u_i(\bar{s}_i, t^k_{-i}) \geq u_i(\bar{s}_i, t_{-i}) - \frac{\epsilon}{2}.$$

Combining these two inequalities we get

$$u_i(\bar{s}_i, t^k_{-i}) > u_i(\bar{t}_i, t_{-i}) + \frac{\epsilon}{2} \geq u_i(\bar{t}^k_i, t^k_{-i}) + \frac{\epsilon}{4},$$

where we used continuity of u_i in the second inequality. This is a contradiction because $\bar{t}^k_i \in B_i(t^k_{-i})$ implies $u_i(\bar{t}^k_i, t^k_{-i}) \geq u_i(s'_i, t^k_{-i})$.

Now, we apply Kakutani's fixed point theorem (Theorem 3) to conclude that there exists s such that $s \in B(s)$. This implies that s is a Nash equilibrium. ■

To see how Theorem 2 can and cannot be applied, consider the following location game. Two shops (players) are locating on the line segment $[0, 1]$ which has a uniform distribution of customers. Once the shops are located, customers go to the nearest shop with tie broken with equal probability. The utility of a shop is the mass of customers that go there. So, strategy sets of both the players are $S_1 = S_2 = [0, 1]$, a convex and compact set. If the shops locate themselves at (s_1, s_2) with $s_1 \leq s_2$, then the utilities of the shops are

$$u_1(s_1, s_2) = \frac{s_1 + s_2}{2}, u_2(s_1, s_2) = 1 - \frac{s_1 + s_2}{2}.$$

Hence, fixing s_2 as s_1 approaches s_2, we see that $u_1(s_1, s_2)$ approaches s_2 but as s_1 crosses s_2 for values arbitrarily close to s_2 it has a value of $1 - s_2$. Hence, u_1 is not continuous in s_1 for all values of $s_2 \neq \frac{1}{2}$. So, Theorem 2 cannot be applied here. But Nash equilibrium exists in such games - $s_1^* = s_2^* = \frac{1}{2}$ is a Nash equilibrium.

Second, consider the Cournot duopoly game with two firms. When firms produce q_1 and q_2, the price in the market is $2 - q_1 - q_2$ and unit costs of the firms are c_1 and c_2 respectively. Then, the utility function of each firm i is

$$u_i(q_1, q_2) = q_i(2 - q_1 - q_2) - c_i q_i.$$

This is continuous in both q_i and q_{-i}. Further, it is concave in q_i. Hence, it satisfies all the conditions of Theorem 2. Further, if we assume that the allowable quantities are some closed interval in the non-negative real line, then the strategy set of each firm is compact and convex. Theorem 2 guarantees that a Nash equilibrium exists.

7 Mixed Strategies

We now consider a game which is *derived* from a finite game. Formally, let

$$\Gamma := (N, \{S_i\}_{i \in N}, \{u_i\}_{i \in N}),$$

be a finite strategic form game (i.e., each S_i is finite). Consider the game derived from Γ by extending the strategy set of each player by allowing them to randomize over S_i.

Formally, the **mixed extension** of Γ is given by

$$\Delta\Gamma := (N, \{\Delta S_i\}_{i \in N}, \{U_i\}_{i \in N}),$$

where for all $i \in N$, the utility function U_i of Player i is a **linear extension** of his utility function u_i in Γ. In particular, if we consider a strategy profile $\sigma \in \in \prod_{i \in N} \Delta S_i$ in the mixed extension $\Delta\Gamma$, we have

$$U_i(\sigma) = \sum_{s\equiv(s_1,\ldots,s_n)\in S} u_i(s)\sigma_1(s_1)\ldots\sigma_n(s_n),$$

where $\sigma_i(s_j)$ is the probability with which Player i plays strategy s_j of game Γ in the strategy σ_i of game $\Delta\Gamma$. Note that the mixed extension of a game is an infinite game - it includes all possible lotteries over pure strategies of a player.

For any finite strategy set S_i of Player i, every $\sigma_i \in \Delta S_i$ is called a **mixed strategy** of Player i. In this case S_i is called the set of **pure strategies** of Player i. In other words, mixed strategies are all the strategies of a player in the mixed extension. A mixed strategy profile is $\sigma \equiv (\sigma_1, \ldots, \sigma_n) \in \prod_{i\in N} \Delta S_i$. Under mixed strategy, players are assumed to randomize independently, i.e., how a player randomizes does not depend on how others randomize.

Consider the following game in Table 12. Suppose Player 1 plays the mixed strategy A with probability $\frac{3}{4}$ and B with probability $\frac{1}{4}$. Suppose Player 2 plays a with probability $\frac{1}{4}$ and b with probability $\frac{3}{4}$. Then, the mixed strategy profile is

$$\sigma \equiv (\sigma_1, \sigma_2) = \Big((\sigma_1(A), \sigma_1(B)), (\sigma_2(a), \sigma_2(b))\Big) = \Big((\frac{3}{4}, \frac{1}{4}), (\frac{1}{4}, \frac{3}{4})\Big).$$

	a	b
A	$(3, 1)$	$(0, 0)$
B	$(0, 0)$	$(1, 3)$

Table 12: Mixed strategies

From this, the probability with which each pure strategy profile is played can be computed (using independence). These probabilities are shown in Table 13. A player computes the utility from a mixed strategy profile using expected utility. The mixed strategy profile σ gives players payoffs as follows:

$$\begin{aligned}
U_1(\sigma) &= u_1(A, a)\sigma_1(A)\sigma_2(a) + u_1(A, b)\sigma_1(A)\sigma_2(b) + u_1(B, a)\sigma_1(B)\sigma_2(a) + u_1(B, b)\sigma_1(B)\sigma_2(b) \\
&= 3\frac{3}{16} + 0 + 0 + 1\frac{3}{16} \\
&= \frac{3}{4} \\
U_2(\sigma) &= u_2(A, a)\sigma_1(A)\sigma_2(a) + u_2(A, b)\sigma_1(A)\sigma_2(b) + u_2(B, a)\sigma_1(B)\sigma_2(a) + u_2(B, b)\sigma_1(B)\sigma_2(b) \\
&= 1\frac{3}{16} + 0 + 0 + 3\frac{3}{16} \\
&= \frac{3}{4}.
\end{aligned}$$

	a	b
A	$\frac{3}{16}$	$\frac{9}{16}$
B	$\frac{1}{16}$	$\frac{3}{16}$

Table 13: Mixed strategies - probability of all pure strategy profiles

7.1 Extending the strategy space

Since $\Delta\Gamma$ is derived from Γ, the first question to ask is what happens to dominated and dominant strategies, and Nash equilibria of Γ when we consider $\Delta\Gamma$. This is a relevant question because the set of strategies in $\Delta\Gamma$ is larger than Γ. The following lemma is useful in understanding this relationship.

Lemma **5 (Indifference Principle)** *Suppose $\sigma_i \in B_i(\sigma_{-i})$ and $\sigma_i(s_i) > 0$. Then, $s_i \in B_i(\sigma_{-i})$.*

Proof: Suppose $\sigma_i \in B_i(\sigma_{-i})$. Let $S_i(\sigma_i) := \{s_i \in S_i : \sigma_i(s_i) > 0\}$. If $|S_i(\sigma_i)| = 1$, then the claim is obviously true. Else, pick $s_i, s_i' \in S_i(\sigma_i)$. We argue that $U_i(s_i, \sigma_{-i}) = U_i(s_i', \sigma_{-i})$. Suppose not and $U_i(s_i, \sigma_{-i}) > U_i(s_i', \sigma_{-i})$. Then,

$$
\begin{aligned}
U_i(\sigma_i, \sigma_{-i}) &= \sum_{s_i'' \in S_i(\sigma_i)} U_i(s_i'', \sigma_{-i})\sigma_i(s_i'') \\
&= U_i(s_i, \sigma_{-i})\sigma_i(s_i) + U_i(s_i', \sigma_{-i})\sigma_i(s_i') + \sum_{s_i'' \in S_i(\sigma_i)\setminus\{s_i, s_i'\}} U_i(s_i'', \sigma_{-i})\sigma_i(s_i'') \\
&< U_i(s_i, \sigma_{-i})\big(\sigma_i(s_i) + \sigma_i(s_i')\big) + \sum_{s_i'' \in S_i(\sigma_i)\setminus\{s_i, s_i'\}} U_i(s_i'', \sigma_{-i})\sigma_i(s_i'') \\
&= U_i(\sigma_i', \sigma_{-i}),
\end{aligned}
$$

where σ_i' is the new mixed strategy of Player i, where he plays s_i with probability $\sigma_i(s_i)+\sigma_i(s_i')$ and s_i' with probability zero, and every other strategy s_i'' in $S_i(\sigma_i)$ is played with probability $\sigma_i(s_i'')$. But this contradicts the fact that $\sigma_i \in B_i(\sigma_{-i})$. ■

This allows us to state the following straightforward results.

Theorem **4** *Suppose $s^* \equiv (s_1^*, \ldots, s_n^*)$ is a strategy profile in the finite game Γ. Then, the following are true.*

1. *If s^* is a Nash equilibrium of Γ, it is also a Nash equilibrium of the mixed extension $\Delta\Gamma$.*

2. *If s_i^* a weakly dominant strategy for Player i in Γ, it is also a weakly dominant strategy for Player i in $\Delta\Gamma$.*

3. *Every strictly dominant strategy of $\Delta\Gamma$ is a pure strategy, i.e., a strategy in Γ.*

Proof: PROOF OF (1). Suppose s^* is not a Nash equilibrium of $\Delta\Gamma$. Then, for some $i \in N$, $s_i^* \notin B_i(s_{-i}^*)$. But by Lemma 5, there is some strategy $s_i' \in S_i$ such that $s_i' \in B_i(s_{-i}^*)$. This means, $u_i(s_i', s_{-i}^*) > u_i(s_i^*, s_{-i}^*)$. This contradicts the fact that s^* is a Nash equilibrium of Γ.

PROOF OF (2). Suppose s_i^* a weakly dominant strategy for Player i in Γ. Suppose s_i^* a not a weakly dominant strategy for Player i in $\Delta\Gamma$. Then, for some σ_{-i}, Lemma 5 implies that there is a strategy s_i such that

$$U_i(s_i, \sigma_{-i}) > U_i(s_i^*, \sigma_{-i}).$$

But this implies that

$$\sum_{s_{-i}} \big[u_i(s_i, s_{-i}) - u_i(s_i^*, s_{-i})\big]\sigma_{-i}(s_{-i}) > 0,$$

where $\sigma_{-i}(s_{-i})$ is the probability with which strategy s_{-i} is played by Players in $N \setminus \{i\}$. But this implies that for some s_{-i}', we must have

$$u_i(s_i, s_{-i}') - u_i(s_i^*, s_{-i}') > 0.$$

This contradicts the fact that s_i^* is a weakly dominant strategy in Γ.

PROOF OF (3). Suppose σ_i is a strategy in $\Delta\Gamma$ but not in Γ (i.e., σ_i is **not** a pure strategy) and σ_i is strictly dominant in $\Delta\Gamma$. Then, by Lemma 5, there are two strategies $s_i \neq s_i'$ belonging to Γ such that $\sigma_i(s_i) > 0$ and $\sigma_i(s_i') > 0$, and for all s_{-i},

$$U_i(s_i, s_{-i}) = U_i(s_i', s_{-i}) = U_i(\sigma_i, s_{-i}).$$

Hence, σ_i is not strictly dominant. ■

Theorem 4 has consequences in computing a Nash equilibrium in the mixed extension of Γ. It says that we can compute Nash equilibria, weakly dominant strategies, strictly dominated strategies, and strictly dominant strategies of Γ, and they continue to maintain their properties in the mixed extension. The following remarks say that mixed extensions may create additional complications.

- A pure strategy that is not dominated by any pure strategy may be dominated by a mixed strategy. To see this, consider the example in Table 14. Strategy C is not dominated by any pure strategy for Player 1. However, the mixed strategy $\frac{1}{2}A$ and $\frac{1}{2}B$ strictly dominates the pure strategy C. Hence, C is a strictly dominated strategy for Player 1 in the mixed extension of the game described in Table 14.

	a	b
A	$(3,1)$	$(0,4)$
B	$(0,2)$	$(3,1)$
C	$(1,0)$	$(1,2)$

Table 14: Mixed strategies may dominate pure strategies

- Even if a group of pure strategies are not strictly dominated, a mixed strategy with only these strategies in its support may be strictly dominated. To see this, consider the game in Table 15. The pure strategies A and B are not strictly dominated. But the mixed strategy $\frac{1}{2}A + \frac{1}{2}B$ is strictly dominated by pure strategy C.

	a	b
A	$(3,1)$	$(0,4)$
B	$(0,2)$	$(3,1)$
C	$(2,0)$	$(2,2)$

Table 15: Mixed strategies may be dominated

7.2 Existence of Nash Equilibrium in Mixed Strategies

In this section, instead of talking about mixed extension of a game, we will refer to the mixed strategies of a player in a game explicitly. Now, we prove Nash's seminal theorem.

Theorem **5 (Nash)** *The mixed extension of every finite game has a Nash equilibrium.*

Note that this theorem is a corollary of our earlier existence theorem - Theorem 2. This is because, it is not difficult to check that the strategy space in the mixed extension of a finite game is a convex set, the utility functions are linear in strategies, and hence, continuous and concave as desired. The proof below is based on a weaker fixed theorem due to Brower. It is also based on the original proof of Nash, and has a useful technique that can be applied in other settings.

Proof: We do the proof in several steps.

STEP 1. For each profile of mixed strategy σ, for each player $i \in N$, and for each pure strategy $s_i \in S_i$, we define

$$g_i(s_i, \sigma) := \max \Big(0, U_i(s_i, \sigma_{-i}) - U_i(\sigma)\Big).$$

The interpretation of $g_i(s_i, \sigma)$ is that it is zero if Player i does not find deviating to s_i from σ profitable. Else, it captures the increase in payoff of Player i from (σ) to (s_i, σ_{-i}). Note that Player i can profitably deviate from σ if and only if it can profitably deviate from σ using a pure strategy - Lemma 5. This implies that σ is a Nash equilibrium if and only if $g_i(s_i, \sigma) = 0$ for all $i \in N$ and for all $s_i \in S_i$.

STEP 2. Now, we show that for each i and each s_i, $g_i(s_i, \cdot)$ is a continuous in the second argument. To see this note that U_i is continuous in σ and σ_{-i}. As a result, $U_i(s_i, \sigma_{-i}) - U_i(\sigma)$ is a continuous function. The max of two continuous functions is continuous. Hence, $g_i(s_i, \cdot)$ is continuous.

STEP 3. Using g_i, we define another map f_i in this step. For every $i \in N$, for every $s_i \in S_i$, and for every σ, define

$$f_i(s_i, \sigma) := \frac{\sigma_i(s_i) + g_i(s_i, \sigma)}{1 + \sum_{s'_i} g_i(s'_i, \sigma)}.$$

The amount $f_i(s_i, \sigma)$ is supposed to hint that if σ_i is not a better response to σ_{-i}, then how much probability on s_i should be assigned - thus, it gives another improved mixed strategy.

It is easy to see that for each i and each s_i, $f_i(s_i, \sigma) \geq 0$. Further,

$$\begin{aligned}
\sum_{s_i \in S_i} f_i(s_i, \sigma) &= \sum_{s_i \in S_i} \frac{\sigma_i(s_i) + g_i(s_i, \sigma)}{1 + \sum_{s'_i \in S_i} g_i(s'_i, \sigma)} \\
&= \frac{\sum_{s_i \in S_i} \sigma_i(s_i) + \sum_{s_i \in S_i} g_i(s_i, \sigma)}{1 + \sum_{s_i \in S_i} g_i(s_i, \sigma)} \\
&= 1.
\end{aligned}$$

Hence, $f_i(\sigma)$ is another mixed strategy of Player i. Further, f_i is a continuous function since both numerator and denominator are non-negative continuous functions. Hence, $f(\sigma) \equiv (f_1(\sigma), \ldots, f_n(\sigma))$ is also a continuous function.

STEP 4. We show that if $f(\sigma) = \sigma$, i.e., σ is a fixed point of f, then for all $i \in N$ and for all s_i,

$$g_i(s_i, \sigma) = \sigma_i(s_i) \sum_{s'_i \in S_i} g_i(s'_i, \sigma).$$

To see this, using the fixed point property and the definition of f_i, we see that

$$\begin{aligned} f_i(s_i, \sigma) &= \sigma_i(s_i) \\ &= \frac{\sigma_i(s_i) + g_i(s_i, \sigma)}{1 + \sum_{s'_i \in S_i} g_i(s'_i, \sigma)}. \end{aligned}$$

Rearranging, we get the desired equality.

STEP 5. In this step of the proof, we show that if σ is a fixed point of f, then σ is a Nash equilibrium. Suppose σ is not a Nash equilibrium. Then, for some Player i, there is a strategy s_i such that $g_i(s_i, \sigma) > 0$ - this uses Lemma 5 because we are claiming that a pure strategy gives more payoff. As a result $\sum_{s'_i \in S_i} g_i(s'_i, \sigma) > 0$. From the previous step, we know that $\sigma_i(s''_i) > 0$ if and only if $g_i(s''_i, \sigma) > 0$ for any s''_i.

Now, note that $U_i(\sigma) = \sum_{s'_i \in S_i} \sigma_i(s'_i) U_i(s'_i, \sigma_{-i})$. Hence,

$$\begin{aligned} 0 &= \sum_{s'_i} \sigma_i(s'_i)\big(U_i(s'_i, \sigma_{-i}) - U_i(\sigma)\big) \\ &= \sum_{s'_i : \sigma_i(s'_i) > 0} \sigma_i(s'_i)\big(U_i(s'_i, \sigma_{-i}) - U_i(\sigma)\big) \\ &= \sum_{s'_i : \sigma_i(s'_i) > 0} \sigma_i(s'_i) g_i(s'_i, \sigma) \\ &> 0, \end{aligned}$$

where the last equality and the strict inequality follows from our earlier observation that $g_i(s'_i, \sigma) > 0$ if and only if $\sigma_i(s'_i) > 0$.

STEP 6. This leads to the last step of the theorem. In this step, we show that a fixed point of f exists. For this, we use the following fixed point theorem due to Brouwer.

THEOREM **6 (Brouwer's fixed point theorem)** *Let X be a convex and compact set in $\mathbb{R}^k$ and let $F : X \to X$ be a continuous function. Then, there exists a fixed point of F.*

Now, we have already argued that f is a continuous function. The domain of f is the set of all strategy profiles. Since this is the set of all mixed strategies of a finite set of pure strategies, it is a compact and convex set. Finally, the range of f belongs to the set of strategy profiles. Hence, by Brouwer's fixed point theorem, there exists a fixed point of f. By the previous step, such a fixed point corresponds to the Nash equilibrium of the finite game. ■

The Brouwer's fixed point theorem is not simple to prove, but you are encouraged to look at its proof. In one-dimension, the Brouwer's fixed point theorem is the *intermediate value theorem.*

7.3 Interpretations of Mixed Strategy Equilibrium

Considering mixed strategies guarantee existence of Nash equilibrium in finite games. However, it is not clear why a player will randomize in the precise way prescribed by a mixed strategy Nash equilibrium, specially given the fact he is indifferent between the pure strategies in the support of such a Nash equilibrium. There are no clear answers to this question. However, following are some arguments to validate that mixed strategies can be part of Nash equilibrium play.

- Players randomize deliberately. For instance, in zero-sum games with two players, players randomize to play their max min strategies. In games like Poker, players have been shown to randomize.

- Mixed strategy equilibrium can be thought to be a belief system - if σ^* is a Nash equilibrium, then σ_i^* describes the belief that opponents of Player i have on Player i's behavior. This means that Player i may not actually randomize but his opponents collectively believe that σ_i^* is the strategy he will play. Hence, a mixed strategy equilibrium is just a steady state of beliefs.

- One can think of a strategic form game being played over time repeatedly (payoffs and actions across periods do not interact). Suppose players choose a best response in each period assuming time average of plays of past (with some initial conditions on how to choose strategies). In particular, they observe that opponents have been playing a strategy A for $\frac{3}{4}$ times and another strategy B for the remaining time. So, they optimally respond by forming this as their belief. It has been shown that such plays eventually converge to a steady state where the average play of each player is some mixed strategy.

- Another interpretation that is provided by Nash himself interprets Nash equilibrium as population play. There are two pools of large population. We draw a player at random from each pool and pair them against each other. The strategy of that player will reflect the expected strategy played by the population and will represent a mixed strategy. So, Nash equilibrium represents some kind of stationary distribution of pure strategies in such population.

7.4 Computing Mixed Strategy Equilibrium - Examples

In general, computing mixed strategy equilibrium of a finite game is computationally difficult. However, couple of thumb-rules make it easier for finding the set of all Nash equilibria. First, we should iteratively eliminate all strictly dominated strategies. As we have learnt, the set of Nash equilibria remains the same after iteratively eliminating strictly dominated strategies. The second is a crucial property that we have already established - the indifference principle in Lemma 5.

We start off by a simple example on how to compute all Nash equilibria of a game. Consider the game in Table 16.

	L	R
T	$(8,8)$	$(8,0)$
B	$(0,8)$	$(9,9)$

Table 16: Nash equilibria computation

First, note that no strategies can be eliminated as strictly dominated. It is easy to verify that (T, L) and (B, R) are two pure strategy Nash equilibria of the game. To compute mixed strategy Nash equilibria, suppose Player 1 plays T with probability p and B with probability $(1-p)$, where $p \in (0,1)$. Then, by playing L, Player 2 gets

$$8p + 8(1-p) = 8.$$

By playing R, Player 2 gets

$$9(1-p).$$

L is best response to $pT + (1-p)B$ if and only if $8 \geq 9(1-p)$ or $p \geq \frac{1}{9}$. Else, R is a best response. Note that Player 2 is indifferent between L and R when $p = \frac{1}{9}$ - this follows from the indifference lemma that we have proved. Hence, if Player 2 mixes, then Player 1 must play $\frac{1}{9}T + \frac{8}{9}B$. But, when Player 2 plays $qL + (1-q)R$, then Player 1 gets 8 by playing

T and $9(1-q)$ by playing B. For Player 1 to mix, Player 2 must make him indifferent between playing T and B, which happens at $q = \frac{1}{9}$. Thus, $(\frac{1}{9}T + \frac{8}{9}B, \frac{1}{9}L + \frac{8}{9}R)$ is also a Nash equilibrium of this game. Note that the payoff achieved by both the players by playing this strategy profile is 8.

There are some strategies of a player which are not strictly dominated, but which can still be eliminated before computing the Nash equilibrium. These are strategies which are *never* best responses.

Definition **6** *A strategy $\sigma_i \in \Delta S_i$ is* **never a best response** *for Player i if for every $\sigma_{-i} \in \Delta S_{-i}$,*

$$\sigma_i \notin B_i(\sigma_{-i}).$$

The following claim is a straightforward observation.

Claim **3** *If a strategy is strictly dominated, then it is never a best response.*

The next claim says that we can remove all pure strategies that are not best responses to compute Nash equilibrium.

Lemma **6** *If a pure strategy $s_i \in S_i$ is never a best response, then any mixed strategy σ_i with $\sigma_i(s_i) > 0$ is not a Nash equilibrium strategy.*

Proof: Suppose $s_i \in S_i$ is never a best response but there is a mixed strategy Nash equilibrium σ with $\sigma_i(s_i) > 0$. By the Indifference Lemma (Lemma 5), s_i is also a best response to σ_{-i}, contradicting the fact s_i is never a best response. ■

The connection between never best response strategies and strictly dominated strategies is deeper. Indeed, in two-player games, a strategy is strictly dominated if and only if it is never a best response. We will come back to this once we discuss zero-sum games. We will now use Lemma 6 to compute Nash equilibria efficiently.

Consider the two player game in Table 17. Computing Nash equilibria of such a game can be quite tedious. However, we can be smart in avoiding certain computations.

	L	C	R
T	$(3,3)$	$(0,0)$	$(0,2)$
M	$(0,0)$	$(3,3)$	$(0,2)$
B	$(2,2)$	$(2,2)$	$(2,0)$

Table 17: Nash equilibria computation

In two player 3-strategy games, we can draw the best response correspondences in a 2-d simplex - Figure 2 represents the simplex of Player 1's strategy space for the game in Table 17. Any point inside the simplex represents a probability distribution over the three strategies of Player 1, and these probabilities are given by the lengths of perpendiculars to the three sides. To see this suppose we pick a point in the simplex with lengths of perpendiculars to sides $(T, B), (T, M), (M, B)$ as p_m, p_b, p_t respectively. The following fact from Geometry is useful.

Fact **1** *For every point inside an equilateral triangle with lengths of perpendiculars* (p_m, p_b, p_t), *the sum of* $p_m + p_b + p_t$ *equals to* $\sqrt{3}a/2$, *where* a *is the length of sides of the equilateral triangle.*

This fact can be proved easily by using the fact the sum of three triangles generated by any point is the same - $\sqrt{3}a^2/4 = \frac{1}{2}a(p_m + p_t + p_b)$. Hence, without loss of generality, we will scale the lengths of the sides of the simplex to $\frac{2}{\sqrt{3}}$. As a result, $p_m + p_t + p_b = 1$ and the numbers p_m, p_t, p_b reflect a probability distribution. We will follow this term to represent strategies in two player 3-strategy games.

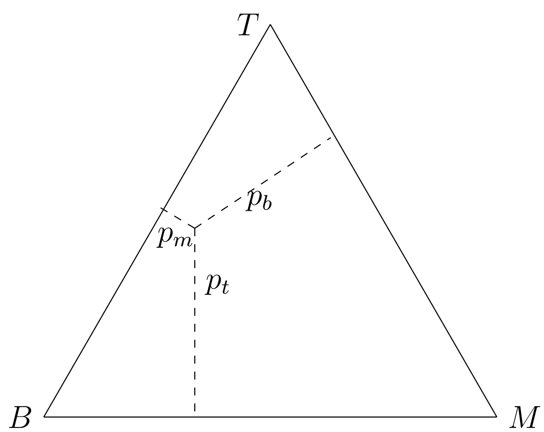

Figure 2: Representing probabilities on a 2d-simplex

Now, let us draw the best response correspondence of Player 1 for various strategies of Player 2: $B_1(\sigma_2)$ will be drawn on the simplex of strategies of Player 2 - see Figure 3. For this, we fix a strategy $\sigma_2 = (\alpha L + \beta C + (1-\alpha-\beta)R)$ of Player 2. We now identify conditions on α and β to identify pure strategy best responses of Player 1. By the Indifference Lemma, the mixed strategy best responses happen at the intersection of these pure strategy best response regions. We consider three cases:

Case 1- T. $T \in B_1(\sigma_2)$ if

$$3\alpha \geq 3\beta$$
$$3\alpha \geq 2.$$

Combining these conditions together, we get $\alpha \geq \frac{2}{3}$ and $\alpha \geq \beta$. The second condition holds if $\alpha \geq \frac{2}{3}$. So, we deduce that the best response region of T are all mixed strategies where L is played with at least $\frac{2}{3}$ probability. This is shown in Figure 3.

Case 2 - M. $M \in B_1(\sigma_2)$ if

$$3\beta \geq 3\alpha$$
$$3\beta \geq 2.$$

This gives us a similar condition to Case 1: $\beta \geq \frac{2}{3}$. The best response region of M is shown in the simplex of Player 2's strategies in Figure 3.

Case 3 - B. Clearly $B \in B_1(\sigma_2)$ in the remaining regions and at all the boundary points where B and T are indifferent and B and M are indifferent. This is shown in Figure 3 in the simplex of Player 2's strategy.

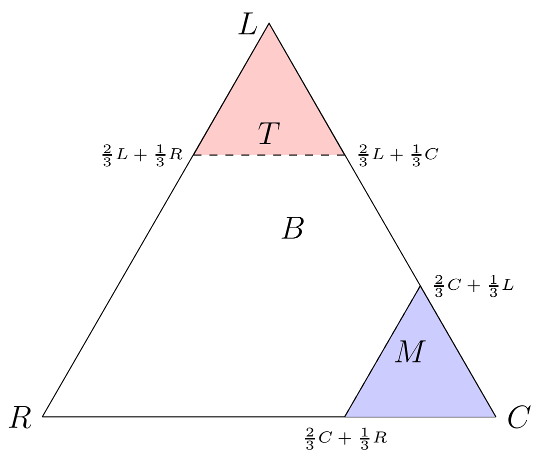

Figure 3: Best response map of Player 1

Once the best response map of Player 1 is drawn, we conclude that no best response involves mixing T and M together. So, every mixed strategy best response involves mixing B.

We now draw the best response map of Player 2. For this we consider a mixed strategy $\alpha T + \beta M + (1 - \alpha - \beta)B$ of Player 1. For L to be a best response of Player 2 against this strategy, we must have

$$
\begin{aligned}
3\alpha + 2(1 - \alpha - \beta) &\geq 3\beta + 2(1 - \alpha - \beta) \\
3\alpha + 2(1 - \alpha - \beta) &\geq 2(\alpha + \beta).
\end{aligned}
$$

This gives us

$$
\begin{aligned}
\alpha &\geq \beta \\
2 &\geq \alpha + 4\beta.
\end{aligned}
$$

The line $\alpha = \beta$ is shown in Figure 3. To draw $2 = \alpha + 4\beta$, we pick two points: (i) $\alpha = 0$ and $\beta = \frac{1}{2}$ and (ii) $\alpha + \beta = 1$ and $\beta = \frac{2}{3}$. The line joining these two points depict $2 = \alpha + 4\beta$. Now, the entire best response region of L is shown in Figure 3.

An analogous argument shows that for C to be a best response we must have

$$
\begin{aligned}
\beta &\geq \alpha \\
2 &\geq \beta + 4\alpha.
\end{aligned}
$$

The best response region of strategy C is shown in Figure 4. The remaining area is the best response region of strategy R (including the borders with L and C).

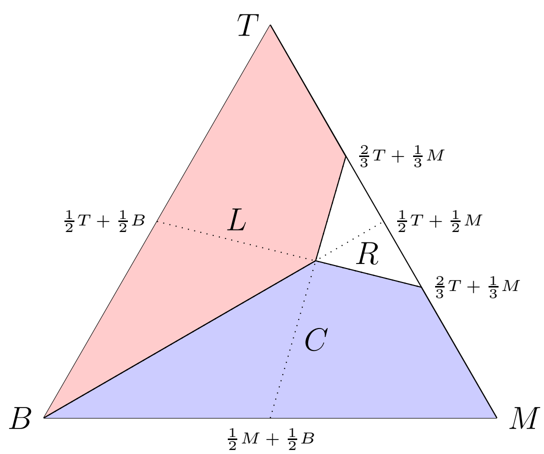

Figure 4: Best response map of Player 2

Computing Nash equilibria. To compute Nash equilibria, we see that there is no best response of Player 1 where T and M are mixed. Further, R is a best response of Player 2 when T and M are mixed. Hence, there cannot be a Nash equilibrium (σ_1, σ_2) such that $\sigma_2(R) > 0$. So, in any Nash equilibrium, Player 2 either plays L or C or mixed L and C but puts zero probability on R.

Since no mixing of T and M is possible for Player 1 in Nash equilibrium, we must look at the best response map of Player 2 when mix of T and B and mix of M and B is played. That corresponds to the two edges of the simplex corresponding to (T, B) and (M, B) in Figure 4. In that region, mixture of L and C is a best response when B is played with probability 1. So, in any Nash equilibrium where L and C is mixed Player 1 plays B for sure. But then looking into the best response map of Player 1 in Figure 3, we see that Player 1 best responds B for sure if Player 2 mixes $\alpha L + (1-\alpha)C$ with $\alpha \in [\frac{1}{3}, \frac{2}{3}]$. The other pure strategy Nash equilibria are (T, L) and (M, C).

So, we can enumerate all the Nash equilibria of the game in Table 17 now:

$$(T, L), (M, C), (B, \alpha L + (1-\alpha)C),$$

where $\alpha \in [\frac{1}{3}, \frac{2}{3}]$.

8 Two Player Zero-Sum Games

The two player zero-sum games occupy an important role in game theory because of variety of reasons. First, they were the first set of games to be theoretically analyzed by von-Neumann and Morgenstern when they came up with the theory of games. Second, the zero-sum games are ubiquitous - examples include any real game where one player's loss is another player's gain. Before formally introducing the notion of a zero-sum game, we describe another concept that we use here.

8.1 The Maxmin Value

Consider a game shown in Table 18. There is a unique Nash equilibrium of this game: (B, R) - verify this. But, will Player 1 play strategy B? What if Player 2 makes a mistake in his belief and plays L? Then, Player 1 will get -100 by playing B. Thinking this, Player 1 may like to play safe, and play a strategy like T that guarantees him a payoff of 2. For Player 2 also, strategy R may be bad if Player 1 decides to play T. On the other hand, strategy L can guarantee him a payoff of 0.

The main message of the example is that sometimes players may choose to play strategy to guarantee themselves some *safe* level of payoff without assuming anything about the rationality level of other players. In particular, we consider the case where every player

	L	R
T	$(2,1)$	$(2,-20)$
M	$(3,0)$	$(-10,1)$
B	$(-100,2)$	$(3,3)$

Table 18: The Maxmin idea

believes that the other players are *adversaries* and are here to punish him - this is a very pessimistic view of the opponents. In such a case, what can a player guarantee for himself?

If Player i chooses a strategy $s_i \in S_i$ in a game, then the worst payoff he can get is

$$\min_{s_{-i} \in S_{-i}} u_i(s_i, s_{-i}).$$

Of course, we are assuming here that the strategy sets and the utility functions are such that a minimum exists - else, we can define an infimum.

DEFINITION **7** *The* **maxmin** *value for Player i in a strategic form game* $(N, \{S_i\}_{i \in N}, \{u_i\}_{i \in N})$ *is given by*

$$\underline{v}_i := \max_{s_i \in S_i} \min_{s_{-i} \in S_{-i}} u_i(s_i, s_{-i}).$$

Any strategy that guarantees Player i a value of $\underline{v}_i$ *is called a* **maxmin** *strategy.*

Note that the above definition allows us to consider games which are mixed extensions of some finite game too. In that case, the max and min over strategy space is well defined because the set of strategies is a compact space and the utility function is linear in (mixed) strategies.

If s_i is a maxmin strategy for Player i, then it satisfies

$$\min_{s_{-i} \in S_{-i}} u_i(s_i, s_{-i}) \geq \min_{s_{-i} \in S_{-i}} u_i(s'_i, s_{-i}) \quad \forall \ s'_i \in S_i.$$

This also means that $u_i(s_i, s_{-i}) \geq \underline{v}_i$ for all $s_{-i} \in S_{-i}$.

In the example in Table 18, we see that $\underline{v}_1 = 2$ and $\underline{v}_2 = 0$. Strategy T is a maxmin strategy for Player 1 and strategy L is a maximin strategy for Player 2. Hence, when players play their maxmin strategy, the outcome of the game is $(2, 1)$. However, there can be more than one maxmin strategies in a game, in which case no unique outcome can be predicted. Consider the example in Table 19. The maxmin strategy for Player 1 is B. But Player 2 has two maxmin strategies $\{L, R\}$, both giving a payoff of 1. Depending on which maxmin strategy Player 2 plays the outcome can be $(2, 3)$ or $(1, 1)$.

	L	R
T	$(3,1)$	$(0,4)$
B	$(2,3)$	$(1,1)$

Table 19: More than one maxmin strategy

It is clear that if a player has a weakly dominant strategy, then it is a maxmin strategy - it guarantees him the best possible payoff irrespective of what other agents are playing. Hence, if every player has a weakly dominant strategy, then the vector of weakly dominant strategies constitute a vector of maxmin strategies. This was true, for instance, in the example involving the second-price sealed-bid auction. Further, if there are strictly dominant strategies for each player (note such strategy must be unique for each player), then the vector of strictly dominant strategies constitute a unique vector of maxmin strategies.

The following theorem shows that a Nash equilibrium of a game guarantees the maxmin value for every player.

Theorem 7 *Every Nash equilibrium s^* of a strategic form game satisfies*

$$u_i(s^*) \geq \underline{v}_i \quad \forall \; i \in N.$$

Proof: For any Player i and for every $s_i \in S_i$, we know that

$$u_i(s_i, s^*_{-i}) \geq \min_{s_{-i} \in S_{-i}} u_i(s_i, s_{-i}).$$

By definition, $u_i(s^*_i, s^*_{-i}) = \max_{s_i \in S_i} u_i(s_i, s^*_{-i})$. Combining with the above inequality, we get

$$u_i(s^*_i, s^*_{-i}) = \max_{s_i \in S_i} u_i(s_i, s^*_{-i}) \geq \max_{s_i \in S_i} \min_{s_{-i} \in S_{-i}} u_i(s_i, s_{-i}) = \underline{v}_i.$$

■

8.2 Zero-sum games

We now look into two-player zero-sum games. Formally, a zero-sum game is defined as follows.

Definition 8 *A finite* **zero-sum game of two players** *is defined as $N = \{1, 2\}$ and (S_1, S_2), (u_1, u_2) with the restriction that for all $(s_1, s_2) \in S_1 \times S_2$, we have*

$$u_1(s_1, s_2) + u_2(s_1, s_2) = 0.$$

Because of this restriction, we can define a zero-sum two player game by a single utility function $u : S_1 \times S_2 \to \mathbb{R}$, where $u(s_1, s_2)$ represents utility of Player 1 and $-u(s_1, s_2)$ represents the utility of Player 2.

	h	t
H	$(1, -1)$	$(-1, 1)$
T	$(-1, 1)$	$(1, -1)$

Table 20: Matching pennies

Consider the two player zero-sum game in Table 20. It is called the *matching pennies* game - the strategies are sides of a coin, if the sides match then Player 1 wins and pays Player 2 Rs. 1, else Player 2 wins and pays Player 1 Rs. 1. There is no pure strategy Nash equilibrium of this game. However, once we start looking at its mixed extension, we observe some interesting facts. Suppose Player 2 plays $\alpha h + (1 - \alpha)t$. To make Player 1 indifferent between H and T, we see that

$$\alpha + (-1)(1 - \alpha) = -\alpha + (1 - \alpha).$$

This gives us $\alpha = \frac{1}{2}$. A similar calculation suggests that if Player 2 has to mix in best response, Player 1 must play $\frac{1}{2}H + \frac{1}{2}T$. Hence, $(\frac{1}{2}H + \frac{1}{2}T, \frac{1}{2}h + \frac{1}{2}t)$ is the unique mixed strategy Nash equilibrium of this game. Note that the payoff achieved by both the players in this Nash equilibrium is zero.

Now, suppose Player 1 plays $\frac{1}{2}H + \frac{1}{2}T$, the worst payoff that he can get from Player 2's strategies (in the mixed extension) can be computed as follows. If Player 2 plays h or t Player 1 gets a payoff of 0. Hence, his worst payoff is 0. As a result, the maxmin value of Player 1 is at least zero. We know (by Theorem 7) that the Nash equilibrium payoff is at least the maxmin value. [2] Hence, the maxmin value is also zero. A similar calculation suggests that the maxmin value of Player 2 is also zero. We show that this is true for *any* finite two player zero-sum game.

The maxmin value of Player 1 in a zero sum game is denoted by

$$\underline{v}_1 := \max_{\sigma_1 \in \Delta S_1} \min_{\sigma_2 \in \Delta S_2} u(\sigma_1, \sigma_2).$$

The maxmin value of Player 2 in a zero sum game is denoted by

$$\underline{v}_2 := \max_{\sigma_2 \in \Delta S_2} \min_{\sigma_1 \in \Delta S_1} -u(\sigma_1, \sigma_2) = - \min_{\sigma_2 \in \Delta S_2} \max_{\sigma_1 \in \Delta S_1} u(\sigma_1, \sigma_2).$$

[2] Theorem 7 continues to hold even we allow consider the mixed extension of a finite game.

Any maxmin and minmax strategies of Player 1 and Player 2 respectively are called **optimal** strategies.

The main result for two person zero-sum game is the following.

THEOREM 8 *Every two player zero-sum game satisfies* $\underline{v}_1 + \underline{v}_2 = 0$. *The payoff from any Nash equilibrium* (σ_1^*, σ_2^*) *corresponds to* $(\underline{v}_1, \underline{v}_2)$. *Further, if* $(\sigma_1^*, \sigma_2)^*$ *is a Nash equilibrium, they are also the optimal (max-min) strategies.*

Proof: Mixed extension of every game has a Nash equilibrium. Hence, a two-player zero sum game will have a mixed strategy Nash equilibrium. We will show that the payoff from *any* Nash equilibrium (σ_1^*, σ_2^*) corresponds to $(\underline{v}_1, \underline{v}_2)$. Since, it is a zero-sum game

$$\underline{v}_1 + \underline{v}_2 = u_1(\sigma_1^*) + u_2(\sigma_2^*) = 0.$$

So, what remains to be shown is that $u_1(\sigma_1^*) = \underline{v}_1$ and $u_2(\sigma_2^*) = \underline{v}_2$. To see this, note that Nash equilibrium implies,

$$u(\sigma_1^*, \sigma_2^*) = \max_{\sigma_1 \in \Delta S_1} u(\sigma_1, \sigma_2^*) \geq \min_{\sigma_2 \in \Delta S_2} \max_{\sigma_1 \in \Delta S_1} u(\sigma_1, \sigma_2) = -\underline{v}_2.$$

Note that by Theorem 7, $-u(\sigma_1^*, \sigma_2^*) \geq \underline{v}_2$. Hence, we have

$$u(\sigma_1^*, \sigma_2^*) = -\underline{v}_2 = \max_{\sigma_2} \min_{\sigma_1} u_2(\sigma_1, \sigma_2).$$

Next, Nash equilibrium also implies that for all $\sigma_2 \in \Delta S_2$, we have $-u(\sigma_1^*, \sigma_2^*) \geq -u(\sigma_1^*, \sigma_2)$. Hence,

$$u(\sigma_1^*, \sigma_2^*) = \min_{\sigma_2 \in \Delta S_2} u(\sigma_1^*, \sigma_2) \leq \max_{\sigma_1 \in \Delta S_1} \min_{\sigma_2 \in \Delta S_2} u(\sigma_1, \sigma_2) = \underline{v}_1.$$

By Theorem 7, $u(\sigma_1^*, \sigma_2^*) \geq \underline{v}_1$. Hence, we get

$$-\underline{v}_2 = u(\sigma_1^*, \sigma_2^*) = \underline{v}_1 = \max_{\sigma_1} \min_{\sigma_2} u(\sigma_1, \sigma_2).$$

Hence, $\underline{v}_1 + \underline{v}_2 = 0$ and σ_1^* and σ_2^* are optimal strategies. ■

9 Correlated Equilibrium

Consider the mixed extension of the following game - usually called the game of "chicken". There are two players - $N = \{1, 2\}$. Player 1 has two pure strategies $S_1 = \{T, B\}$ and Player 2 has two pure strategies $S_2 = \{L, R\}$. The payoffs are shown in Table 21. The story that accompanies this game is that two drivers are racing towards each other on a single lane.

Each driver can either stay on or move away from the road. If both move away, then they get a payoff of 6 each. If both stay on, then they get a payoff of 0. If one of them stays on but the other moves away, then the one who stays on gets a payoff of 7 but the other one gets a payoff of 2.

	L	R
T	$(6,6)$	$(2,7)$
B	$(7,2)$	$(0,0)$

Table 21: Game of chicken

There are three Nash equilibria of this game: $(T,R),(B,L),\left(\frac{2}{3}T+\frac{1}{3}B,\frac{2}{3}L+\frac{1}{3}R\right)$. Notice that the mixed strategy Nash equilibrium puts a probability of $\frac{1}{9}$ with which the worst possible payoff profile (B,R) will be played. Now, consider the following "extended" game. There is an outside observer. The observer recommends each player *privately* a pure strategy to play. Note that no player observes the recommendation of the other player. Given his own recommended strategy, a player forms belief about the recommended strategy of the other player, assuming that the other player follows the recommendation. He follows his recommended strategy if and only if it is a best response given his belief about other player's recommended strategy.

Two natural confusions arise - (a) How does the observer recommend? and (b) How do the players form beliefs? It is assumed that the observer has access to a randomization device which is public, i.e., players know the distribution from which the recommendations are derived. Given the distribution of recommendation, players form beliefs by using Bayes' rule - they compute conditional probabilities.

In the game in Table 21, suppose the observer recommends pure strategy profiles in Nash equilibrium: (T,R) and (B,L) with probability p and $(1-p)$. Then, given his recommended strategy each player can uniquely infer the recommended strategy of the other player. Player 1 gets a recommendation of T means, Player 2 must have received a recommendation of R. So, Player 1 forms a belief that Player 2 plays R with probability 1. But (T,R) is a Nash equilibrium means, T is a best response to R. A similar logic shows that Player 1 will also accept B if it is recommended. Same argument applies to Player 2. Hence, *any* convex combination of pure strategy Nash equilibrium can be sustained as a *correlated* equilibrium of this extended game. In particular $p(T,R)+(1-p)(B,L)$ for any p is an equilibrium of this game. The set of payoffs that can be obtained are convex combination of $(7,2)$ and $(2,7)$.

Can we get other equilibrium? Suppose the observer recommends $(T,R),(B,L)$, and (T,L) with probability $\frac{1}{3}$ each. Then, if Player 1 observes T as a recommendation, then he can infer that Player 2 will have R as recommendation with probability $\frac{1}{2}$ and L as recommendation with probability $\frac{1}{2}$. Hence, he forms belief that Player 2 plays $\frac{1}{2}R + \frac{1}{2}L$. Is T a best response of Player 1 to this strategy? Playing T gives him 4 and playing B gives him 3.5. So, T is a best response, and Player 1 accepts the recommendation. If Player 1 receives B as a recommendation, then he forms a belief that Player 2 must receive L as recommendation. Since (B, L) is a Nash equilibrium, B is a best response to L. For Player 2, if he receives R as a recommendation, then he infers Player 1 must have received T and that being a Nash equilibrium, he accepts the recommendation. If Player 2 receives L as a recommendation, then he believes Player 1 must have received T as recommendation with probability $\frac{1}{2}$ and B as recommendation with probability $\frac{1}{2}$. Indeed, L is a best response to this strategy. Hence, both the players agree to accept the recommendations of the observer using this randomization device. The equilibrium payoff of both players from this is $(5, 5)$ which could not be obtained if we just randomize over Nash equilibria. Hence, an observer using a public randomizing device allows players to get payoff outside the convex hull of Nash equilibrium payoffs.

As the previous example illustrated, using public randomization allowed the players to avoid the worst payoff $(0, 0)$ by putting zero probability on that profile. This is impossible in a mixed strategy - independent randomization. To be able to play strategy profile (T, R), Player 2 must play R with some probability and that will mean playing (B, R) with some probability.

9.1 Correlated Strategies

A crucial assumption in mixed strategies is that players randomize independently. Each of them have access to a randomizing device (say, a coin to toss or a random number generating computer program) and these devices are independent. In some circumstances, players may have access to the same randomizing device. For instance, players observe some common event in the nature and decide to play their strategies based on this common event - say weather in a particular area.

Consider the same example in Table 12. Suppose Player 1 plays A and Player 2 plays a if it rains and Player 1 plays B and Player 2 plays b if it does not rain. Suppose the probability of rain is $\frac{1}{2}$. This means that the strategy profiles (A, a) and (B, b) is played with probability $\frac{1}{2}$ each but other strategy profiles are played with zero probability. There is strong correlation between the strategies played by both the players. Formally, a correlated strategy ρ is a map $\rho : S \to [0, 1]$ with $\sum_{s \in S} \rho(s) = 1$. The correlated strategy discussed above is shown in Table 22.

	a	b
A	$\frac{1}{2}$	0
B	0	$\frac{1}{2}$

Table 22: Correlated strategies - probability of all pure strategy profiles

An important fact to note is that a correlated strategy may not be obtained from a mixed strategy. For instance, consider the correlated strategy in Table 22. If Player 1 and Player 2 play mixed strategies that generates the same distribution over strategy profile as in Table 22, then either 1 must put zero weight on A or 2 must put zero weight on b. This implies that we cannot get the distribution in Table 22.

In general, the correlated strategy $\rho \in \Delta\Big(\prod_{i\in N} S_i\Big)$ and a mixed strategy $\sigma \in \prod_{i\in N} \Delta S_i$. Every mixed strategy generates a correlated strategy. Hence, the set of distributions over strategy profiles that can be obtained by correlated strategy is larger than the set of distributions generated by mixed strategies. Player i evaluates a correlated strategy ρ using expected utility:

$$U_i(\rho) = \sum_{s\in S} u_i(s)\rho(s).$$

9.2 Formal Definition

We will now define a correlated equilibrium based on the notion of correlated strategies. Let $\Gamma \equiv (N, \{S_i\}_{i\in N}, \{u_i\}_{i\in N})$ be a finite strategic form game. To avoid confusion, we will refer to strategies in S_i for each i as **actions** of Player i.

For every probability vector (correlated strategy) p over $S \equiv S_1 \times \ldots \times S_n$, an **extended game of** Γ is defined as:

- An outside observer chooses a profile of pure actions $s \in S$ using the correlated strategy p.
- It reveals to each player i, his recommendation s_i but not s_{-i}.
- Each player i chooses an action $s_i' \in S_i$ after receiving his recommendation.

We denote this extended game as $\Gamma(p)$. Hence, formally a strategy in this extended game is a different object compared to the strategy in a strategic form game.

Definition 9 *A strategy of Player i in the extended game $\Gamma(p)$ is a map $\psi_i : S_i \to S_i$, i.e., giving an action for every possible recommended action.*

One strategy is the **obedient** strategy map - for every $s_i \in S_i$, $\psi_i^*(s_i) = s_i$ for each i. Below, we show the mathematical implication of the fact that ψ^* is a Nash equilibrium of $\Gamma(p)$. What does it mean to say that ψ^* is a Nash equilibrium of $\Gamma(p)$? It means that given that everyone else is playing the strategy ψ^*, payoff of an agent i is maximized by playing ψ^*. Since there is uncertainty about the recommendation of other players, payoff of agent i has to be computed by taking expectation over all possible recommendations.

THEOREM **9** *The strategy profile ψ^* is a Nash equilibrium of $\Gamma(p)$ if and only if for every $i \in N$, for every $s_i, s_i' \in S_i$, we have*

$$\sum_{s_{-i} \in S_{-i}} p(s_i, s_{-i}) u_i(s_i, s_{-i}) \geq \sum_{s_{-i} \in S_{-i}} p(s_i, s_{-i}) u_i(s_i', s_{-i})$$

Proof: The strategy profile ψ^* is a Nash equilibrium if and only if no player i can unilaterally deviate from his recommended action. If Player i receives recommendation s_i, then his conditional belief that other players received recommendation s_{-i} is

$$\frac{p(s_i, s_{-i})}{\sum_{t_{-i}} p(s_i, t_{-i})},$$

where the denominator is positive from the fact that $p(s_i, s_{-i}) > 0$. Then, his expected payoff from following $\psi_i^*(s_i) = s_i$ (given others are following recommendation) is

$$\sum_{s_{-i} \in S_{-i}} \frac{p(s_i, s_{-i})}{\sum_{t_{-i}} p(s_i, t_{-i})} u_i(s_i, s_{-i}).$$

His expected payoff from playing s_i' (given others are following recommendation) is

$$\sum_{s_{-i} \in S_{-i}} \frac{p(s_i, s_{-i})}{\sum_{t_{-i}} p(s_i, t_{-i})} u_i(s_i', s_{-i}).$$

Since the denominator is positive, we can say that s_i is best response if and only if

$$\sum_{s_{-i} \in S_{-i}} p(s_i, s_{-i}) u_i(s_i, s_{-i}) \geq \sum_{s_{-i} \in S_{-i}} p(s_i, s_{-i}) u_i(s_i', s_{-i}).$$

■

This leads to the definition of a correlated equilibrium.

DEFINITION **10** *A correlated strategy p over S is a* **correlated equilibrium** *if for every $i \in N$, for every $s_i, s_i' \in S_i$,*

$$\sum_{s_{-i} \in S_{-i}} p(s_i, s_{-i}) u_i(s_i, s_{-i}) \geq \sum_{s_{-i} \in S_{-i}} p(s_i, s_{-i}) u_i(s_i', s_{-i}).$$

In other words, a correlated strategy p over S is a **correlated equilibrium** *if the strategy profile ψ^* is a Nash equilibrium of the extended game $\Gamma(p)$.*

This shows that the set of correlated equilibria are solutions to a finite set of inequalities in a finite game. As result, they form a convex and compact set (in particular, a *polytope*, defined by a system of linear inequalities).

Every Nash equilibrium σ^* of Γ induces a probability distribution p_{σ^*}, where for every $(s_1, \ldots, s_n)$,

$$p_{\sigma^*}(s_1, \ldots, s_n) = \sigma_1^*(s_1) \times \ldots \times \sigma_n^*(s_n).$$

Below, we formally show that every Nash equilibrium induces a distribution over strategy profiles that is a correlated equilibrium.

THEOREM **10** *For every Nash equilibrium σ^* of Γ, the induced correlated strategy p_{σ^*} is a correlated equilibrium of $\Gamma(p_{\sigma^*})$.*

Proof: Note that $p_{\sigma^*}(s) > 0$ if and only if for every $i \in N$, s_i is in the support of σ^*. Pick agent i, $s_i, s_i' \in S_i$. We see that

$$\sum_{s_{-i} \in S_{-i}} p_{\sigma^*}(s_i, s_{-i}) u_i(s_i, s_{-i}) = \sum_{s_{-i} \in S_{-i}} \sigma_1^*(s_1) \times \ldots \times \sigma_n^*(s_n) u_i(s_i, s_{-i}) = \sigma_i^*(s_i) U_i(s_i, \sigma_{-i}^*).$$

Further,

$$\sum_{s_{-i} \in S_{-i}} p_{\sigma^*}(s_i, s_{-i}) u_i(s_i', s_{-i}) = \sum_{s_{-i} \in S_{-i}} \sigma_1^*(s_1) \times \ldots \times \sigma_n^*(s_n) u_i(s_i', s_{-i}) = \sigma_i^*(s_i) U_i(s_i', \sigma_{-i}^*).$$

Since s_i is in the support of Nash equilibrium at σ^*, it implies that $\sigma_i^*(s_i) > 0$. Further, by the indifference lemma, s_i is a best response to σ_{-i}^*, and hence,

$$U_i(s_i, \sigma_{-i}^*) \geq U_i(s_i', \sigma_{-i}^*).$$

This gives us that

$$\sum_{s_{-i} \in S_{-i}} p_{\sigma^*}(s_i, s_{-i}) u_i(s_i, s_{-i}) \geq \sum_{s_{-i} \in S_{-i}} p_{\sigma^*}(s_i, s_{-i}) u_i(s_i', s_{-i}),$$

as required. ∎

10 A Foundation for Iterated Elimination

In this section, we discuss a foundation for eliminating strictly dominated strategies in a finite game. The foundation is inspired by the idea of correlated strategies discussed earlier - it extends the idea of correlated strategies to beliefs.

To fix ideas, we are given a finite strategic form game: $\Gamma := (N, \{S_i\}_{i \in N}, \{u_i\}_{i \in N})$. We are going to consider the mixed extension of this game. But we will only be concerned with eliminating pure strategies from this mixed extension. To remind, Theorem 4 has told us that if a pure strategy s_i is strictly dominated for Player i, every mixed strategy σ_i with s_i in its support is also strictly dominated. Hence, eliminating a pure strategy also eliminates *some* strictly dominated mixed strategies. However, as we have seen earlier, it may not eliminate *all* strictly dominated mixed strategies. The foundation we provide here is about iterated elimination of strictly dominated *pure* strategies.

We remind ourselves what we mean by iterated elimination of strictly dominated (pure) strategies. First a pure strategy s_i is **strictly dominated** for Player i in $\Delta\Gamma$ if there exists $\sigma_i \in \Delta S_i$ such that

$$U_i(\sigma_i, \sigma_{-i}) > U_i(s_i, \sigma_{-i}) \ \forall \ \sigma_{-i} \in \Delta S_{-i}.$$

The following lemma suggests a weakening of this definition.

Lemma **7** *A strategy s_i is strictly dominated for Player i if there exists $\sigma_i \in \Delta S_i$ such that*

$$U_i(\sigma_i, s_{-i}) > U_i(s_i, s_{-i}) \ \forall \ s_{-i} \in S_{-i}. \tag{1}$$

Proof: Suppose there exists $\sigma_i \in \Delta S_i$ such that Inequality 1 holds. Consider σ_{-i} and note that

$$\begin{aligned} U_i(\sigma_i, \sigma_{-i}) &= \sum_{s_{-i}} U_i(\sigma_i, s_{-i}) \sigma_{-i}(s_{-i}) \\ &> \sum_{s_{-i}} U_i(s_i, s_{-i}) \sigma_{-i}(s_{-i}) \qquad \text{(By Inequality 1)} \\ &= U_i(s_i, \sigma_{-i}). \end{aligned}$$

Hence, s_i is strictly dominated. ■

By Lemma 7, we will now say that a strategy s_i is strictly dominated for Player i if there exists $\sigma_i \in \Delta S_i$ such that

$$U_i(\sigma_i, s_{-i}) > U_i(s_i, s_{-i}) \ \forall \ s_{-i} \in S_{-i}.$$

10.1 Correlated Beliefs

Just like correlated strategies allow for probability distribution over strategy profiles, a general system of belief for Player i must allow a probability distribution over S_{-i} - it specifies a probability of each of the strategy profile s_{-i} being played. Such probabilities need not be computed using independence of strategies of other players. So, belief of player i is a map $\mu_i : S_{-i} \to [0,1]$, with $\sum_{s_{-i}} \mu_i(s_{-i}) = 1$. Note that a mixed strategy profile σ induces a belief for every player i: $\mu_i(s_{-i}) := \times_{j \neq i} \sigma_j(s_j)$ for all s_{-i}. These beliefs are generated by independent probabilities of each player $j \neq i$. In general, beliefs may allow correlations.

A strategy $s_i \in S_i$ is a **best response with respect to** a belief μ_i if

$$\sum_{s_{-i}} U_i(s_i, s_{-i})\mu_i(s_{-i}) \geq \sum_{s_{-i}} U_i(\sigma_i, s_{-i})\mu_i(s_{-i}) \ \forall \ \sigma_i \in \Delta S_i. \tag{2}$$

Note here that we check for mixed strategies σ_i for best response. But this is not necessary as the following lemma suggests.

Lemma **8** *Suppose strategy s_i satisfies the following inequalities with respect to a belief μ_i:*

$$\sum_{s_{-i}} u_i(s_i, s_{-i})\mu_i(s_{-i}) \geq \sum_{s_{-i}} u_i(\sigma_i, s_{-i})\mu_i(s_{-i}) \ \forall \ \sigma_i \in S_i. \tag{3}$$

Then, s_i is a best response with respect to belief μ_i.

Proof: Suppose Inequality 3 holds. Then, for every σ_i, we can write

$$\begin{aligned}
\sum_{s_{-i}} U_i(\sigma_i, s_{-i})\mu_i(s_{-i}) &= \sum_{s_{-i}} \sum_{s'_i \in S_i} u_i(s'_i, s_{-i})\sigma_i(s'_i)\mu_i(s_{-i}) \\
&= \sum_{s'_i \in S_i} \sigma_i(s'_i) \sum_{s_{-i}} u_i(s'_i, s_{-i})\mu_i(s_{-i}) \\
&\leq \sum_{s'_i \in S_i} \sigma_i(s'_i) \sum_{s_{-i}} u_i(s_i, s_{-i})\mu_i(s_{-i}) \qquad \text{(By Inequality 3)} \\
&= \sum_{s_{-i}} u_i(s_i, s_{-i})\mu_i(s_{-i}) \qquad \text{(Using } \sum_{s'_i} \sigma_i(s'_i) = 1\text{)} \\
&= \sum_{s_{-i}} U_i(s_i, s_{-i})\mu_i(s_{-i}).
\end{aligned}$$

So, s_i is a best response with respect to belief μ_i. ■

We define the notion of never best response using correlated beliefs now.

Definition 11 *A strategy $s_i \in S_i$ is a* **never-best response** *in the mixed extension of the strategic form game $(N, \{S_i\}_{i\in N}, \{u_i\}_{i\in N})$ if it is not a best response with respect to any belief μ_i.*

We prove the equivalence of never-best response strategies and strictly dominated strategies. To show the subtle nature of the result, consider couple of examples first. Consider the two player game in Table 23 - the table only shows payoff of Player 1. Strategy C is **not** strictly dominated. We will show there are **beliefs** of Player 1 on the strategies of Player 2 for which he should play C. It is clear that if the beliefs put entire probability on a or b, C cannot be a best response. But if he puts $\frac{1}{2}$ probability on a and $\frac{1}{2}$ probability on b, then C is a best response.

	a	b
A	1	0
B	0	1
C	0.6	0.6

Table 23: Correlated beliefs example

To make things more interesting, consider a finite game with 3 players. Player 1 has three strategies: $\{A, B, C\}$, Player 2 has two strategies $\{a, b\}$, and Player 3 has two strategies $\{a', b'\}$. We only show the payoff of Player 1 in Table 24.

	(a, a')	(b, a')	(a, b')	(b, b')
A	4	2	2	1
B	1	2	2	4
C	3	0	0	3

Table 24: Correlated beliefs example

Now, a belief of Player 1 is a function μ_1 which assigns the following non-negative numbers adding to 1:

$$\mu_1(a, a'),\ \mu_1(b, a'),\ \mu_1(a, b'),\ \mu_1(b, b').$$

We see that strategy C is **not** strictly dominated for Player 1 - if $\alpha A + (1 - \alpha)B$ strictly dominates C, then when other players play (a, a'), we must have $4\alpha + 1 - \alpha > 3$ and when others play (b, b') we must have $\alpha + 4(1 - \alpha) > 3$. Adding these two inequalities gives us $5 > 6$, which is a contradiction.

Now, suppose we only consider independent beliefs: so, Player 1 believes that Player 2 plays a with probability p_2 and b with probability $(1-p_2)$; Player 3 plays a' with probability p_3 and b' with probability $(1-p_3)$. This results in the following beliefs:

$$\mu_1(a,a') = p_2p_3,\ \mu_1(b,a') = (1-p_2)p_3,\ \mu_1(a,b') = p_2(1-p_3),\ \mu_1(b,b') = (1-p_2)(1-p_3).$$

The payoffs of Player 1 from this belief is given below:

$$\begin{aligned}
\mathcal{U}_1(A,\mu_1) &= 4p_2p_3 + 2\Big[(1-p_2)p_3 + p_2(1-p_3)\Big] + (1-p_2)(1-p_3) \\
\mathcal{U}_1(B,\mu_1) &= p_2p_3 + 2\Big[(1-p_2)p_3 + p_2(1-p_3)\Big] + 4(1-p_2)(1-p_3) \\
\mathcal{U}_1(C,\mu_1) &= 3\Big[p_2p_3 + (1-p_2)(1-p_3)\Big]
\end{aligned}$$

The difference in expected payoff are

$$\begin{aligned}
\mathcal{U}_1(A,\mu_1) - \mathcal{U}_1(C,\mu_1) &= 1 + p_2 + p_3 - 5p_2p_3 \\
\mathcal{U}_1(B,\mu_1) - \mathcal{U}_1(C,\mu_1) &= 4(p_2+p_3) - 2 - 5p_2p_3
\end{aligned}$$

We see that utility from A is better than B if and only if

$$\Big[\big[\mathcal{U}_1(A,\mu_1) - \mathcal{U}_1(C,\mu_1)\big] - \big[\mathcal{U}_1(B,\mu_1) - \mathcal{U}_1(C,\mu_1)\big] \geq 0\Big] \Leftrightarrow \Big[p_2 + p_3 \leq 1\Big].$$

But if $p_2 + p_3 \leq 1$, then $1 \geq p_2 + p_3 \geq 2\sqrt{p_2p_3}$. Hence, $p_2p_3 \leq \frac{1}{4}$. Using this, we get

$$\begin{aligned}
\mathcal{U}_1(A,\mu_1) - \mathcal{U}_1(C,\mu_1) &= 1 + p_2 + p_3 - 5p_2p_3 \\
&= 1 + p_2(1-p_3) + p_3(1-p_2) - 3p_2p_3 \\
&\geq 1 - 3p_2p_3 \\
&\geq \frac{1}{4}.
\end{aligned}$$

A similar argument shows that if $p_2 + p_3 \geq 1$, then $\mathcal{U}_1(B,\mu_1) - \mathcal{U}_1(C,\mu_1) > 0$. Hence, independent beliefs imply that C cannot be a best response to such beliefs.

However, consider the following correlated beliefs.

$$\mu_1(a,a') = \frac{1}{2},\ \mu_1(b,a') = 0,\ \mu_1(a,b') = 0,\ \mu_1(b,b') = \frac{1}{2}.$$

Payoffs of Player 1 are now: $\mathcal{U}_1(A,\mu_1) = \mathcal{U}_1(B,\mu_1) = 2.5, \mathcal{U}_1(C,\mu_1) = 3$. Hence, C is a best response. Notice that this correlated belief cannot be generated using independent beliefs.

This idea extends generally. If a pure strategy is not strictly dominated, then it is best response to some *correlated* belief.

THEOREM **11** *A pure strategy of a player in a strategic form game is a never-best response if and only if it is strictly dominated.*

Proof: Clearly, every strictly dominated strategy is a never-best response strategy. For the other direction, fix a player j in a strategic form game $\Gamma \equiv (N, \{S_i\}_{i\in N}, \{u_i\}_{i\in N})$ and a strategy $\bar{s}_j \in S_j$. Consider a new game in which there are just two players j and $-j$. The set of strategies available to Player j is $S'_j := S_j \setminus \{\bar{s}_j\}$ and to Player $-j$ is S_{-j} (i.e., every strategy profile of players in $N \setminus \{j\}$ is interpreted as a strategy of Player $(-j)$). The utility of Player j at strategy profile (s_j, s_{-j}) is:

$$v_j(s_j, s_{-j}) = u_j(s_j, s_{-j}) - u_j(\bar{s}_j s_{-j}).$$

The payoff to Player $-j$ is negative of payoff to Player j - hence, it is a zero-sum game. Denote this game as Γ' and consider its mixed extension $\Delta\Gamma'$. We will abuse notation and denote the payoff from a mixed strategy profile (σ_j, σ_{-j}) to Player j as $v_j(\sigma_j, \sigma_{-j})$. Let $(\sigma_j^*, \sigma_{-j}^*)$ be a Nash equilibrium of this game.

Now, since Γ' is a two-person zero-sum game. Hence, a mixed strategy of Player $-j$ corresponds to a *correlated belief* of Player j in the original game. Hence, strategy $\bar{s}_j$ is a never-best response implies for every mixed strategy σ_{-j} of Player $-j$, there exists a strategy σ_j such that $U_j(\sigma_j, \sigma_{-j}) - U_j(\bar{s}_j, \sigma_{-j}) > 0$, or, $v_j(\sigma_j, \sigma_{-j}) > 0$. But this implies that the Nash equilibrium payoff of Player j is positive: $v_j(\sigma_j^*, \sigma_{-j}^*) > 0$. By using Theorem 8, we conclude that

$$v_{-j}(\sigma_j^*, \sigma_{-j}^*) = -v_j(\sigma_j^*, \sigma_{-j}^*) < 0.$$

Since $(\sigma_j^*, \sigma_{-j}^*)$ is a Nash equilibrium, we get that for every strategy σ_{-j} of Player $(-j)$, we have

$$-v_j(\sigma_j^*, \sigma_{-j}) \leq -v_j(\sigma_j^*, \sigma_{-j}^*) < 0,$$

which gives

$$u_j(\sigma_j, \sigma_{-j}) > u_j(\bar{s}_j, \sigma_{-j}) \ \forall \ \sigma_{-j}.$$

Hence, $\bar{s}_j$ is strictly dominated for Player j. ■

Remember that this equivalence is only valid if we allow for correlated beliefs - of course, for two-player games these correlated belief is same as independent belief.

10.2 Correlated Rationalizability

The fact that pure strategy Nash equilibrium does not exist makes it problematic as a solution concept some times. Mixed strategies are not entirely convincing since players play pure strategies at the end of the game anyway. The notion of correlated rationalizability is developed as a "set theoretic" pure strategy Nash equilibrium. In stead of predicting a unique strategy to be played by each player, we will say that a player *may* play any strategy from a set as long as it is best response with respect to *some* belief over the strategy sets chosen by other players.

DEFINITION **12** *A profile of set of strategies* $(Z_1, \ldots, Z_n)$ *is* **rationalizable** *in the strategic form game* $(N, \{S_i\}_{i \in N}, \{u_i\}_{i \in N})$ *if for every* $i \in N$ *and every* $s_i \in Z_i$ *there is a belief* μ_i *whose support is a subset of* $\times_{j \neq i} Z_j$ *such that*

$$\mathcal{U}_i(s_i, \mu_i) \geq \mathcal{U}_i(s'_i, \mu_i) \qquad \forall\ s'_i \in S_i,$$

i.e., s_i *is a best response with respect to belief* μ_i*.*

Note that the strategies in Z_j for each j are only used to form beliefs - strategy profiles involving strategies outside them get zero probability. The best response is with respect to all the strategies.

Also, note that if a profile of set of strategies $(Z_1, \ldots, Z_n)$ is rationalizable and another profile of set of strategies $(Z'_1, \ldots, Z'_n)$ is rationalizable then the profile of set of strategies $(Z_1 \cup Z'_1, \ldots, Z_n \cup Z'_n)$ is also rationalizable. Hence, the set of rationalizable strategies is the largest collection of $\{Z_j\}_j$ that can be rationalized.

Consider the example in Table 25. $(\{A\}, \{a\})$ is not a set of rationalizable strategies. This is because here there is only one degenerate belief: Player 1 must believe Player 2 plays a and Player 2 must believe that Player 1 plays A. But a is not a best response if Player 1 plays A. On the other hand, $(\{A, C\}, \{a, b\})$ is a set of rationalizable strategies. How do we verify this? A is a best response if a is played and C is a best response if b is played. Similarly, for Player 2, a is a best response if C is played and b is a best response if A is played.

	a	b	c
A	$(6, 2)$	$(0, 6)$	$(4, 4)$
B	$(2, 12)$	$(4, 3)$	$(2, 5)$
C	$(0, 6)$	$(10, 0)$	$(2, 2)$

Table 25: Two Player Game

The idea here is that we observe various (pure) strategies being played by each player. When can we say that these strategies being played are best response to some belief of players over the strategies played by other players? This is the rationalizability question.

An immediate claim is the following.

LEMMA **9** *Every strategy in the support of a Nash equilibrium is rationalizable.*

Proof: Suppose s_i is a strategy of Player i in the support Nash equilibrium σ^*. Now for every j, Z_j are all the strategies in the support of the Nash equilibrium σ^* and the belief μ_j is the product

$$\times_{k \neq j} \sigma^*_k(s_k) \ \forall \ s_{-j}.$$

By the definition of Nash equilibrium and the indifference lemma, each s_j in the support of σ^*_j is a best response of j with respect to the belief μ_j. ■

One can also show that strategies used with positive probability in a correlated equilibrium are also rationalizable - this follows directly from the definition of correlated equilibrium. In general, finding the set of rationlizable strategies can be quite cumbersome. Below, we provide an easy method with the help of a cute result.

Couple of quick observations are worth making. First, if a strategy is strictly dominated, then it cannot be rationalizable. But we can say more. We now remind ourselves the definition of the iterated elimination procedure, but restricting attention to pure strictly dominated strategy elimination in the mixed extension of Γ.

DEFINITION **13** *The profile of set of strategies* $(X_1, \ldots, X_n)$ *survives* **iterated elimination of strictly dominated pure strategies** *if* $X \equiv \times_{j \in N} X_j$ *and there is a collection* $(\{X^t_j\}_{j \in N})^T_0$ *of sets that satisfy for each* $j \in N$ *the following:*

- $X^0_j = S_j$ *and* $X^T_j = X_j$,
- $X^{t+1}_j \subseteq X^t_j$ *for each* $t < T$,
- *for each* $t < T$, *every strategy in* $X^t_j \setminus X^{t+1}_j$ *is strictly dominated in the mixed extension of the game* $\Gamma^t \equiv (N, \{X^t_i\}_i, \{u^t_i\}_i)$, *where* u^t_i *is the restriction of* u_i *to strategy profiles in this game.*
- *No strategy in* X^T_j *is strictly dominated.*

The theorem says that rationalizability and iterated elimination of strictly dominated pure strategies are equivalent.

THEOREM **12** *The set of rationalizable strategies of a player is the set of strategies available after iterated elimination of strictly dominated strategies.*

Proof: Let $(Z_1, \ldots, Z_n)$ be the *largest* set of rationalizable strategies for each player. We will argue that Z_i survives iterated elimination of strictly dominated strategies for each $i \in N$. Suppose not. Then, consider the first stage t where strategy $s_i \in Z_i$ of some Player i gets eliminated in iterated elimination procedure. Since this is the first period where such a strategy is getting eliminated, all the strategies Z_j of $j \neq i$ still exists in the game Γ^t in period t. Hence, s_i is strictly dominated in Γ^t implies (by Theorem 11) that s_i is a never-best-response strategy for this game. Since s_i is rationalizable, there is a belief μ_i over $\times_{j\neq i} Z_j$ such that s_i is a best response with respect to μ_i. This is a contradiction to s_i being a never-best-response in Γ^t.

Now, we turn to the other direction, where we intend to show that for each Player i, Z_i is the only set of strategies that survives iterated elimination procedure. For this, pick $s_i \in X_i$, where X_i is the set of strategies that survive iterated elimination procedure (remember, we want to show $Z_i = X_i$). By definition every strategy in X_i is not strictly dominated in the game Γ^T with strategy sets X_i. So, by Theorem 11, every strategy in X_i is a best response among strategies in X_i to some belief μ_i over X_{-i}. So,

$$\mathcal{U}_i(s_i, \mu_i) \geq \mathcal{U}_i(s'_i, \mu_i) \qquad \forall\ s'_i \in X_i = X_i^T.$$

We will show that for all $t \in \{0, \ldots, T\}$,

$$\mathcal{U}_i(s_i, \mu_i) \geq \mathcal{U}_i(s'_i, \mu_i) \qquad \forall\ s'_i \in X_i^t.$$

Suppose this is not true. Then, there is some period t where

$$\mathcal{U}_i(s_i, \mu_i) \geq \mathcal{U}_i(s'_i, \mu_i) \qquad \forall\ s'_i \in X_i^{t+1}. \tag{4}$$

but

$$\mathcal{U}_i(s_i, \mu_i) < \max_{s'_i \in X_i^t} \mathcal{U}_i(s'_i, \mu_i).$$

Let $\hat{s}_i$ be a such that

$$\mathcal{U}_i(s_i, \mu_i) < \mathcal{U}_i(\hat{s}_i, \mu_i) = \max_{s'_i \in X_i^t} \mathcal{U}_i(s'_i, \mu_i).$$

Since strategies $X_j \subseteq X_j^t$ for all $j \in N$ and μ_i has support over $\times_{j\neq i} X_j$, $\hat{s}_i$ is a best response of Player i with respect to belief μ_i in Γ^t. By Theorem 11, $\hat{s}_i$ is not strictly dominated in Γ^t. So, $\hat{s}_i$ is not strictly dominated. Hence, $\hat{s}_i \in X_i^{t+1}$. But this contradicts Inequality 4.

This implies that $\mathcal{U}_i(s_i, \mu_i) \geq \mathcal{U}_i(s'_i, \mu_i)$ for all $s'_i \in X_i^0 = S_i$. Hence, s_i is a best response with respect to belief μ_i (with support $\times_{j\neq i} X_j$) in Γ. Hence, the collection of sets of strategies $(X_1, \ldots, X_n)$ is rationalizable. ■

Since the procedure we defined for iterated elimination did not specify any order of elimination, this also implies that order of elimination of strictly dominated strategies does not matter.

11 Bayesian Games

Often, the strategic form game depends on some external factor. These factors may be known to some agents with varying certainty. To make ideas clear, consider a situation in which two agents are deciding where to meet. Each agent privately observes the weather in his city but does not know the weather of the other agent's city. Based on the weather in the city, an agent has a set of *actions* available to him, and his utility will depend on the weather in both the cities and the actions chosen by both the agents. Here, the weather in each city is a *signal* that is privately observed by the player. The signal determines the action set of the strategic game. The utility in the strategic form game is determined by the signals realized by all the agents and the actions taken.

The kind of uncertainty in this example is about the weather in the cities. Each agent uses a *common prior* to evaluate uncertainty using expected utility. In this example, there is a probability distribution about the weather in both the cities. Note that since an agent only observes weather in his own city, he can use Bayes rule to update the conditional probabilities.

Note that the strategy of a player and his payoff functions are complicated objects in this environment because (a) it depends on the signals players receive and (b) there is uncertainty about the signals of other players. Harasanyi was the first to formally define an analogue of a strategic game in this uncertain environment.

Definition **14** *A* **Bayesian game (game of incomplete information)** *is defined by*

- N*: a finite set of players,*
- T_i*: set of* **types** *(signals) for each player* i*, and* $T = \times_{i \in N} T_i$ *is the set of type vectors,*
- p*: a common probability distribution (***belief or prior***) over* T *with the restriction that* $\pi_i(t_i) := \sum_{t_{-i} \in T_{-i}} p(t_i, t_{-i}) > 0$ *for each* $t_i \in T_i$ *and for each* $i \in N$*,*
- $A_i(t_i)$*: the set of actions available to each Player* i *with type* t_i*,*
- $u_i(t, a)$*: the payoff assigned by each Player* i *at type profile* $t \equiv (t_1, \ldots, t_n) \in T$ *when action profile* $a \equiv (a_1, \ldots, a_n)$*, where each* $a_j \in A_j(t_j)$ *for all* $j \in N$*, is played.*

A Bayesian game proceeds in a sequence where some of the associated uncertainties are resolved.

- The type vector $t \in T$ is chosen (by nature) using the probability distribution p.
- Each player $i \in N$ observes his own type t_i but does not know the types of other agents.
- After observing their types, each player i plays an action $a_i \in A_i(t_i)$.
- Each player i receives an utility equal to $u_i(t, a)$ when the type profile realized is $t \equiv (t_1, \ldots, t_n)$ and the action profile is $a \equiv (a_1, \ldots, a_n)$.

In most of the examples, we will make the assumption that for all t_{-i}, t'_{-i},

$$u_i((t_i, t_{-i}), a) = u_i((t_i, t'_{-i}), a) \text{ for all } a, \text{ for all } t_i, \text{ for all } i \in N$$

This is called a **private values** model. It rules out the possibility that a player's utility depends directly on the type of other players. Notice that the action chosen by a Player may depend on his type in the game, and hence, indirectly, Player i's utility will depend on the type of other players (though the actions chosen by other players).

Because of uncertainty, the players do not even know the action set available to other players. So, they do not know which strategic form game is being played. Note that the action set depends on the type of the player. Further, the utility depends on the type vector realized and the actions taken by all the players.

Strategies in such games are complicated objects. To remind, a strategy must describe the action to be taken for every possible contingency. Hence, here also, a strategy must describe what action to take for every signal/type that the player receives.

A **strategy** of Player i in a Bayesian game is a map $s_i : T_i \to \cup_{t_i \in T_i} A_i(t_i)$ such that $s_i(t_i) \in A_i(t_i)$ for all $t_i \in T_i$. Thus, a strategy prescribes one action for every type.

What is the payoff of Player i from a strategy profile $s \equiv (s_1, \ldots, s_n)$? There are two ways to think about it: ex-ante payoff, which is computed before realization of the type, and interim payoff, which is computed after realization of the type. Ex-ante payoff from strategy profile s is

$$U_i(s) := \sum_{t \in T} p(t) u_i(t, (s_1(t_1), \ldots, s_n(t_n))).$$

Here, if type profile t is realized, then action profile $(s_1(t_1), \ldots, s_n(t_n))$ is played according to the strategy profile s. Hence, the payoff realized by Player i at type profile t is just $u_i(t, (s_1(t_1), \ldots, s_n(t_n)))$. Then, $U_i(s)$ computed using expectation from this.

The interim payoffs are computed by updating beliefs after realizing the types. In particular, once Player i knows his type to be $t_i \in T_i$, he computes his conditional probabilities as follows. For every $t_{-i} \in T_{-i}$,

$$p_i(t_{-i}|t_i) := \frac{p(t_i, t_{-i})}{\sum_{t'_{-i} \in T_{-i}} p(t_i, t'_{-i})} = \frac{p(t_i, t_{-i})}{\pi_i(t_i)},$$

where we will denote $\pi_i(t_i) \equiv \sum_{t'_{-i} \in T_{-i}} p(t_i, t'_{-i})$ and note that it is positive by our assumption. The interim payoff of Player i with type t_i from a strategy profile s_{-i} of other players and when he takes action $a_i \in A_i(t_i)$ is thus

$$U_i((a_i, s_{-i})|t_i) := \sum_{t'_{-i} \in T_{-i}} p_i(t'_{-i}|t_i) u_i(t, (a_i, s_{-i}(t'_{-i}))).$$

If the beliefs are independent, then observing own type gives no extra information to the players. Hence, no updating of prior belief is required by the players.

An easy consequence of this definition is the following. Consider Player i and a strategy profile (s_i, s_{-i})

$$\begin{aligned}\sum_{t_i \in T_i} U_i((s_i(t_i), s_{-i})|t_i)\pi_i(t_i) &= \sum_{t_i \in T_i} \pi_i(t_i) \sum_{t_{-i} \in T_{-i}} p_i(t_{-i}|t_i) u_i(t, (s_i(t_i), s_{-i}(t_{-i}))) \\ &= \sum_{t \in T} p(t) u_i(t, (s_i(t_i), s_{-i}(t_{-i}))) = U_i(s). \qquad (5)\end{aligned}$$

Note: The above expressions are for finite type spaces, but similar expressions (using integrals) can also be written with infinite type spaces. Below is an application with infinite type spaces.

11.1 Two examples

We give two examples. One corresponding to a finite type space and the other for a non-finite type space.

Toy example. Suppose there are two players $N := \{1, 2\}$. Type space of Player 1 is $T_1 = \{x_1, t_1\}$ and of Player 2 is $T_2 = \{x_2, t_2, w_2\}$. So, the set of type profiles is the following set

$$T := \{(x_1, x_2), (x_2, t_2), (x_1, w_2), (t_1, x_2), (t_1, t_2), (t_1, w_2)\}.$$

Now, the **common prior** over T is a probability distribution p. For instance, p can be a *uniform* distribution over T, implying that it assigns equal ($\frac{1}{6}$) probability to each type profile. From this, we can compute the conditional probabilities of each player. We show it for Player 1.

For Player 1, his $\pi_1(x_1) = p(x_1, x_2) + p(x_1, t_2) + p(x_1, w_2) = \frac{1}{2}$. Similarly, $\pi_1(t_1) = \frac{1}{2}$. Hence, his conditional probabilities are

$$p_1(x_2|x_1) = \frac{p(x_2, x_1)}{\pi_1(x_1)} = \frac{1}{3}.$$

Similarly, we can compute the other conditional probabilities as

$$p_1(t_2|x_1) = p_1(w_2|x_1) = \frac{1}{3}$$

and

$$p_1(x_2|t_1) = p_1(t_2|t_1) = p_1(w_2|t_1) = \frac{1}{3}.$$

Suppose $A_1(x_1) = A_1(t_1) = \{a_1, b_1\}$ and $A_2(x_2) = A_2(t_2) = A_2(w_2) = \{a_2, b_2, c_2\}$. What are the strategies of players in this game? For Player 1, we list **all** the strategies it can have:

$$s_1(x_1) = s_1(t_1) = a_1; s'_1(x_1) = s'_1(t_1) = b_1; s''(x_1) = a_1, s''(t_1) = b_1; s'''(x_1) = b_1, s'''(t_1) = a_1.$$

So, the set of strategies of Player 1 can be any of the four strategies listed above. Similarly, Player 2 can have 27 possible strategies. We list two of them below for illustration:

$$s_2(x_2) = a_2, s_2(t_2) = b_2, s_2(w_2) = c_2; s'_2(x_2) = s'_2(t_2) = a_2, s'_2(w_2) = b_2.$$

How are the utilities evaluated? Consider a strategy profile (s_1, s_2). The ex-ante payoff of Player 1 from (s_1, s_2) is computed as follows:

$$U_1(s_1, s_2) = \sum_{(y_1,y_2)\in T} p(y_1, y_2) u_1((y_1, y_2), (s_1(y_1), s_2(y_2))).$$

On the other hand, the interim payoff of Player 1 with type x_1 and action $s_1(x_1)$ is computed as follows:

$$\begin{aligned} U_1((s_1(x_1), s_2)|x_1) &= p_1(x_2|x_1)u_1((x_1, x_2), (s_1(x_1), s_2(x_2))) \\ &\quad + p_1(t_2|x_1)u_1((x_1, t_2), (s_1(x_1), s_2(t_2))) \\ &\quad + p_1(w_2|x_1)u_1((x_1, w_2), (s_1(x_1), s_2(w_2))). \end{aligned}$$

Using these definitions, we will later define the notion of a Bayesian equilibrium.

First-price auction. We give an informal description of a Bayesian game before describing the equilibrium concepts. This Bayesian game is in the context of an auction - a popular subfield of economic theory where game theory has been applied successfully in practice and theory. There is an indivisible object for sale to a set of buyers (players). Each buyer has a value v_i for the object. The value is the type of the object, and hence, every buyer only knows his own value but not the value of others. The values are drawn using a distribution p over the set of all value profiles.

The set of actions available to a player in this game is the set of all non-negative real numbers. Such actions are called **bids** in auction literature. A bid specifies the amount a buyer is willing to pay. The buyer with the highest bid (ties broken with equal probability) wins the object. If a buyer i with value v_i wins the object by bidding b_i, then his utility is $v_i - b_i$ times the probability of winning. A losing buyer gets zero utility. Note that the amount a biddder bids may depend on his type. Whether a buyer wins or not depends on the bids of all the players. The utility of a player depends on this probability of winning and his own type.

To be a little more specific, let us study strategies which are commonly referred to as **symmetric monotone bidding strategies**. Assume that type space $T_i = \mathbb{R}_+$. A symmetric monotone strategy is a map $b : T_i \to \mathbb{R}_+$. Note that every bidder is using the same bid function (strategy). We further assume that b is strictly increasing and differentiable. Suppose each bidder draws his type independently from T_i using a distribution F (identical distribution for all bidders). If types of all the buyers are $(v_1, \ldots, v_n)$, the the probability of this type vector is $F(v_1) \times \ldots \times F(v_n)$. Since bids are monotone functions, a bidder with type v_i wins when everyone follows this strategy if $v_i > \max_{j \neq i} v_j$. The probability of this event is $F(v_i)^{n-1}$. The interim payoff of Player i with type v_i from this strategy is

$$F(v_i)^{n-1}(v_i - b_i(v_i)).$$

Hence, the ex-ante payoff from of Player i with type v_i from this strategy is

$$\int_{v_i} F(v_i)^{n-1}(v_i - b_i(v_i))f(v_i)dv_i,$$

where f is the density function. Note that this expression is independent of the uncertainty about other players' types. This is because of the particular strategies (symmetric and monotone) strategies that we are considering.

11.2 Bayesian Equilibrium

As we saw, there are two points at which a player may evaluate his utility: ex-ante or interim. Depending on that the notion of equilibrium can be defined. The ex-ante notion coincides with the idea of a Nash equilibrium.

Definition **15** *A strategy profile s^* is a* **Nash equilibrium** *in a Bayesian game if for each player i and each pure strategy s_i,*

$$U_i(s_i^*, s_{-i}^*) \geq U_i(s_i, s_{-i}^*).$$

There is also an interim way of defining the equilibrium. This is called the Bayesian equilibrium, and is the common way of defining equilibrium in Bayesian games.

Definition **16** *A strategy profile s^* is a* **Bayesian equilibrium** *in a Bayesian game if for each player i, each type $t_i \in T_i$, and each action $a_i \in A_i(t_i)$,*

$$U_i((s_i^*(t_i), s_{-i}^*)|t_i) \geq U_i((a_i, s_{-i}^*)|t_i) \quad \forall\ t_i \in T_i.$$

Informally, it says that a player i of type t_i maximizes his expected/interim payoff by following s_i^* given that all other players follow s_{-i}^*.

The first property that we show is that (with finite type spaces) a strategy profile is a Nash equilibrium if and only if it is a Bayesian equilibrium. In other words, a player has a profitable deviation in Bayesian game before he learns his type if and only if he has a profitable deviation after he learns his type. This result will use the fact that probability of every type occurring is positive.

Theorem **13** *Suppose type space of each player is finite. A strategy profile is a Bayesian equilibrium if and only if it is a Nash equilibrium.*

Proof: Consider a strategy profile s^*. Suppose s^* is a Bayesian equilibrium. Then, for every $i \in N$, for every $t_i \in T_i$, and every $a_i \in A_i(t_i)$, we have

$$U_i((s_i^*(t_i), s_{-i}^*)|t_i) \geq U_i((a_i, s_{-i}^*)|t_i).$$

For any strategy $s_i : T_i \to \cup_{t_i} A_i(t_i)$ with $s_i(t_i) \in A_i(t_i)$ for all t_i, we know from Equality 5 that

$$U_i(s_i, s_{-i}^*) = \sum_{t_i \in T_i} \pi_i(t_i) U_i((s_i(t_i), s_{-i}^*)|t_i) \leq \sum_{t_i \in T_i} \pi_i(t_i) U_i(s_i^*(t_i), s_{-i}^*|t_i) = U_i(s_i^*, s_{-i}^*).$$

Hence, s^* is a Nash equilibrium.

Now, suppose that s^* is a Nash equilibrium. Assume for contradiction that s^* is not a Bayesian equilibrium. Then, there is some $i \in N$ and some $t_i \in T_i$ with $a_i \in A_i(t_i)$ such that

$$U_i((a_i, s_{-i}^*)|t_i) > U_i((s_i^*(t_i), s_{-i}^*)|t_i). \tag{6}$$

Now, construct a new strategy s_i such that $s_i(t_i) = a_i$ but $s_i(t'_i) = s'_i(t'_i)$ for all $t'_i \neq t_i$.

Now, observe the following:

$$\begin{aligned}
U_i((s_i, s^*_{-i})) &= \sum_{t'_i \neq t_i} \pi_i(t'_i) U_i((s_i(t'_i), s^*_{-i})|t'_i) + \pi_i(t_i) U_i((s_i(t_i), s^*_{-i})|t_i) \\
&= \sum_{t'_i \neq t_i} \pi_i(t'_i) U_i((s^*_i(t'_i), s^*_{-i})|t'_i) + \pi_i(t_i) U_i((s_i(t_i), s^*_{-i})|t_i) \\
&> \sum_{t'_i \neq t_i} \pi_i(t'_i) U_i((s^*_i(t'_i), s^*_{-i})|t'_i) + \pi_i(t_i) U_i((s^*_i(t_i), s^*_{-i})|t_i) \\
&= U_i(s^*_i, s^*_{-i}),
\end{aligned}$$

where the strict inequality followed from Inequality 6 and the fact that $\pi_i(t_i) > 0$ for all i and for all t_i. This contradicts the fact that s^* is a Nash equilibrium. ■

The equivalence result needs type spaces to be finite. In general, we will consider Bayesian games where type space is not finite. In such games a Bayesian equilibrium will continue to imply a Nash equilibrium but the converse need not hold. So, we will use the solution concept Bayesian equilibrium in all the Bayesian games that we analyze.

But do all Bayesian games admit a Bayesian equilibrium? Which Bayesian games admit a Bayesian equilibrium? There is a long literature on this topic, which we will skip. Just like Nash equilibrium, there is a well-behaved class of games that admit a Bayesian equilibrium. One simple way to think of an existence result is to allow for *mixed actions* for every type. Essentially, we enrich the action space but keep the finite nature of type space. That is, after every t_i, the set of actions available to a player is $\Delta A_i(t_i)$, where $A_i(t_i)$ is finite. This will be the analogue of the mixed strategy. Formally, a mixed strategy of Player i is a map $\sigma_i : T_i \to \cup_{t_i \in T_i} \Delta A_i(t_i)$ such that for every $t_i \in T_i$, $\sigma_i(t_i) \in \Delta A_i(t_i)$. The utility of player i from such a mixed strategy will be evaluated by taking expectation. A mixed strategy Bayesian equilibrium always exists if action spaces are finite and type spaces are finite - a result which we will not prove.

12 Analysis of First-price Auction

We will study a model of selling a single indivisible object. Each agent derives some utility by acquiring the object - we will refer to this as his **valuation**. In the terminology of the Bayesian games, the valuation is the type of the agent.

We will study auction formats to sell the object. This will involve payments. A central assumption in auction theory is that utility from monetary payments is **quasi-linear**, i.e., if an agent gets utility v from the object and pays an amount p, then his net utility is

$$v - p.$$

Implicitly, this assumes risk neutral bidders - the net utility of a bidder is his net payoff.

Another fundamental assumption that is commonly made is that of **no externality**, i.e., if an agent does not win the object then he gets zero utility. The auction that we will study will involve zero payments by the agent who does not win the object. We will assume that all the bidders draw their value from some interval $[0, w]$ using a distribution F (same for all the bidders). We also assume that F admits a density function f such that $f(x) \neq 0$ for all $x \in [0, w]$. It is possible that the interval is the whole non-negative real line, in which case, we will abuse notation to let $w = \infty$. But the mean of this distribution will be finite.

A random variable that will come handy is the highest of $(n-1)$ values: we denote it by G. In particular, we will be interested to know what is the probability that $(n-1)$ bidders have *value* less than or equal to x: this is precisely

$$G(x) = [F(x)]^{n-1}.$$

Like in the Vickrey auction, the highest buyer wins the object in the first-price auction too. We assume that in case of a tie for the highest bid, each bidder gets the good with equal probability. We denote the probability of winning at a profile of bids $b \equiv (b_1, \ldots, b_n)$ as $\phi_j(b)$ for each buyer $j \in N$. Note that $\phi_j(b) = 1$ if $b_j > \max_{k \neq j} b_k$ and $\phi_j(b) = 0$ if $b_j < \max_{k \neq j} b_k$.

Given a profile of bids $b \equiv (b_1, \ldots, b_n)$ of bidders, the payoff to bidder j with value x_j is given by

$$\pi_j(b) = \phi_j(b)\big[x_j - b_j\big]$$

Unlike the Vickrey auction, the first-price auction has no weakly dominant strategy (verify). Obviously, bidding your true value guarantees a payoff of zero, and there are obvious ways to generate positive expected payoff. Hence, we adopt the weaker solution concept of Bayesian equilibrium. In fact, we will restrict ourselves to equilibria where bidders use the same *bidding function* which are technically well behaved.

In particular, for any bidder $j \in N$, a strategy $\beta_j : [0, w] \to \mathbb{R}_+$ is his bidding function. The focus in our study will be **monotone symmetric equilibria**, where every bidder uses the same bidding function. So, we will denote the bidding function (strategy in the Bayesian game) by simply $\beta : [0, w] \to \mathbb{R}_+$. We assume $\beta(\cdot)$ to be strictly increasing and differentiable.

Bayesian equilibrium requires that if every bidder except bidder i follows $\beta(\cdot)$ strategy, then the expected payoff maximizing strategy (over all strategies, including non-symmetric ones) for bidder i must be $\beta(x)$ when his value is x. Note that if bidder i with value x bids $\beta(x)$, and since everyone else is using $\beta(\cdot)$ strategy, increasing β ensures that the probability of winning for bidder i is equal to the probability that x is the highest value, which in turn is equal to $G(x)$. Thus, we can define the notion of a symmetric (Bayesian) equilibrium in this case as follows.

DEFINITION **17** *A strategy profile* $\beta : [0, w] \to \mathbb{R}_+$ *for all agents is a* **symmetric Bayesian equilibrium** *if for every bidder i and every type* $x \in [0, w]$

$$G(x)(x - \beta(x)) \geq \textit{Probability of winning by bidding } b \ (x - b) \qquad \forall \ b \in \mathbb{R}_+,$$

where the probability of winning is calculated by assuming bidders other than bidder i are following $\beta(\cdot)$ *strategy.*

Remember that due to symmetry, $G(x)$ indicates the probability of winning in the auction when the bidder bids $\beta(x)$, and $(x - \beta(x))$ is the resulting payoff.

THEOREM **14** *A symmetric equilibrium in a first-price auction is given by*

$$\beta^I(x) = \frac{1}{G(x)} \int_0^x yg(y)dy.$$

REMARK. The interpretation of this bid function is as follows. A bidder with type/value x for the object bids an amount equal to his *conditional* expectation of the highest value of other bidders, where the conditioning is done on the fact that he has the highest value.

Proof: Suppose every bidder except bidder 1 follows the suggested strategy. The suggested strategy generates non-negative payoff. Let bidder 1 bid b. Notice that if other bidders follow β, the maximum they can bid is $\beta(w)$. So, bidder 1 can focus on bidding no more than $\beta(w)$ - any bid strictly more than $\beta(w)$ can be improved by bidding $\beta(w)$. So, it is without loss of generality to consider $b \in [0, \beta(w)]$. Hence, any bid b can be mapped to a $z = \beta^{-1}(b)$. Then the expected payoff from bidding $\beta(z) = b$ when his true value is x is

$$\begin{aligned}
\pi(b, x) &= G(z)\big[x - \beta(z)\big] \\
&= G(z)x - \int_0^z yg(y)dy \\
&= G(z)x - zG(z) + \int_0^z G(y)dy \\
&= G(z)\big[x - z\big] + \int_0^z G(y)dy,
\end{aligned}$$

where, we have integrated by parts in the fourth equality [3]. Hence, we can write

$$\pi(\beta(x), x) - \pi(\beta(z), x) = G(z)(z - x) - \int_x^z G(y)dy \geq 0.$$

Notice that the previous inequality holds whether $z \leq x$ or $z \geq x$. Hence, bidding according to $\beta(\cdot)$ is a symmetric equilibrium. ■

We now prove that this is the unique symmetric equilibrium in the first-price auction. Now, consider any bidder, say 1. Assume that he realizes a true value x, and wants to determine his optimal bid value b using a symmetric (increasing and differentiable) bidding function β.

Notice that when a bidder realizes a value zero, by bidding a positive amount, he makes a loss. So, $\beta(0) = 0$. Bidder 1 wins whenever his bid $b > \max_{i \neq 1} \beta(x_i)$, where x_i is the value of bidder i. This is equivalent to saying that $b > \beta(\max_{i \neq 1} x_i)$ (since β is increasing) or $\max_{i \neq 1} x_i < \beta^{-1}(b)$. The probability of this event is $G(\beta^{-1}(b))$. Hence, his expected payoff is

$$G(\beta^{-1}(b))(x - b).$$

A necessary condition for maximum is the first order condition, which is obtained by differentiating with respect to b.

$$\frac{g(\beta^{-1}(b))}{\beta'(\beta^{-1}(b))}(x - b) - G(\beta^{-1}(b)), \tag{7}$$

where we used g is the density function of G and $\beta(\beta^{-1}(b)) = b$. If β is an equilibrium bidding strategy Expression 7 must equal zero for all x when $b = \beta(x)$. This implies for all x we must have

$$G(x)\beta'(x) + g(x)\beta(x) = xg(x)$$
$$\Leftrightarrow \frac{d}{dx}(G(x)\beta(x)) = xg(x).$$

Integrating both sides, and using $\beta(0) = 0$, we get

$$\beta(x) = \frac{1}{G(x)} \int_0^x yg(y)dy.$$

Hence, this is the unique symmetric equilibrium in the first-price auction.

[3] To remind, integration by parts $\int h_1(y)h_2'(y)dy = h_1(y)h_2(y) - \int h_1'(y)h_2(y)dy$.

The equilibrium bid in the first-price auction can be rewritten as

$$\beta^I(x) = x - \int_0^x \frac{G(y)}{G(x)} dy.$$

This amount is less than x. From the proof of the Theorem 14, it can be seen that if a bidder with value x bids $\beta(z')$ with $z' > z$, then his loss in payoff is the shaded area above the $G(\cdot)$ curve in Figure 5. On the other hand, if he bids $\beta(z'')$ with $z'' < z$, then his loss in payoff is the shaded area below the $G(\cdot)$ curve in Figure 5.

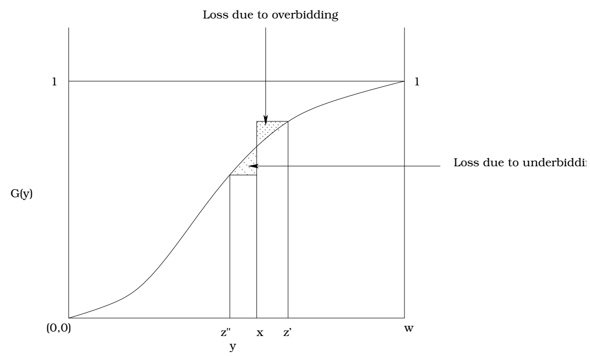

Figure 5: Loss in first-price auction by deviating from equilibrium

Another observation is that $\frac{G(y)}{G(x)} = (\frac{F(y)}{F(x)})^{n-1}$. Hence, the amount of lowering of bid vanishes to zero as the number of bidders increase, and the equilibrium bid amount approaches the true valuation.

Hence, the expected payment in the first price auction for a bidder with value x can be written as

$$\pi^I(x) = G(x)\beta(x) = \int_0^x yg(y)dy.$$

Now, consider the second-price auction. A bidder with value x pays an amount equal to the highest of other $(n-1)$ values *if* he is the highest valued bidder. Hence, his expected payment is the probability that he has the highest value (which is $G(x)$) times the conditional expected value of the highest of other values:

$$\pi^{II}(x) = G(x)\frac{1}{G(x)} \int_0^x yg(y)dy = \int_0^x yg(y)dy = \pi^I(x).$$

This leads to an important observation.

Theorem **15 (Revenue equivalence)** *Suppose values of bidders are distributed independently and identically. Then the expected revenue from the first-price auction and the second-price auction is the same.*

It is instructive to look at the following example. Suppose values are distributed uniformly in $[0,1]$. So, $F(x) = x$ and $G(x) = x^{n-1}$. So, $\beta(x) = x - \frac{1}{x^{n-1}} \int_0^x y^{n-1} dy = x - \frac{x}{n} = \frac{n-1}{n} x$. So, in equilibrium, every bidder bids a constant fraction of his value.

12.1 Analysis of Bilateral Trading

The bilateral trading is one of the simplest model to study Bayesian games. It involves two players: a buyer (b) and a seller (s). The seller can produce a good with cost c and the buyer has a value v for the good. Suppose both the value and the cost are distributed *uniformly* in $[0,1]$.

Now, consider the following Bayesian game. The buyer announces a price p_b that he is willing to pay and the seller announces a price p_s that she is willing to accept. Trade occurs if $p_b > p_s$ at a price equal to $\frac{p_b+p_s}{2}$. If $p_b \leq p_s$, then no trade occurs.

The type of the buyer is his value $v \in [0,1]$ and the type of the seller is his cost $c \in [0,1]$. A strategy for each agent is to announce a price given their types. In other words, the strategy of the buyer is a map $p_b : [0,1] \to \mathbb{R}$ and $p_s : [0,1] \to \mathbb{R}$.

If no trade occurs, then both the agents get zero payoff. If trade occurs at price p, then the buyer gets a payoff of $v - p$ and the seller gets a payoff of $p - c$.

Theorem **16** *There is a Bayesian equilibrium (p_b^*, p_s^*) in the bilateral trading problem with uniformly distributed types in $[0,1]$, where for every $v, c \in [0,1]$,*

$$p_b^*(v) = \frac{2}{3}v + \frac{1}{12}, \quad p_s^*(c) = \frac{2}{3}c + \frac{1}{4}.$$

Proof: Suppose the seller follows strategy p_s^*. Then he never quotes a price above $\frac{2}{3} + \frac{1}{4} = \frac{11}{12}$. So, the buyer should never quote a price above $\frac{11}{12}$ as a best response - this is because any price strictly above $\frac{11}{12}$ can be improved by lowering it a little further, and hence, cannot be a best response. Similarly, the seller quotes a minimum price of $\frac{1}{4}$. So, the buyer can guarantee himself zero payoff by quoting a price of $\frac{1}{4}$. Since quoting any price below $\frac{1}{4}$ also ensures zero payoff, it is without loss of generality to consider those strategies where the buyer never quotes a price below $\frac{1}{4}$.

Suppose he quotes a price π_b when his value is v. Then, trade occurs if the $p_s^*(c) < \pi_b$ or $c < \frac{3}{2}\pi_b - \frac{3}{8}$. Note that since $\frac{1}{4} \leq \pi_b \leq \frac{11}{12}$, we have $0 \leq \frac{3}{2}\pi_b - \frac{3}{8} \leq 1$.

Let $x_b \equiv \frac{3}{2}\pi_b - \frac{3}{8}$. Then the expected payoff of buyer from bidding π_b at type v is

$$\begin{aligned}\int_0^{x_b}\Big(v - \frac{\pi_b + p_s^*(c)}{2}\Big)dc &= \int_0^{x_b}\Big(v - \frac{\pi_b + \frac{2}{3}c + \frac{1}{4}}{2}\Big)dc \\ &= \big(v - \frac{\pi_b}{2} - \frac{1}{8}\big)x_b - \frac{1}{6}x_b^2 \\ &= \big(v - \frac{1}{3}x_b - \frac{1}{4}\big)x_b - \frac{1}{6}x_b^2 \\ &= (v - \frac{1}{4})x_b - \frac{1}{2}x_b^2.\end{aligned}$$

This is a strictly concave function in π_b, hence, the first order condition gives the unique maximum of the unconstrained problem. The first order condition gives $(v - \frac{1}{4}) - x_b = 0$. This implies that $x_b = \frac{3}{2}\pi_b - \frac{3}{8} = v - \frac{1}{4}$. Hence, $\pi_b = \frac{2}{3}v + \frac{1}{12}$. Note that $\pi_b \in [\frac{1}{12}, \frac{9}{12}]$ satisfies our constraint. Hence, it is a best response to p_s^* strategy of the seller.

A similar optimization exercise solves the seller's problem. Suppose the buyer follows strategy p_b^*. Then, the buyer quotes a minimum of $\frac{1}{12}$ and a maximum of $\frac{3}{4}$. Then the seller should never quote less than $\frac{1}{12}$ because such a strategy will not maximize his expected payoff. Suppose he quotes π_c, then trade occurs if $\pi_c < \frac{2}{3}v + \frac{1}{12}$, which reduces to $v > \frac{3}{2}\pi_c - \frac{1}{8} \geq 0$ since $\pi_c \geq \frac{1}{12}$. Further, $\frac{3}{2}\pi_c - \frac{1}{8} \leq 1$ since $\pi_c \leq \frac{3}{4}$. Denote $x_c = \frac{3}{2}\pi_c - \frac{1}{8}$. Hence, the expected payoff of the seller at type c is given by

$$\begin{aligned}\int_{x_c}^{1}\Big(\frac{\pi_c + \frac{2}{3}v + \frac{1}{12}}{2} - c\Big)dv &= \int_{x_c}^{1}\Big(\frac{1}{2}\pi_c + \frac{1}{24} - c + \frac{1}{3}v\Big)dv \\ &= \int_{x_c}^{1}\Big(\frac{1}{3}\pi_c + \frac{1}{12} - c + \frac{1}{3}v\Big)dv \\ &= \Big(\frac{1}{3}\pi_c + \frac{1}{12} - c\Big)(1 - x_c) + \frac{1}{6}(1 - x_c^2).\end{aligned}$$

Again this is a strictly concave function and its maximum can be found by solving the first order condition. The first order condition gives us

$$\frac{1}{3}(1 - x_c) - \Big(\frac{1}{3}\pi_c + \frac{1}{12} - c\Big) - \frac{1}{3}x_c = 0.$$

This gives us $x_c = \frac{3}{2}\pi_c - \frac{1}{8} = c + \frac{1}{4}$, which gives the unique best response as $\pi_c = \frac{2}{3}c + \frac{1}{4}$. ■

There are other Bayesian equilibria of this game. However, this equilibrium can be shown to be unique in the class of strategies where players use strategies linear in their type. One notable feature of this equilibrium is that trade occurs when $p_b^*(v) > p_s^*(c)$, which is equivalent to requiring $\frac{2}{3}v + \frac{1}{12} > \frac{2}{3}c + \frac{1}{4}$. This gives $v - c > \frac{1}{4}$. Note that efficiency will require trade to happen when $v > c$. Hence, there is some loss in efficiency. This is in general an impossibility

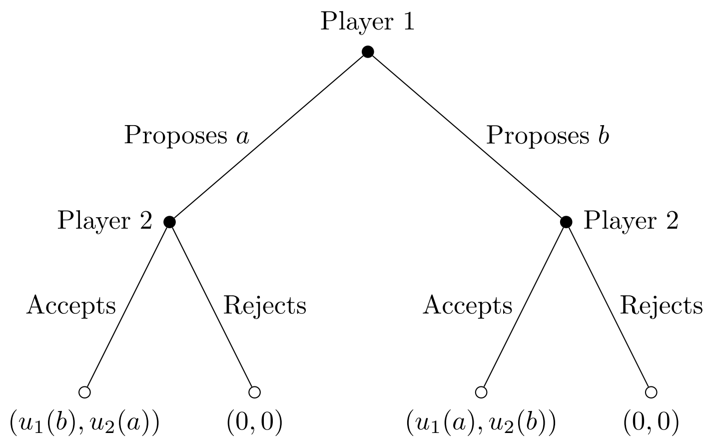

Figure 6: Extensive form game with perfect information

- you cannot construct any Bayesian game whose equilibrium will have efficiency in Bayesian equilibrium in this model (more on this in some advanced course).

13 Extensive Form Games

In many situations strategic interactions between agents happen sequentially. Unlike in strategic form games, agents move sequentially in such games. We consider some examples first.

Suppose two players are deciding how to share two indivisible objects $\{a, b\}$. First, Player 1 proposes an allocation. Player 2 observes the proposal of Player 1 and then decides whether to accept or reject the proposal. If Player 2 rejects, then no player gets any object. If Player 2 accepts the proposal, then each receives the proposed allocation of Player 1. Each player $i \in \{1, 2\}$ only cares about his own object and has a utility function $u_i \equiv (u_i(a), u_i(b))$, indicating his utility for the objects.

This situation can be modeled as an extensive game of perfect information. This is usually depicted by a game tree.

An important feature of this game is that Player 2 has completely observed what Player 1 has proposed. His action is contingent on what he has observed so far in the game. Such games are called extensive form games with perfect information, i.e., where every player has perfectly observed what has happened so far in the game at every point. The outcomes of the game are realized after the game ends. Players assign payoffs to this terminal stages of the game - this will involve assigning payoffs to every possible sequence of moves in the game.

Figure 6 depicts the extensive form game using a tree. The payoffs of the agents are written in the *leaf* nodes.

A strategy in such a game is a complex object. It must state the action to be taken for every contingent path that can be taken in this game.

We now look at another example where perfect information is absent. Suppose two friends are trying to meet. Friend 1 observes the weather in his city, which is either rain or sunny. Then, he decides to either go to Friend 2's place or stay at home. If Friend 1 stays at home, Friend 2 does not do anything and the game ends. If Friend 1 comes to Friend 2's place, she either takes him for dinner or cooks at home. Crucial here is the fact that Friend 2 does not observe the weather in Friend 1's city, which Friend 1 has observed. However, Friend 2 observes whether he Friend 1 has come to her place or not. But Friend 2 does not know if Friend 1 has come from a sunny city or rainy city. In that sense, though the game has sequential nature, the information is not perfect in this game.

There is a way to represent this game as an extensive form game with imperfect information. This is done by introducing the dummy player (Nature) who creates the imperfect information. Nature makes the first move by taking either the action "Rainy" or "Sunny". The action of Nature is observed by Friend 1 but not by Friend 2. After observing the action of Nature, Friend 1 takes either of the actions "Stay home" or "Go to Friend 2". Friend 1 can now come to Friend 2 from a Sunny city or a Rainy city. This idea is captured by an *information set*, where a bunch of nodes in the game are combined together to capture Friend 2's uncertainty about where she is in the game. Irrespective of where she is in the game, she observes that Friend 1 has come to her place, and then she chooses one of the actions "go out" or "stay in".

Figure 7 shows the extensive form game with information set. The information set of Player 2 is shown in dashed rectangle - it consists of two nodes in the game tree. At this information set, Player 2 does not know if Player 1 has come from a sunny city or rainy city.

Each of the possible paths in the game are assigned a payoff for each player. Further, games of imperfect information also specify probabilities/priors of uncertain moves of Nature. These are used to compute expected payoffs on information sets.

14 Extensive Form Games with Perfect Information

We now formally define the notion of an extensive form game. We start from the most basic extensive game - a perfect information game, where every player at every node in the game knows what path/history has brought him to that node.

To formally define an extensive form game, we need to define a *cycle-free* graph. A graph

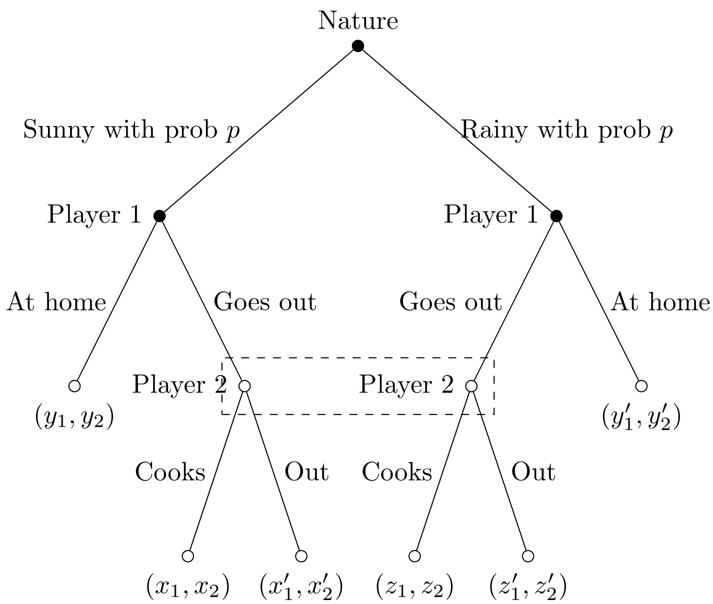

Figure 7: Extensive form game with information sets

$G = (V, E)$ is a set of a vertices V and subset of unordered pairs $E \subseteq V \times V$ such that for all $\{i, j\} \in E$, $i \neq j$. A cycle in a graph G is a sequence of distinct vertices $v_1, \ldots, v_k$ with $k > 2$ such that $\{v_1, v_2\}, \ldots, \{v_{k-1}, v_k\}, \{v_k, v_1\}$ are all edges of the graph. A graph G is cycle-free if there are no cycles in G.

A path in a graph G is a sequence of distinct vertices $v_1, \ldots, v_k$ such that $\{v_1, v_2\}, \ldots, \{v_{k-1}, v_k\}$ are all edges of the graph. A graph is connected if there is a path from every vertex to every other vertex. A connected and cycle-free graph is called a *tree*.

An important property of a tree graph is that there is a *unique* path from every vertex to every other vertex. From every tree $G = (V, E)$, we can construct a *rooted tree* by choosing a root vertex $r \in V$. A rooted tree is represented by $G \equiv (V, E, r)$. In a rooted tree, G, a vertex v is called the *child* of v' if there is an edge $\{v, v'\}$ and v' is in the unique path from root r to v. The set of all children of a vertex v is denoted by $C(v)$. Any vertex v with no children, i.e., $C(v) = \emptyset$ is called a *leaf* vertex.

An example of a rooted tree is shown in Figure 8. The root of this tree is shown. The leaves of the tree are $\{v_3, v_6, v_7, v_8, v_9, v_{10}\}$. For child: v_5 is the only child of v_2, whereas v_1 has two children: $\{v_3, v_4\}$.

The backbone of an extensive form game is a rooted tree.

Definition 18 *An* **extensive form game of perfect information** *is*

$$\Gamma \equiv \Big(N, (V, E, r), \{V_i\}_{i \in N}, \{A(x)\}_{x \in V}, \{u_i\}_{i \in N}\Big),$$

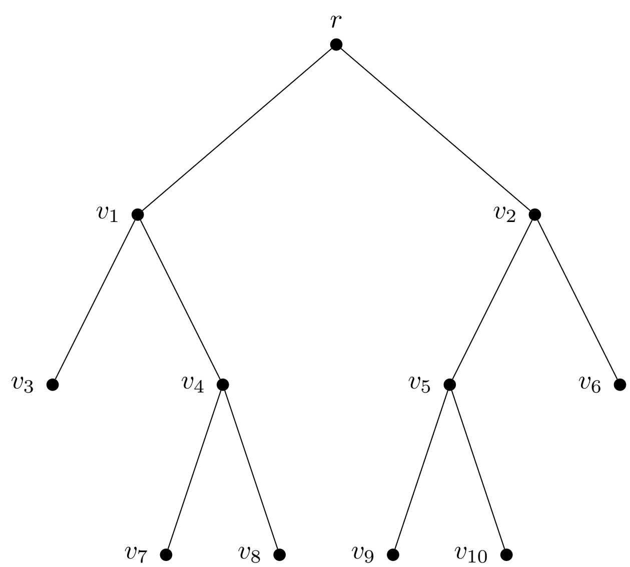

Figure 8: An example of a rooted tree

where

- N *is the set of players*
- (V, E, r) *is a game tree, where*
  - *Each non-leaf vertex* $x \in V$ *specifies a player, called the decision maker at* x*, in* N *who will take an action at this vertex.*
  - *Each leaf or terminal vertex* $x \in V$ *is a* **payoff** *vertex.*
  - *Each edge* $\{x, y\} \in E$ *represents an action, in particular decision maker at* x *takes an action specified by this edge to reach vertex* y*.*
  - *Root vertex* r *specififes the first player in* N *to take an action.*
- $A(x)$ *is the set of actions available at vertex* x *(they identify the set of edges from* x *which lie on the path from* x *to all the leaf nodes). Note that if* x *is a leaf vertex, then* $A(x)$ *is an empty set.*
- $\{V_i\}_{i \in N}$ *is a partitioning of the set of decision vertices. Hence,* V_i *represents the set of decision vertices where Player* i *takes action.*
- *For every player* $i \in N$*,* $u_i(x)$ *assigns a payoff for every terminal vertex* x *to Player* i*.*

We note here that the set of vertices/edges in a game tree may be infinite. This can happen because of two reasons: (1) the set of actions available at a vertex may be infinite and/or (2) the set of stages (i.e., lengths of paths) of the game may be infinite. At every vertex x in an extensive form game, the unique path from root r to vertex x conveys a lot of information: it contains information about who are the players who have taken what action to reach from r to x. It is standard to denote this information on the path as **history** h_x at vertex x. In fact, an alternate representation of an extensive form game is to just specify the history at every vertex.

Consider the following example of Figure 6. There is only one vertex, the root vertex, where Player 1 is the decision maker. For all other non-leaf nodes, Player 2 is the decision maker. Player 1 has two actions available to him - the two proposals he can make to Player 2. In each of his vertices, Player 2 has the same two actions (Accept, Reject) available to him. The payoffs of both the players are shown on the leaf vertices.

A strategy for a player in an extensive game must specify what he will do at each of his decision vertices. Hence, you can imagine a Player telling a computer to play on his behalf. In that case, he does not know ex-ante which decision vertices will be reached. So, he gives the computer a complete contingent plan of what actions must be taken at every decision vertex.

Formally, a **strategy** of player $i \in N$ is a map

$$s_i : V_i \to \cup_{x \in V_i} A(x) \text{ such that } s_i(x) \in A(x) \ \forall \ x \in V_i.$$

Notice that there are certain games, where every player moves only once - these games are said to satisfy the *single move property.* However, there are games in which the single move property is not satisfied. In those games, if a strategy specifies a certain action at a decision vertex, that may ensure that certain decision vertex is never reached. But that does not exclude us from describing what action to take in those unreached vertices.

To see this, consider the game in Figure 9, where Player 2 moves twice. If Player 2 plays a strategy where he says he "Calls Player 1" at the first vertex, then exactly one more of his decision vertex will be reached. But a strategy for Player 2 must specify his action at *all* the decision vertices. This is crucial to evaluating his and his opponent's options.

15 Equilibrium for Extensive Form Games

We discuss equilibrium concepts for extensive form games. One naive way of doing that is to represent it as a strategic form game, and then apply the solution concepts of strategic form games. Representing an extensive form game as a strategic form game is quite easy:

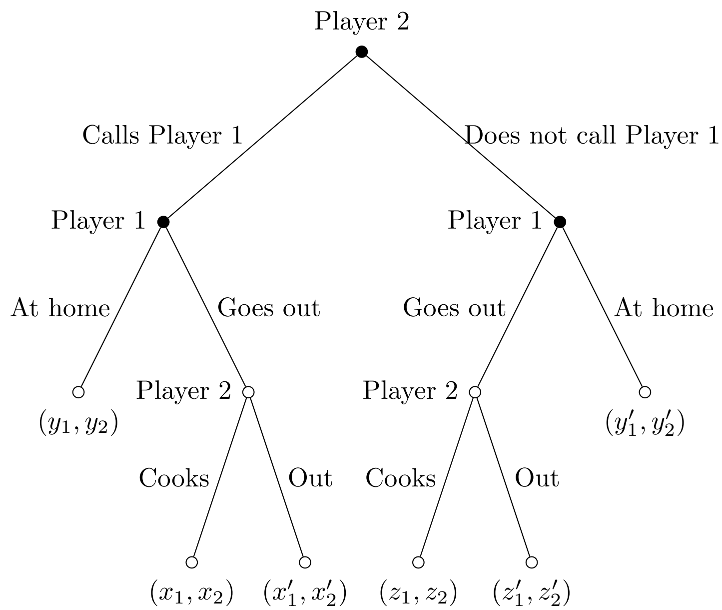

Figure 9: Extensive form game without single move property

for every player i and every strategy of i in the extensive form game corresponds to a pure strategy in the strategic form game. The payoff from a strategy profile can then be computed from the game tree. This is because each strategy profile in the extensive form game maps to a unique terminal vertex of the game tree. This is called the **reduced normal/strategic form** of the extensive game. For a strategy profile s in an extensive form game Γ, we let x_s as the terminal vertex reached because of the strategy profile s. Then, the payoff of agent i from a strategy profile s is $u_i(x_s)$.

DEFINITION **19** *A strategy profile* $s \equiv (s_1, \ldots, s_n)$ *is a Nash equilibrium of* Γ *if for all* $i \in N$ *and for all* s'_i

$$u_i(x_{(s_i, s_{-i})}) \geq u_i(x_{(s'_i, s_{-i})}).$$

This definition just says that consider the reduced-form strategic form game and consider the Nash equilibrium of that game. In other words, it ignores all the extensive form (sequential) play of players actions in the game. Hence, Nash equilibrium is not the correct solution concept for extensive form games. We illustrate this with an example.

Consider the game in Figure 10. The reduced strategic form representation of this game is shown in Table 27. From this, one concludes that the game has two pure strategy Nash equilibria: (U, L) and (D, R).

But note that once the game has reached the information set of Player 2, he will play R. So, playing L is not *credible* for Player 2. Then, Player 1 can take this information into

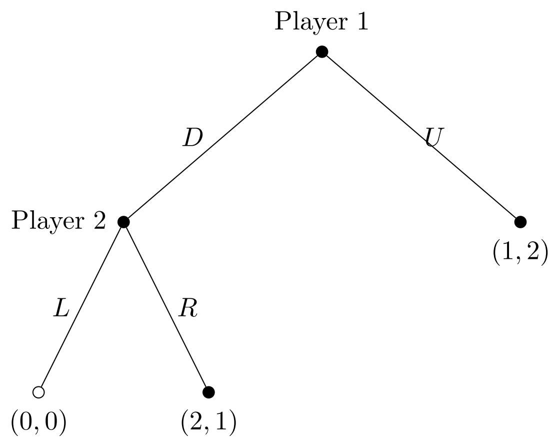

Figure 10: Nash equilibrium

	L	R
U	$(1,2)$	$(1,2)$
D	$(0,0)$	$(2,1)$

Table 26: Reduced strategic form of the game in Figure 10

account while choosing his action. Player 1 clearly prefers playing D over U since Player 2 cannot threaten him credibly to play L. Hence, the equilibrium (U, L) is not a good prediction of the game.

The main idea here is that the equilibrium (U, L) specifies a strategy L for Player 2 which is not a credible strategy - once the decision vertex of Player 2 is reached, he will never play this.

As we discussed above, a strategy profile leads to a unique terminal vertex with a unique path from root to the terminal vertex. Hence, an equilibrium strategy profile will not touch on many decision vertices - these are called *off-equilibrium path* decision vertices. One primary requirement in extensive form game equilibrium is that action of every player must be optimal starting at *every* decision vertex, and *not just* decision vertices reached on equilibrium path.

15.1 Subgame Perfect Equilibrium

We now discuss a *refinement* to Nash equilibrium for extensive form game. This is the single-most important solution concept for extensive form games. it enforces and formalizes the idea of credibility by using the notion of subgames.

The **subgame** of an extensive form game of perfect information

$$\Gamma \equiv \Big(N, (V, E, r), \{V_i\}_{i \in N}, \{A(x)\}_{x \in V}, \{u_i\}_{i \in N}\Big),$$

starting at $x \in V$, where x is not a leaf vertex, is an extensive form game

$$\Gamma(x) \equiv \Big(N, (V(x), E(x), x), \{V_i(x)\}_{i \in N}, \{A(x')\}_{x' \in V(x)}, \{u_i\}_{i \in N}),$$

where the the (x) in the above notation means that the restriction of the original game starting from vertex x and its children, and children of its children etc.

Note that a game is a subgame of itself. So, every game has a subgame. Game in Figure 9 has many subgames: there are two subgames starting with Player 1's two decision nodes; there are three subgames starting with Player 2's three decision nodes. In general, the number of subgames in a game equals the number of decision nodes in the game.

DEFINITION **20** *A strategy profile s is a* **subgame perfect equilibrium (SPE)** *of the extensive form game Γ if for every subgame of Γ the strategy profile s restricted to that subgame is a Nash equilibrium of the subgame.*

Since Γ itself is a subgame of the game Γ, it follows that every SPE is a Nash equilibrium - hence, SPE is a refinement of Nash equilibrium. We document this as a fact below.

FACT **2** *Every subgame perfect equilibrium is a Nash equilibrium.*

The game in Figure 10 has a unique SPE. To see this, the subgame starting from decision vertex of Player 2 has only one player. In that, Player 2 playing R is a dominant strategy. So, out of the two Nash equilibria of the entire game (subgame), only the one with R being played by Player 2 survives. Hence, (D, R) is the unique SPE.

Figuring out Nash equilibrium of subgames can be quite a complicated task. In games with perfect information, this can be avoided because of a well known equivalence of subgame perfect equilibrium with two other notions. The first is the idea of *sequential rationality*.

DEFINITION **21** *A strategy s_i of Player i is* **sequentially rational** *given s_{-i} if for each decision vertex x of Player i, s_i restricted to subgame at x is a best response to s_{-i} restricted to the subgame at x.*

The main difference between subgame perfect equilibrium and sequential rationality is that sequential rationality requires that at each subgame starting at decision vertex x, only the *decision maker* of decision vertex x must be choosing a best response. On the other hand, the subgame perfect equilibrium requires at every subgame, strategy of *every* player must be a best response given strategies of other players.

We now define a further weakening of subgame perfect equilibrium.

DEFINITION **22** *A strategy s_i of Player i satisfies* **one-shot deviation principle** *given s_{-i} if for each decision vertex x of Player i and each strategy s'_i of Player i which only differs from s_i by the action chosen at x, we have*

$$u_i(x_{(s_i,s_{-i})}) \geq u_i(x_{(s'_i,s_{-i})}).$$

A fundamental result is that all these notions are the same. We now prove a useful theorem for understanding subgame perfect equilibrium. The result below allows for a decision vertex to have arbitrary (possibly infinite) number of actions - hence, the game tree may have infinite number of vertices. However, it restricts itself to games having finite number of *stages*. To understand the notion of stage, let $L(x)$ denote the length of the longest path from a decision vertex x to any terminal vertex reachable from x. We will say a game Γ has **finite number of stages** if $L(x)$ is finite for each decision vertex x.

THEOREM **17** *Let Γ be an extensive form game of perfect information and finite number of stages. Then the following are equivalent.*

1. *s is a subgame perfect equilibrium.*

2. *For every $i \in N$, s_i is sequentially rational given s_{-i}.*

3. *For every $i \in N$, s_i satisfies one-shot deviation principle given s_{-i}.*

Proof: The implications (1) $\Rightarrow$ (2) and (2) $\Rightarrow$ (3) are immediate from definitions. To see this note that sequential rationality requires that at every subgame starting with vertex x only the decision maker at x must best respond - but subgame perfect equilibrium needs that everyone must best respond. Hence, (1) $\Rightarrow$ (2). For (2) $\Rightarrow$ (3), consider the optimization problem done in the backward induction procedure and sequential rational strategies. Suppose Player i is a decision maker of a decision vertex x. Denote the strategy profile s restricted to subgame from x as s^x. To verify sequential rationality of s^x_i given s^x_{-i}, we need to check for deviations at all decision vertices in the subgame. For backward induction, as we had argued earlier, we only need to check deviations of one decision vertex at a time. We now establish the other directions.

(2) $\Rightarrow$ (1). Suppose s is a sequentially rational strategy profile. Let x be a decision vertex of agent i. We need to show that s^x is a Nash equilibrium of the subgame $\Gamma(x)$. Count the length of the paths from x to every possible terminal vertex reachable from x, and denote the length of the maximal path by $\ell(x)$. We do the proof by induction on $\ell(x)$. If $\ell(x) = 1$,

then the proof follows from sequential rationality itself. Assume $\ell(x) > 1$ and suppose that the claim is true for all y with $\ell(y) < \ell(x)$.

First, note that by definition, s_i^x is a best response to s_{-i}^x. Consider any player $j \neq i$. If j does not have a decision vertex in the subgame $\Gamma(x)$, then his strategy is vacuously a best response in this subgame. If j has a decision vertex in this subgame, let y be the first such decision vertex when we go from x to a terminal vertex. By induction and sequential rationality, s_j^y is a best response to s_{-j}^y.

Now, j's strategy in the subgame at x is the union of his strategies in each such y. Since each of them is a best response to s_{-j}^y for each y by induction, his strategy in the subgame $\Gamma(x)$ is also a best response to s_{-j}^x. This shows that s is a Nash equilibrium.

(3) $\Rightarrow$ (2). Fixing the strategies of other players at s_{-i}, we will show that if for strategy s_i there is no strategy s_i' which is a one-shot profitable deviation, then s_i is sequentially rational given s_{-i}. Assume for contradiction that this is not true. Then, there is a decision vertex x of agent i such that s_i^x has a profitable deviation in subgame $\Gamma(x)$. Let $s_i'^x$ be such a strategy resulting in a profitable deviation in $\Gamma(x)$.

Consider all paths from x to a terminal vertex in the subgame $\Gamma(x)$ and let $L(x)$ denote the maximum number decision vertices along any such path that belongs to i. If $L(x) = 1$, then this will imply that s_i^x has a profitable one-shot deviation, contradicting one-shot deviation principle. So, $L(x) > 1$. We can now use induction on $L(x)$. Suppose the claim is true for all decision vertices y of agent i with $L(y) < L(x)$ - the base case of induction has already been established.

Now, we travel from x to all terminal vertices in the game tree and record the *first* decision vertex of Player i after x that we encounter along these paths. Denote the set of all these decision vertices as Y - note that this can potentially be an infinite set if the action set at some decision vertex is infinite. Since $s_i'^x$ is not a one-shot deviation, it must be the case that Player i takes different action in $s_i'^x$ than s_i^x in a subgame $\Gamma(y)$ for some $y \in Y$. But note that $L(y) < L(x)$ for all $y \in Y$. By our induction hypothesis, s_i^x restricted $\Gamma(y)$ is a best response to s_{-i} restricted to $\Gamma(y)$ for each $y \in Y$. This means that for every $y \in Y$, if we change the action specified in s_i' to the action specified in s_i in subgame $\Gamma(y)$, the payoff of Player i will increase. Call such a strategy $s_i''^x$. Since $s_i'^x$ is a profitable deviation, $s_i''^x$ is also a profitable deviation. But note that $s_i''^x$ is either a one-shot deviation strategy or $s_i''^x = s_i^x$ since $s_i''^x$ and s_i^x are identical in subgame $\Gamma(y)$ for all $y \in Y$. This means s_i^x does not satisfy the one-shot deviation property given s_{-i} at decision vertex x, which is a contradiction. ■

Finally, an easy method to compute a strategy profile satisfying one-shot deviation principle in finite extensive form game is the following. Start with a decision vertex just before a terminal vertex. Specify an action that leads to the highest payoff for the decision maker of that vertex among all possible actions - in case of ties, all possible actions leading to highest payoff are specified. If such an optimal action leads to terminal vertex z, then replace this decision vertex and the subsequent subgame by terminal vertex z. Repeat this procedure. If indifferences occur, this will lead to multiple strategy profiles surviving. This procedure is called the **backward induction** procedure.

DEFINITION **23** *A strategy profile that survives the above procedure is said to be a strategy profile* **surviving the backward induction procedure**.

In the game in Figure 10, Player 2 plays R. Then we replace the subgame starting at the decision vertex of Player 2 by payoff $(2, 1)$. Now, Player 1 chooses D in this new game. Hence, the unique outcome of the backward induction procedure is (D, R).

Consider the game in Figure 11. There are three players: two entrant firms and one incumbent firm. The entrants decide sequentially whether to stay out (O or o) or enter the market (E or e). If they stay out they get zero. If they enter, then the incumbent can fight ($f/f'/f''$) or accommodate ($a/a'/a''$). If both entrants stay out, the incumbent gets 5. If the entrant accommodates, the per firm profit is 2 for duopoly and -1 for triopoly. On top of this, if the incumbent fights, then it costs 1 for the incumbent and 3 for entrants. The game is described in Figure 11.

If we solve this game by backward induction procedure, then the incumbent always accommodates. Given this, entrant firm 2 enters in his left-most information set but stays out in the right-most information set. Given this, entrant firm 1 enters. This illustrates the idea of a first-mover advantage in extensive form games.

How do we describe the subgame perfect equilibrium of this game? We need to specify the actions at every information set: $(E, (e, o'), (a, a', a''))$. You can verify that there are many Nash equilibria of this game. Hence, Nash equilibrium has very less predictive power in this game but the subgame perfect equilibrium leads to a unique outcome.

Backward induction can be a very demanding solution in games where players need to move many times. This is because it requires players to anticipate actions down the game tree. A sharp example of this fact is given a well known game called the **centipede game**. Two players start with 1 unit of money each. Each player can either decide to continue C or stop S. If anyone stops, then the game ends and each take their piles. If a player continues, then the opponent gets to take action but his pile is reduced by 1 while the opponent's pile is increased by 2. The play ends when any player reaches 100. Suppose Player 1 moves first.

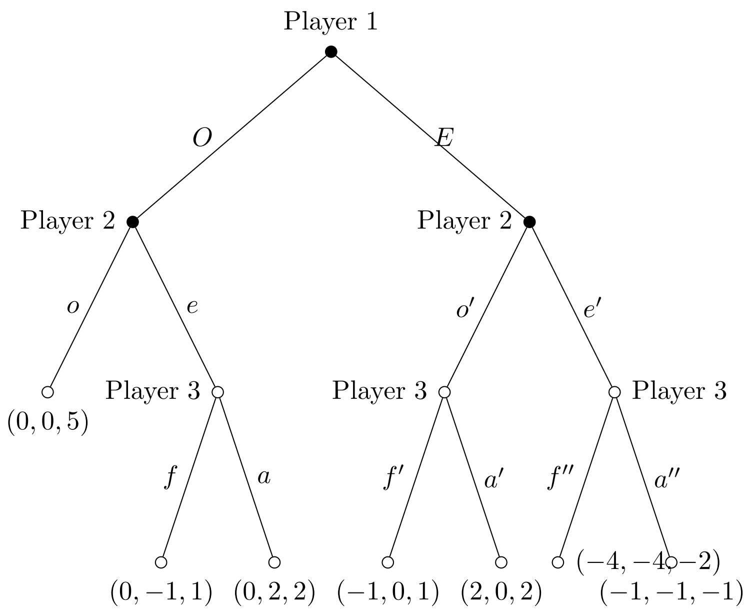

Figure 11: Backward induction

Unique prediction due to backward induction is Player 1 stops in the first chance resulting in $(1, 1)$. The subgame perfect equilibrium specifies action S at every decision vertex. This is also the unique Nash equilibrium of this game.

In lab experiments, agents have usually continued for some time. This is a general critique of equilibrium in extensive form game that no satisfactory refinement can predict such an outcome.

We will often refer to all these notions to be the definition of a subgame perfect equilibrium in such games. An immediate corollary of Theorem 17 is that a subgame perfect equilibrium in pure strategies always exist - this follows from the fact that the backward induction procedure always generates at least one pure strategy profile. If there are no indifferences in payoffs, the backward induction procedure generates a unique strategy profile, which is referred to as the backward induction *solution*.

16 Mixed and Behavior Strategies

We have defined pure strategies in an extensive form game as a map that defines what action a player will take in each of his decision vertices. There are two natural ways to define *randomized* strategies in this environment. The first one says that we define a probability distribution over the set of all pure strategies. This is the notion of a mixed strategy.

Formally, a **mixed strategy** of Player i is $\sigma_i \in \Delta \prod_{x \in V_i} A(x)$.

Consider the game in Figure 12. Player 1 has two pure strategies - we roughly write it as $\{x, y\}$ to denote that in his only decision vertex, he can either choose action x or action y. Similarly, the pure strategies of Player 2 can be written as $\{Aa, Ar, Ra, Rr\}$, where Aa indicates that in his left-most decision vertex he plays A and in the other decision vertex, he plays a - similar interpretations can be made for other pure strategies. A mixed strategy of Player 1 will be $\sigma_1(x), \sigma_1(y)$ such that $\sigma_1(x) + \sigma_1(y) = 1$. A mixed strategy of Player 2 will be $\sigma_2(Aa), \sigma_2(Ar), \sigma_2(Ra), \sigma_2(Rr)$ such that

$$\sigma_2(Aa) + \sigma_2(Ar) + \sigma_2(Ra) + \sigma_2(Rr) = 1.$$

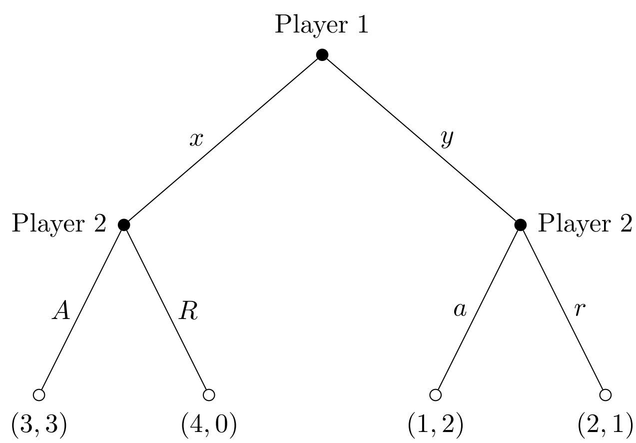

Figure 12: Extensive form game with perfect information

Another way to specify random behavior in this game is to specify a probability distribution at each decision vertex. A **behavior strategy** of Player i specifies a probability distribution b_i^x over $A_i(x)$ for each of his decision vertices x. Hence, $b_i \in \prod_{x \in V_i} \Delta A(x)$. Notice that every behavior strategy naturally induces a probability distribution over pure strategies, and hence, is a mixed strategy.

In the game in Figure 12, Player 2 will have to specify two maps: $b_2^1(A), b_2^1(R)$ with $b_2^1(A) + b_2^1(R) = 1$ and $b_2^2(a), b_2^2(r)$ with $b_2^2(a) + b_2^2(r) = 1$. Note that the induced mixed strategy of Player 2 can be computed by multiplying the respective probabilities: for instance, $\sigma_2(Aa) = b_2^1(A)b_2^2(a)$. Thus, specifying randomization using a behavior strategy assumes independence across decision vertices - when a player reaches his decision vertex, he randomizes over the actions at that decision vertex only.

Since mixed strategies allow for correlation, not every mixed strategy can be induced from behavior strategies. To see this, consider the game in Figure 12. Suppose $b_2^1(A) = \frac{1}{2} = b_2^1(R)$

and $b_2^2(a) = \frac{1}{3}$, $b_2^2(r) = \frac{2}{3}$. The mixed strategy generated is

$$\sigma_2(Aa) = \frac{1}{6}, \sigma_2(Ar) = \frac{1}{3}, \sigma_2(Ra) = \frac{1}{6}, \sigma_2(Rr) = \frac{1}{3}.$$

Now, consider the following mixed strategy of Player 2,

$$\sigma_2(Aa) = \frac{1}{3}, \sigma_2(Ar) = \frac{1}{6}, \sigma_2(Ra) = 0, \sigma_2(Rr) = \frac{1}{2}.$$

If there is a behavior strategy of Player 2 that generates this mixed strategy, then we must have $b_2^1(R) = 0$ or $b_2^2(a) = 0$, which will then imply that either $\sigma_2(Rr)$ or $\sigma_2(Aa)$ is zero, a contradiction. The main idea here is that behavior strategy does not allow for correlation present in this mixed strategy.

But such correlation is strategically unnecessary. This is because decision vertices are reached sequentially. To make ideas precise, fix a player i and a mixed strategy σ_{-i} of other players. By specifying a behavior strategy b_i, we induce a probability distribution over the terminal vertices of the game tree by the play (b_i, σ_{-i}). Similarly, each σ_i also induces a probability distribution over terminal vertices by the play (σ_i, σ_{-i}).

Formally, let $\rho(x; \sigma)$ denote the probability that a terminal vertex x is reached by playing a strategy profile σ. How is ρ computed? Remember, there is a unique path from the root vertex to x in Γ. Then, $\rho(x; \sigma)$ is the multiplication of playing each of the actions along this path (which can be computed from σ).

We illustrate with the above example. In the above example, suppose Player 1 plays the behavior/mixed strategy where he plays x and y with equal probability. Suppose Player 2 plays strategy σ_2. Then what is the probability of reaching the terminal vertex with payoff $(3, 3)$? It can be reached if Player 1 plays x and Player 2 either plays Aa or Ar. Hence, the required probability is

$$\sigma_1(x) \times \big[\sigma_2(Aa) + \sigma_2(Ar)\big] = \frac{1}{4}.$$

A similar calculation reveals the following distribution over terminal vertices

$$(\frac{1}{4}, \frac{1}{4}, \frac{1}{6}, \frac{1}{3}),$$

where we have written the probabilities of terminal vertices from left to right.

A similar calculation for behavioral strategies can also be done. It can be verified that both the mixed strategy and the behavior strategies give rise to the same distribution over terminal vertices. When computing the probability of a terminal node, we somehow constructed a behavior strategy by adding up all the pure strategies in the support of the pure strategy that lead to this terminal vertex. It so turned out that it was indeed a behavior strategy that we had earlier stated.

Definition 24 *A behavior strategy b_i and a mixed strategy σ_i of Player i are* **outcome equivalent** *if for every mixed strategy σ_{-i} of other players, the probability distributions induced over the terminal vertices by (b_i, σ_{-i}) and (σ_i, σ_{-i}) are the same.*

Formally, Harold Kuhn established the following theorem.

Theorem 18 *In every extensive game of perfect information, every mixed strategy of a player is outcome equivalent to a behavior strategy.*

The proof involves constructing particular behavior strategies for every mixed strategy. Though the proof is notationally quite involved, the idea is relatively straightforward. We illustrate this with an example. Consider Player 2 in the game in Figure 13. Consider a mixed strategy of Player 2 as $\sigma_2(L\ell) = \sigma_2(Lr) = \frac{1}{3}$, $\sigma_2(R\ell) = \frac{1}{12}, \sigma_2(Rr) = \frac{1}{4}$.

Suppose Player 1 plays p_u (for U) and p_d (for D) as his mixed strategy. We need to construct behavior strategies which is outcome equivalent to this. Consider the decision vertex 2 of Player 2. A natural candidate of his behavior strategy is the *conditional* probability of agent 2 playing ℓ (and r can be computed similarly) given that this decision vertex is reached:

$$\frac{\sigma_2(L\ell)p_u}{\sigma_2(L\ell)p_u + \sigma_2(Lr)p_u} = \frac{\sigma_2(L\ell)}{\sigma_2(L\ell) + \sigma_2(Lr)}.$$

Similarly, the candidate behavior strategy for the first decision node of Player 2 is

$$(\sigma_2(L\ell) + \sigma_2(Lr)).$$

These candidates for behavior strategy generates the following probability of reaching the decision vertex with payoff $(4, 1)$:

$$(\sigma_2(L\ell) + \sigma_2(Lr))p_u \frac{\sigma_2(L\ell)}{\sigma_2(L\ell) + \sigma_2(Lr)} = p_u \sigma_2(L\ell),$$

which is also the probability of reaching this decision vertex by strategy profile (p_u, σ_2).

Doing the calculations reveal that the probability distribution induced on terminal vertices $(3, 1), (3, 0), (4, 1), (2, 2)$ respectively are $\frac{1}{3}, p_u\frac{1}{3}, p_u\frac{1}{3}, p_d\frac{2}{3}$.

Clearly, to achieve these probabilities Player 2 must play $\frac{1}{3}$ on R at his first decision vertex. So, he plays L with probability $\frac{2}{3}$. Then, to ensure equivalent outcome, he should play ℓ and r with probability $\frac{1}{2}$ each. Hence, we computed behavior strategy of playing ℓ of Player 2 at his second decision vertex by the following conditional probability:

$$\frac{\sigma_2(L\ell)}{\sigma_2(L\ell) + \sigma_2(Lr)} = \frac{1}{2}.$$

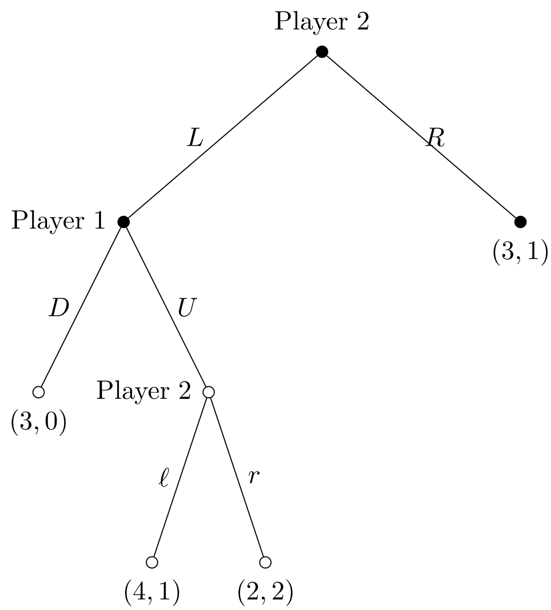

Figure 13: Extensive form game: illustration of Kuhn's theorem

The proof of Kuhn's theorem formalizes this and shows that such computations are always possible.

Because of this result, we will only talk about behavior strategies from now onwards. Notice that the equivalence between one-shot deviation property, sequential rationality, and subgame perfect equilibrium (Theorem 17) continues to hold even with behavior strategies since we allowed for infinite action sets in Theorem 17. However, conceptually, a behavior strategy in an extensive form game is a complicated object - after all, players observe others playing a pure action and not the randomization. One way to think of it is that though players choose pure actions, the randomization device they use is public - this is referred to as *public randomization*. This issue is bypassed by the backward induction procedure because it is based on beliefs down a decision vertex.

Indifference. If there are indifferences, then many pure and mixed strategies will survive backward induction and all of them will be subgame perfect equilibrium. To illustrate this, consider the following example in Figure 14.

In the game in Figure 14, Player 2 is indifferent between his strategies L and R. Suppose he plays L, then optimal strategy for Player 1 is to play U. On the other hand if Player 2 plays R, then Player 1 chooses D. So, (U, L) and (D, R) are two subgame perfect equilibria. If Player 2 randomizes $\alpha L + (1-\alpha)R$. Player 1 gets 0 by playing U and $1-2\alpha$ by playing D.

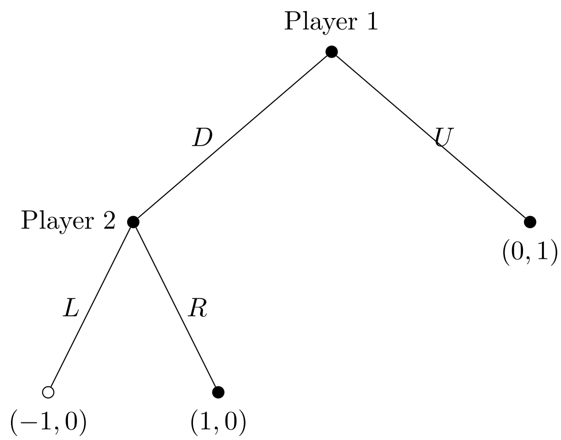

Figure 14: Backward induction with indifference

If $\alpha > \frac{1}{2}$, then Player 1 playing U is optimal. If $\alpha < \frac{1}{2}$, then Player 1 playing D is optimal. If $\alpha = \frac{1}{2}$, then Player 1 randomizing $\beta L + (1-\beta)D$ for any $\beta \in [0,1]$ is optimal. All these correspond to subgame perfect equilibria of this game.

Infinite horizon and action sets. There are extensive games where the number of stages is infinite. For such games, the process of backward induction is not defined. However, the notion of subgame perfect equilibrium is still well defined. We need to consider subgames, and the strategies should consist of equilibrium behavior in each subgame.

Another important remark is that with finite number of stages, backward induction is well defined even if agents have infinite set of actions in a decision vertex. However, the optimal response may be empty with infinite set of actions. So, wherever the optimal response map is non-empty, we can easily define the backward induction process. The following application illustrates this point clearly.

16.1 Alternative Offers Bargaining

We now visit an application of subgame perfect equilibrium. In this problem, two players are bargaining over 1 unit of money. They will bargain for $T+1$ periods starting from period 0. In even periods (starting at 0), Player 1 offers a split $(o_t, 1-o_t)$, where $o_t \in [0,1]$ is Player 1's share. If Player 2 accepts, the game ends. Else, we move to the next period. In odd periods, Player 2 offers a split. If no split is accepted at the end of period T, then the game ends with each player getting 0. Money received in period t is discounted by δ^t, where $\delta \in (0,1)$.

This game has perfect information, finite number of stages, but infinite set of actions at each decision vertex. There are many tied utilities too. But surprisingly, it has a unique subgame perfect equilibrium.

To understand the game better, consider just a one-period $T = 1$ case. Player 1 offers a split $(o_1, 1 - o_1)$ and Player 2 can either accept or reject. In all the decision vertices, where Player 2 gets a positive offer, he accepts. In the decision vertex where Player 2 gets zero offer, he is indifferent. Knowing this, we now apply backward induction on Player 1. Player 1's optimal is not clearly to give a positive split to Player 2 because that is dominated. If Player 2 rejects a zero offer with positive probability y, then Player 1 gets a payoff of $1 - y$, which is dominated by Player 1 offering $(1 - \frac{y}{2}, \frac{y}{2})$. Hence, again Player 2 rejecting a zero offer with positive probability and accepting a positive offer implies Player 1 has *no* optimal action at his decision vertex. Hence, the backward induction procedure does not provide any strategy of Player 1 for such a strategy of Player 2. On the other hand, if Player 2 accepts Player 1's zero offer with probability 1, then Player 1's optimal action is to offer $(1, 0)$. This will be a subgame perfect equilibrium. This forms the basis of the theorem below.

THEOREM **19** *In the alternative offers bargaining game, there is a unique subgame perfect equilibrium, where the initial offer is accepted. As $T \to \infty$, the equilibrium payoffs converge to $(\frac{1}{1+\delta}, \frac{\delta}{1+\delta})$.*

Proof: Suppose T is even. Then, in the last period, Player 1 offers. Consider the subgame from this period. It consists of a decision vertex for Player 1 where he offers a split $(o_T, 1-o_T)$ and a decision vertex for Player 2 for each offer of Player 1. In the decision vertex, Player 2 must accept any positive offer. But it can accept, reject, or randomize on zero offer. Then, consider the offer of Player 1. Player 1 cannot offer positive amount to Player 2 since he can improve it by giving half of that - hence, there is a one-shot deviation. So, Player 1 must offer 0 amount to Player 2. Now, if Player 2 rejects such an offer, then both get zero. Hence, if Player 2 randomizes with α probability reject and $(1 - \alpha)$ probability accept, then Player 1 offering 0 gets a payoff of $(1 - \alpha)\delta^T$. But Player 1 can do better by offering Player 2 an amount $\frac{1}{2}\alpha$ (which Player 2 will accept). Hence, if Player 2 rejects with positive probability, then offering 0 is not a best response of Player 1. So, offering 0 and getting rejected with some probability is not a subgame perfect equilibrium. Thus, offering 0 and accepting 0 is the unique subgame perfect equilibrium outcome from period T.

We now repeat this idea. Essentially, at each subgame an offer must be made such that the opponent is indifferent between accepting and rejecting and the opponent must accept. By backward induction, we proceed as follows.

1. In period T, Player 1 offers $(1, 0)$, which Player 2 accepts. Resulting payoffs are $(\delta^T, 0)$.

2. In period $(T-1)$, Player 1 can assure himself of δ^T. So, he accepts any offer giving him at least δ^T. So, Player 2 offers $(\delta, 1-\delta)$ which gives payoff $(\delta^T, \delta^{T-1}-\delta^T)$.

3. In period $(T-2)$, Player 2 can assure himself of $\delta^{T-1}-\delta^T$. So, Player 1 offers $(1-\delta+\delta^2, \delta-\delta^2)$, which gives payoff $(\delta^{T-2}-\delta^{T-1}+\delta^T, \delta^{T-1}-\delta^T)$.

   Continuing in this manner, we get

4. In period 0, Player 1 offers $(1-\delta+\delta^2-\ldots+\delta^T, \delta-\delta^2+\ldots-\delta^T) \equiv (\frac{1+\delta^{T+1}}{(1+\delta)}, \frac{\delta-\delta^{T+1}}{(1+\delta)})$, which is accepted by Player 2. Note that the limit of $T \to \infty$ is $(\frac{1}{1+\delta}, \frac{\delta}{(1+\delta)})$.

If T is odd, a similar analysis yields an offer by Player 1 equal to $(\frac{1-\delta^{T+1}}{(1+\delta)}, \frac{\delta+\delta^{T+1}}{(1+\delta)})$, whose limit $T \to \infty$ is also $(\frac{1}{1+\delta}, \frac{\delta}{(1+\delta)})$. ■

17 Games with Imperfect Information

In games with imperfect information a player may not observe the entire history at every decision vertex. Hence, when he reaches his decision vertex, there is uncertainty about which decision vertex he is really in. To make complete sense of this uncertainty, the set of actions available at each of these uncertain decision vertices must be same. This idea is captured by the notion of an information set. Consider the following examples given below.

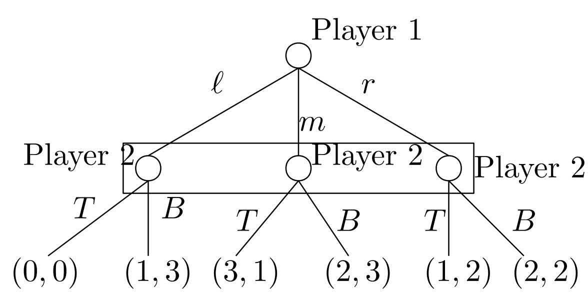

Figure 15: Strategic form game as an extensive form game

1. Strategic form games. Every strategic form game can be represented as an extensive form game of imperfect information. To see this consider a strategic form game of two players: $N = \{1, 2\}$. In the strategic form game, each player $i \in N$ chooses an action from his strategy S_i simultaneously. So, think of an extensive form game, where one of the players, say 1, moves first. However, the action of Player 1 is not observed by Player 2. This can be depicted by an extensive form game. Suppose $S_1 = \{\ell, m, r\}$ and $S_2 = \{T, B\}$. Then, the game is shown in Figure 15.

Notice that when Player 2 takes her action in Figure 15, she does not know which decision vertex she is in - so her three decision vertices are bundled in one *information set*.

2. Bayesian games. In Bayesian games, there is a clear sequential nature of play. First, Nature draws the type of each player, but informs them *privately*. Hence, the action of Nature's move is observed to corresponding players only. We illustrate this with a simpler version of bilateral trading, where seller's cost c is known to both the buyer and the seller. At the beginning, buyer's value v is drawn from $\{v_L, v_H\}$ with probabilities π_L and π_H respectively. Then, the seller announces one of two prices: p_L and p_H. The buyer observes the prices and chooses either to accept or reject the offer. If accepted, trade happens at the announced price of the seller. Else, no trade happens.

   This game is shown as a game of imperfect information in Figure 16. Here, the imperfect information is generated by the private nature of Nature's move. Since the seller does not know the type of the buyer, he does not know which decision vertex he is in when he announces a price. Hence, his decision vertices are bundled in one information set.

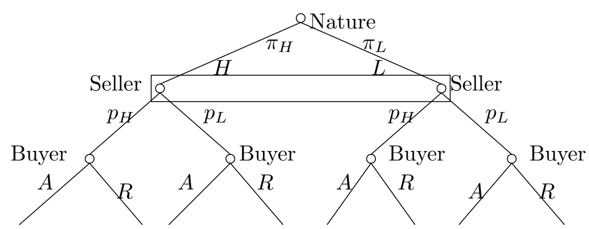

Figure 16: Bilateral trading (one-sided asymmetry) as an extensive form game

The idea of an information set is formalized below.

Definition **25** *In an extensive form game the information set of Player i is a non-empty subset $U_i \subseteq V_i$ and a subset of actions $A(U_i)$, such that at each $x \in U_i$ we have $A(x) = A(U_i)$.*

The only additional information in an extensive form game with imperfect information is a specification of information sets. In particular, for every player i, we specify a partition $\{U_i^j\}_j$ of the decision vertices V_i of Player i, where each U_i^j is an information set. Now, set of actions are specified for each information set. Another important specification is that we allow for moves by a player, who we denote by 0, called *Nature*. So, there will be a subset of decision vertices V_0, where Nature takes some actions. The probability of these actions are specified and known to all players in the game - Nature is not strategic.

Formally, an extensive form game of imperfect information can be defined similar to a game of perfect information with some minor modifications given as follows.

Definition 26 *An* **extensive form game of imperfect information** *is*

$$\Gamma \equiv (N, V, E, r, \{V_i\}_{i\in N\cup\{0\}}, \{U_i^j\}_{i\in N}^j, \{A(U_i^j)\}_{i\in N}^j, \{p_x\}_{x\in V_0}, \{A(x)\}_{x\in V_0}, \{u_i\}_{i\in N}),$$

where

- $\{U_i^j\}_{i\in N}^j$ *is a partition of* V_i *for each Player* $i \in N$*,*
- $A(U_i^j)$ *specifies the actions available at each information set* U_i^j *for Player* i*,*
- p_x *specifies a probability distribution at each of Nature's decision vertex* $x \in V_0$ *over his set of actions* $A(x)$*.*

Note that if every information set contains a single vertex, then the game is of perfect information.

The strategy and the idea of subgame is suitably changed in a game of imperfect information. Since the player is unsure about the vertex he has reached in an information set, his strategy must specify an action at every information set. We will denote by $\mathcal{U}_i \equiv \{U_i^1, \ldots, U_i^k\}$ the collection of information sets of Player i.

Formally, a strategy of player $i \in N$ is a map $s_i : \mathcal{U}_i \to \cup_{U_i^j\in\mathcal{U}_i} A(U_i^j)$ such that $s_i(U_i^j) \in A(U_i^j)$ for all $U_i^j \in \mathcal{U}_i$.

In the game in Figure 7, each player's information set is a singleton, except for Player 2, who has a single information set with two vertices. His strategy must specify what he will do at this information set.

The definition of a subgame is just the subtree starting from a decision vertex. If the game is of imperfect information, we need to worry about information sets. In particular, when we consider a subtree, for every Player and every information set of this player, all the vertices of this information set either belongs to the subtree or does not intersect with the subtree. So, $\Gamma(x)$ will be a subgame if for every $i \in N$ and for every $U_i^j \in \mathcal{U}_i$ either U_i^j lies in the subtree in $\Gamma(x)$ or it has an empty intersection with the subtree in $\Gamma(x)$.

The game in Figure 7 has only one subgame, i.e., the game itself. This is because every other subgame will only have part of the information set of Player 2.

17.1 Perfect Recall

Consider the following game in Figure 17. Player 2 is forgetful here. He forgets whether he had called Player 1 or not earlier. As a result, when Player 1 reaches his home, he does not know whether Player 2 has come because of his call or without his call. Thus, Player 2 has an information set consisting of two decision vertices.

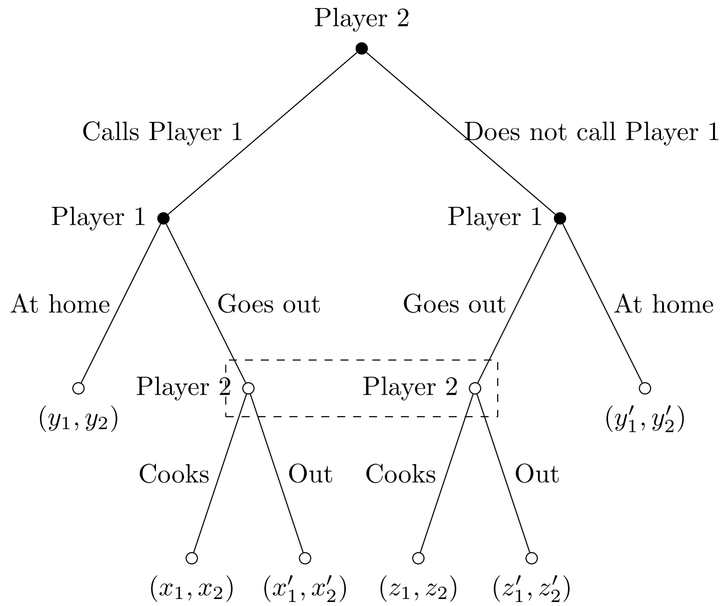

Figure 17: Extensive form game without perfect recall

Games in which players remember the entire sequence of information (history) from root to their every information set are players with **perfect recall**. Formally, Player i has perfect recall if at every information set U_i^j and every pair of vertices $x, x' \in U_i^j$, the information observed by Player i to reach x and x' from root are identical. An extensive form game in which all the players have perfect recall is called a game with perfect recall. We will exclusively focus attention on games in which all the players have perfect recall.

18 Equilibria for Games of Imperfect Information

In games where there is imperfect information, subgame perfect equilibrium can still be applied but backward induction is not well-defined in such games. Moreover, subgame perfect equilibrium may be a useless solution concept in which there is imperfect information. To see this, consider the game in Figure 18. This game has only one subgame. Hence, the set of Nash equilibria are equivalent to the set of subgame perfect equilibria. The problem with subgame perfect equilibrium in this game is that it does not use any *beliefs* of Player 2. As a result, it puts no restriction on his optimal choice when his information set is reached. To appropriately define behavior in information sets, any equilibrium must also define beliefs and equilibrium choices must be consistent with these beliefs. This is the basic idea behind defining equilibrium refinements in games of imperfect information.

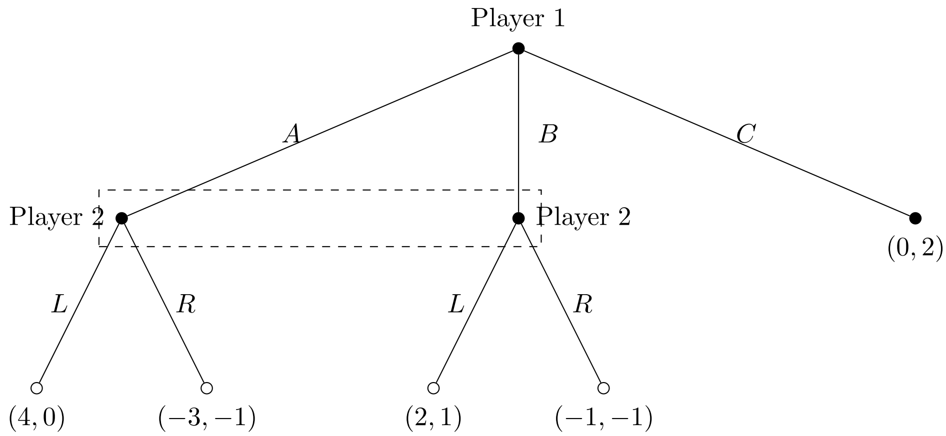

Figure 18: Imperfect Information

18.1 Perfect Bayesian Equilibrium

To understand the problem with subgame perfect equilibrium further in such games, consider the reduced-form strategic-form game of the game in Figure 18. It is shown in Table 27.

	L	R
A	$(4, 0)$	$(-3, -1)$
B	$(2, 1)$	$(-1, -1)$
C	$(0, 2)$	$(0, 2)$

Table 27: Reduced strategic form of the game in Figure 10

The Nash equilbria of this strategic-form game consists of (A, L), $(C, \alpha L + (1-\alpha)R)$, where $\alpha \leq \frac{1}{3}$. The idea of *sequential rationality* requires that each player must behave rationally once his information set is reached. To be able to do this, players must form beliefs about where they are inside their information set, and act optimally according to this belief. The nature of beliefs that is permissible results in different solution concepts.

For instance, if we specify a strategy profile, where Player 1 plays A with probability $\frac{1}{3}$ and B with probability $\frac{1}{2}$, then this equilibrium knowledge is enough to pin down the beliefs of Player 2. Remember, that Player 2 has correct belief about equilibrium behavior of Player 1. Hence, his belief of the information set can be deduced from this: total probability of reaching this information set is $\frac{5}{6}$, and individual conditional probabilities are $(\frac{2}{5}, \frac{3}{5})$. Of course, here we cannot apply this principle if a strategy profile does not reach a particular information set since conditional probabilities are not defined at those information sets. So, sequential rational behavior can be with respect to *any* belief at such information sets.

Formally, in an extensive form game with imperfect information, the belief of Player i is a map $\mu_i^j : U_i^j \to [0,1]$ for each j such that $\sum_{x \in U_i^j} \mu_i^j(x) = 1$ for all j. We write the collection of beliefs of Player i as μ_i: this specifies a probability distribution for each of his information sets.

Given a strategy profile σ, we can compute the probability with which each decision vertex is reached in an extensive form game. We denote this as $P_\sigma(x)$. The probability with with an information set U_i^j is reached given σ is $P_\sigma(U_i^j) = \sum_{x \in U_i^j} P_\sigma(x)$.

DEFINITION **27** *Belief μ_i of Player i is* **Bayesian** *given a strategy profile σ if for every information set U_i^j reached with positive probability in the strategy profile σ, we have for all $x \in U_i^j$,*

$$\mu_i^j(x) = \frac{P_\sigma(x)}{P_\sigma(U_i^j)}.$$

Sequential rationality now extends to this setting as follows.

DEFINITION **28** *A strategy σ_i of Player i at information set U_i^j is* **sequentially rational** *given strategies σ_{-i} and beliefs μ_i if for all σ'_i, we have*

$$\sum_{x \in U_i^j} \mu_i^j(x) u_i(\sigma_i, \sigma_{-i}|x) \geq \sum_{x \in U_i^j} \mu_i^j(x) u_i(\sigma'_i, \sigma_{-i}|x).$$

A strategy σ_i of Player i is **sequentially rational** *given σ_{-i} and μ_i if it is sequentially rational at all information sets.*

An equilibrium here in an imperfect information extensive form game involves specifying **strategies and beliefs**. Beliefs have to be consistent in the form of Bayesian and strategies have to be sequentially rational. The pair of strategy profile and belief profile is called an *assessment*.

DEFINITION **29** *An assessment (σ, μ) is a* **perfect Bayesian equilibrium (PBE)** *if for every Player i*

- *μ_i is Bayesian given σ*

- *σ_i is sequentially rational given σ_{-i} and μ_i.*

In the game in Figure 18, for every belief of Player 2, L is a strictly dominant action. Given this, Player 1 must play A irrespective of his beliefs. Hence, the unique PBE of this game is $(A, L, \mu_2(B) = 1)$. In general, a PBE does not allow players to play a strictly dominated action, while a Nash equilibrium does not preclude this off equilibrium path. A fact that we do not prove here but state is: every PBE is a Nash equilibrium.

18.2 Sequential Equilibrium

However, PBE allows for any arbitrary beliefs off equilibrium path. This can lead to unsatisfactory predictions in certain games. The following example illustrates this. Consider the game in Figure 19. In this game, what beliefs of Player 2 induce him to play ℓ? Suppose he puts μ probability on his left decision vertex and $(1-\mu)$ on the other. Then, his payoff by playing ℓ is $2-\mu$ and his payoff from playing r is $3-4\mu$. So he plays ℓ if $\mu > \frac{1}{3}$, r if $\mu < \frac{1}{3}$, and mixes ℓ and r otherwise. But Player 1 plays his dominant strategy D in his second information set. So, what should Player 1 play in PBE in the first information set? Suppose he mixes $\alpha L + (1-\alpha)R$, where $\alpha > 0$. Then, $\mu = 1$ is the only Bayesian belief - note this information set is reached in equilibrium now. Then Player 2 must play ℓ. This means that $\alpha = 1$. If Player 1 plays R, then any belief is allowed for Player 2. But for Player 1 to choose R in equilibrium, Player 2 must play r - if he plays ℓ, then he is better of choosing L and then D to get payoff 2. For Player 2 to play r, the belief should be $\mu \leq \frac{1}{3}$. There are other PBE where Player 2 mixes also.

Now, let us consider the PBE $((R, D), r; \mu \leq \frac{1}{3})$. It is not reasonable to assume that Player 2 plays r in his information set since he knows that U is never played by Player 1. Another amazing feature of this game is its subgame perfect equilibrium. The subgame starting with the second information set of Player 1 has one Nash equilibrium - Player 1 chooses his dominant strategy D and Player 2 best responds with ℓ. Given this, Player 1 chooses L in the first information set. Hence, $((L, D), \ell)$ is a unique subgame perfect equilibrium of this game. Thus, the PBE is *not* a refinement of subgame perfect equilibrium.

To get rid of this unpleasant feature of PBE, a refinement is proposed. The refinement aims to put some consistent beliefs on information sets that are not reached in equilibrium.

Definition **30** *An assessment* (σ, μ) *is a* **sequential equilibrium** *if*

1. μ *is* **consistent** *given* σ*: There exists a sequence of completely mixed strategy profile* $\{\sigma^k\}_k$ *such that (i)* $\lim_k \sigma^k = \sigma$ *and if* μ^k *are the unique Bayesian beliefs for* σ^k*, then* $\lim_k \mu^k = \mu$.

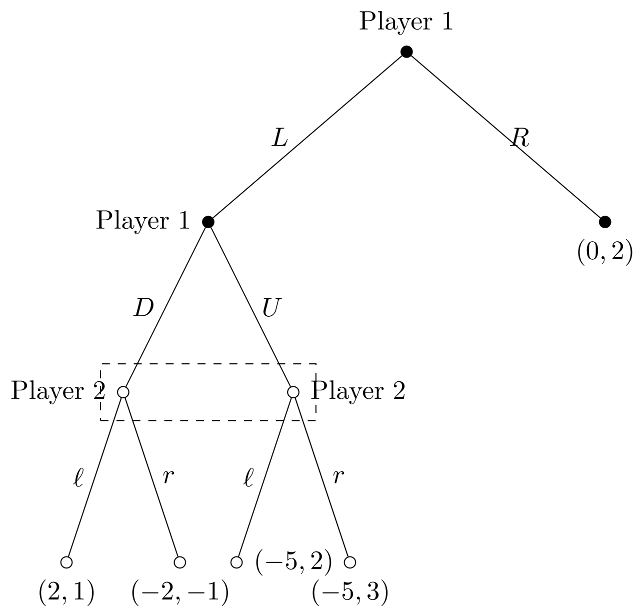

Figure 19: Problems with PBE

2. σ is sequentially rational given μ.

The new condition here from PBE is consistency, which requires that if Players make some small mistakes from equilibrium, the beliefs should be close to the Bayesian beliefs corresponding to those small mistakes. Note that the sequence we construct need not be unique, and different sequences may lead to different beliefs.

The following proposition says that every sequential equilibrium is also a perfect Bayesian equilibrium.

PROPOSITION **1** *If μ is consistent given σ, it is Bayesian given σ. Hence, every sequential equilibrium is also a perfect Bayesian equilibrium.*

Proof: VERY INFORMAL. For this, we pick an information set U_i^j of Player i which is reached with positive probability in σ. Bayesian belief says that for every $x \in U_i^j$,

$$\mu_i^j(x) = \frac{P_\sigma(x)}{P_\sigma(U_i^j)}.$$

Any perturbation σ^ϵ, will generate a belief μ^ϵ, which is computed by computing $P_{\sigma^\epsilon}(x)$ and $P_{\sigma^\epsilon}(U_i^j)$. As the perturbations approach zero, $P_{\sigma^\epsilon}(x)$ and $P_{\sigma^\epsilon}(U_i^j)$ approach $P_\sigma(x)$ and $P_\sigma(U_i^j)$ respectively - this happens because σ^ϵ approaches σ and the linear way in which probabilities are computed. So, as long as $P_\sigma(U_i^j)$ is non-zero, these limits give you $\mu_i^j(x)$. ■

In extensive form games with imperfect information, the one-shot deviation principle continues to hold. Hence, in such games, it is enough to check for deviations at one information set at a time.

The following theorem, whose proof we skip, establishes that a sequential equilibrium is refinement of subgame perfect equilibrium.

THEOREM **20** *Every sequential equilibrium is a subgame perfect equilibrium. Every completely mixed strategy Nash equilibrium is a sequential equilibrium.*

The second part of Theorem 20 follows trivially by taking the sequence of strategies same as the equilibrium strategy.

Let us now revisit the game in Figure 19. First, look at the subgame perfect equilibrium $((L, D), \ell)$. If we consider mixed strategies, where $\sigma_1^k(R) = \epsilon_R^k, \sigma_1^k(L) = 1 - \epsilon_R^k$ and $\sigma_1^k(D) = 1 - \epsilon_D^k, \sigma_1^k(U) = \epsilon_D^k$. Then,

$$\mu = \frac{(1 - \epsilon_k^D)(1 - \epsilon_R^k)}{1 - \epsilon_R^k} \to 1.$$

Note that perturbation of Player 2's strategy is not necessary here. Hence, $\mu = 1$ is a consistent belief given this strategy profile. We already know that this strategy profile is sequentially rational given μ. Hence, it is a sequential equilibrium.

Now, can there be a sequential equilibrium where Player 1 chooses (R, D) and Player 2 chooses r. If we perturb the strategies of Player 1, then we reach the information set of Player 2 with positive probability where the belief on the (L, D) decision vertex must be very high. As a result, Player 2 must choose ℓ here to be sequentially rational. Hence, no sequential equilibrium will choose Player 2 playing r with positive probability if Player 1 plays (R, D).

A comment about existence of PBE and sequential equilibrium is that if games have perfect recall, then these equilibria always exist.

18.3 EXAMPLE: A SIGNALING GAME

We give an example to illustrate the notions of PBE and sequential equilibrium. This example is usually called a simpler version of the *signaling game*. There are two agents in this example - see Figure 20. Agent 1 has two types - High or Low, their probabilities are as shown in Figure 20. Agent 1's type is not observed by Agent 2 but his action, which is either N or E, is observable by Agent 2. After observing Agent 1's action, Agent 2 takes an action, which is either U or D. The payoffs are as shown in Figure 20.

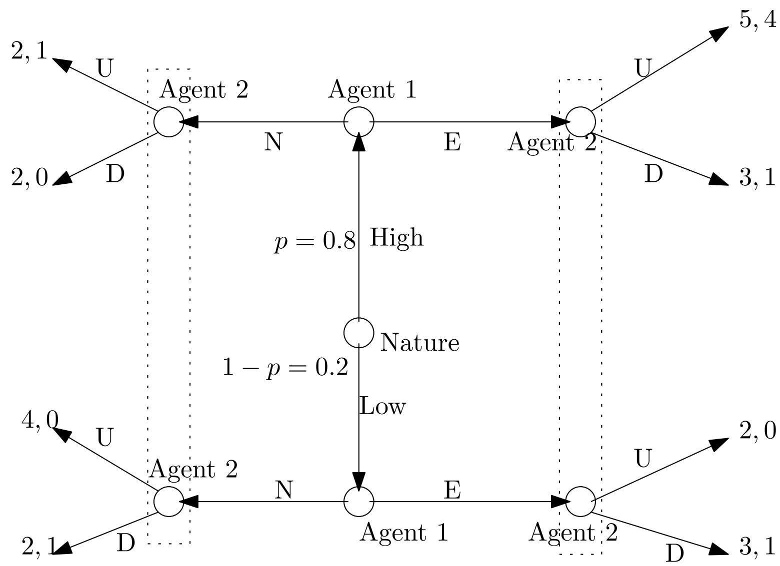

Figure 20: Signaling game

We now compute some of the PBE of this game. Before doing so, we observe that Agent 1 of type High strictly prefers E to N. Hence, in any PBE, Agent 1 must choose E at his decision vertex corresponding to High type. We now look at various PBE of this game. Denote the belief of Agent 2 on his left information set as μ_L for the top decision vertex and $1 - \mu_L$ for the bottom decision vertex. Similarly, denote the belief of Agent 2 on his right information set as μ_R for the top decision vertex and $1 - \mu_R$ for the bottom decision vertex.

- **Separating PBE.** High type Agent 1 chooses E but Low type Agent 1 chooses N. If such a PBE exists, then all the information sets of Agent 2 is reached in equilibrium. By Bayesian rationality, Agent 2's belief must satisfy: $\mu_L = 0, \mu_R = 1$. Then, sequential rationality of Agent 2 implies that he must choose D in the left information set and U in the right information set. Finally, we verify that Agent 1 is sequentially rational. As argued, the High type choosing E is sequentially rational. For the Low type, choosing N gives a payoff of 2 and choosing E gives a payoff of 2 also. Hence, Agent 1's strategy is sequentially rational. So, we can describe the separating PBE as:

$$(High : E, Low : N, Left : D, Right : U, \mu_L = 0, \mu_R = 1).$$

  This PBE is trivially a sequential equilibrium since every information set is reached with positive probability in this equilibrium.

- **Pooling PBE.** Both High and Low type Agent 1 choose E. If such a PBE exists, then left information set of Agent 2 is not reached in equilibrium and right information set is reached with probability 1. By Bayesian rationality, Agent 2's belief in right information set must be: $\mu_R = p = 0.8$. Then, sequential rationality of Agent 2 in the right information set implies he must choose U: choosing U gives a payoff equal to 0.8(4) compared to a payoff of 1 by choosing D. For Agent 1 to choose N when he is of Low type, Agent 2 must choose D - this is because if Agent 2 chooses U, then Agent 1 is better off choosing N when he is of Low type. So, sequential rationality of Low type Agent 1 forces Agent 2 to choose D in his left information set. But such a choice is possible with sequential rationality if $1 - \mu_L \geq \mu_L$ or $\mu_L \leq 0.5$.

Hence, there is a class of pooling PBE:

$$(High : E, Low : E, Left : D, Right : U, \mu_L \leq 0.5, \mu_R = p = 0.8).$$

Any such PBE is also a sequential equilibrium. Fix a particular PBE with a particular value of μ_L. For this, we think of a perturbation of Agent 1's actions to reach the left information set of Agent 2. But this perturbation must generate beliefs μ_L in the limit. A possible way to generate this belief is to choose perturbations as follows:

$$High : \epsilon' N + (1 - \epsilon')E; Low : \epsilon N + (1 - \epsilon)E,$$

where $\epsilon' = \epsilon \frac{\mu_L}{4(1-\mu_L)}$. Notice that this choice of ϵ and ϵ' exactly generates μ_L belief by Bayesian rationality. Hence, as $\epsilon \to 0$ (and, hence, $\epsilon' \to 0$), we get the beliefs approaching μ_L.

- **Mixing at Low type.** High type Agent 1 chooses E but Low type agent mixes N and E. If such a PBE exists, then let Low type Agent 1 mixes as $\sigma_E E + (1 - \sigma_E)N$, where $\sigma_E \in (0, 1)$. As a result, all information sets of Agent 2 is reached in equilibrium. Bayesian rationality implies that

$$\mu_L = 0, \mu_R = \frac{0.8}{0.8 + 0.2\sigma_E}.$$

Then, sequential rationality of Agent 2 requires that he must choose D in the left information set. Sequential rationality of Agent 1 at Low type requires that he must be indifferent between N and E (because he mixes). This is only possible if Agent 2 chooses U at his right information set. But then, $4\mu_R \geq 1$ or $3.2 \geq 0.8 + 0.2\sigma_E$ or $\sigma_E \leq 1.2$, which is always true. Hence, independent of the mixing probability of Agent 1 of Low type, Agent 2 prefers U at his right information set. So, for any $\sigma_E \in (0, 1)$, we have the following PBE:

$$(High : E, Low : \sigma_E E + (1 - \sigma_E)N, Left : D, Right : U, \mu_L = 0, \mu_R = \frac{0.8}{0.8 + 0.2\sigma_E}).$$

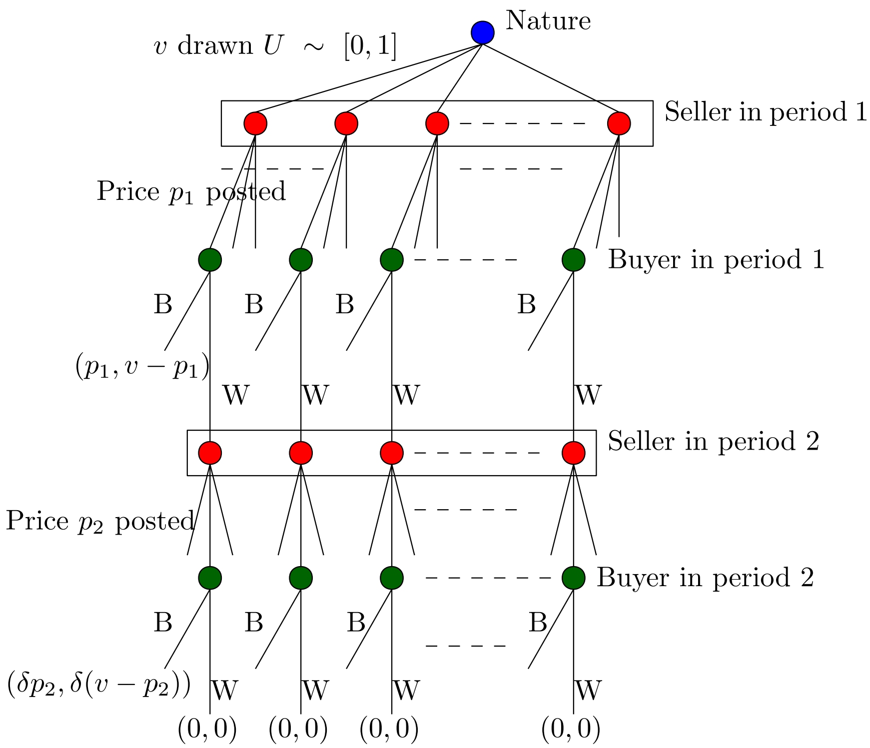

Figure 21: Sale across two periods

Since every information set is reached with positive probability in such PBE, they are also sequential equilibria.

18.4 Example: Two period sale

This example illustrates perfect Bayesian equilibrium in a game with infinite set of actions. A seller has an object to sell. She has zero value for the object, which is common knowledge. There is a single buyer, whose for the object is uniformly distributed in $[0, 1]$ - this is common knowledge. However, the value for the buyer is known privately to the buyer.

There are two periods. In the first period, the seller posts a price p_1 and the buyer chooses one of the actions: BUY (B) or WAIT (W). If the buyer chooses B, then the game ends with the seller getting a payoff of p_1 and the buyer of type v getting a payoff of $v - p_1$. If the buyer chooses W, then the game proceeds to period 2, where the seller posts a price p_2. The buyer can again choose one of the two actions: BUY (B) or WAIT (W). The game now ends. If the buyer chooses W, then both the players get a payoff of zero. But if the buyer chooses B, then the seller gets a payoff δp_2, whereas the buyer of type v gets a payoff of $\delta(v - p_2)$, where $\delta \in (0, 1)$ is a common discount factor. The game is shown in Figure 21.

What is a strategy for a player in this game?

Seller. The seller has exactly two information sets, one corresponding to each period. At each information set, he posts a price. So, his strategy in period 1 is $p_1 \in \mathbb{R}_+$ and in period 2 it is $p_2 \in \mathbb{R}_+$ having posted a price p_1 (note, in period 2, the seller has infinite number of information sets - one corresponding to each choice of p_1).

Buyer. The buyer has two sets of decision vertices: corresponding to period 1 and period 2. In period 1, his decision vertex depends on (a) his own type and (b) the price the seller posts. So, a period 1 decision vertex can be described by (v, p_1) and for every (v, p_1), the buyer either chooses B or W. In period 2, the decision vertex is characterized by (v, p_1, p_2), and then the buyer either chooses B or W.

We now solve for a perfect Bayesian equilibrium of this game in steps.

Bayesian rationality of seller in period 1. This just requires that his beliefs must be same as nature probabilities: probability that seller is at a decision vertex corresponding to a buyer of type less than or equal to x (note: continuous distribution of types) is x (due to uniform distribution).

Sequential rationality of buyer in period 2. Sequential rationality in period 2 for buyer implies that a buyer of type v must buy if $v > p_2$ and wait if $v < p_2$.

Sequential rationality of buyer in period 1. Consider a buyer in period 1 who has value v and sees price p_1. Given strategy of the seller, Bayesian rational belief of seller, and his own sequentially rational action in period 2, if he finds sequentially rational to buy, then every type $v' > v$ must also find it sequentially rational to buy at price p_1. To see this, fix the strategy of the seller as p_1 and p_2 given p_1. The payoff of a buyer of type p_1 by buying today is $v - p_1$ and waiting for next period is $\max(\delta(v - p_2), 0)$. If $v - p_1 > \max(\delta(v - p_2), 0)$, then for all $v' > v$, we also have $v' - p_1 > \max(\delta(v' - p_2), 0)$. Similarly, if $v - p_1 < \max(\delta(v - p_2), 0)$, then for all $v' < v$, we also have $v' - p_1 < \max(\delta(v' - p_2), 0)$. This suggests a **cutoff-action** to be optimal in period 1 for the buyer. For every price p_1, there is a cutoff value $v(p_1)$ such that all buyer types above it buy and all buyer types below it wait.

Bayesian rationality of seller in period 2. Given the strategy of the buyer, the seller in period 2 knows that only a buyer with value $v < v(p_1)$ will be active in period 2. Hence, the conditional probability of being at decision vertex where value of buyer is less than or equal to x (we compute cdf because there are infinite number of decision vertices in this information set) is given by (using conditional uniform distribution):

$$\frac{x}{v(p_1)}.$$

Further, such a buyer chooses BUY in period 2 if $v > p_2$. Hence, expected payoff of seller by setting a price p_2 in period 2 given a price p_1 in period 1 is given by

$$p_2 \frac{v(p_1) - p_2}{v(p_1)}.$$

SEQUENTIAL RATIONALITY OF SELLER IN PERIOD 2. Sequential rationality of seller in period 2 who has already posted a price p_1 is to maximize her expected payoff given his Bayesian rational beliefs. This leads to maximizing

$$p_2 \frac{v(p_1) - p_2}{v(p_1)},$$

over all $p_2 \in \mathbb{R}_+$. The maximum of this expression happens at

$$p_2 = \frac{1}{2} v(p_1).$$

SEQUENTIAL RATIONALITY OF BUYER IN PERIOD 1 (AGAIN). Having computed p_2 as a function of $v(p_1)$, we can now be more precise about buyer's action in period 1. We know that the **cutoff** type will be indifferent between BUY and WAIT in period 1. Hence,

$$v(p_1) - p_1 = \delta(v(p_1) - p_2) = \delta\big(v(p_1) - \frac{1}{2} v(p_1)\big) = \frac{\delta}{2} v(p_1).$$

This gives us

$$v(p_1) = \frac{1}{1 - \frac{\delta}{2}} p_1.$$

SEQUENTIAL RATIONALITY OF SELLER IN PERIOD 1. Finally, sequential rationality of the seller must require that the seller must maximize her expected payoff (given the strategy of buyer and his beliefs in period 2) in period 1. His expected payoff by posting a price p_1 is (denoting $1 - \frac{1}{2}\delta = K$ below):

$$\begin{aligned} p_1(1 - v(p_1)) + \delta p_2(v(p_1) - p_2) &= p_1(1 - v(p_1)) + \delta \frac{1}{2} v(p_1) \frac{1}{2} v(p_1) \\ &= p_1(1 - \frac{1}{K} p_1) + \frac{\delta}{4} \frac{1}{K^2} (p_1)^2. \end{aligned}$$

Taking the first order condition with respect to p_1 and setting it equal to zero, we get

$$1 - \frac{2}{K}p_1 + \frac{\delta}{4K^2}2p_1 = 0.$$

This gives us

$$p_1 = \frac{2K^2}{4K-\delta} = \frac{1}{2}\frac{(1-\frac{\delta}{2})^2}{(1-\frac{3\delta}{4})}.$$

This also gives us the complete specification of the equilibrium:

$$p_1 = \frac{1}{2}\frac{(1-\frac{\delta}{2})^2}{(1-\frac{3\delta}{4})}; v(p_1) = \frac{1}{2}\frac{(1-\frac{\delta}{2})}{(1-\frac{3\delta}{4})}; p_2 = \frac{1}{4}\frac{(1-\frac{\delta}{2})}{(1-\frac{3\delta}{4})};$$

supplemented by beliefs for seller 1: in period 1 information set, her belief is the same as Nature's probability; in period 2 information set, her belief of being at a vertex corresponding to buyer type less than or equal to x and price p_1 is 0 if $x > v(p_1)$ and $\frac{v(p_1)-x}{v(p_1)}$ if $x < v(p_1)$.

For a class of $\delta \in (0,1)$, $v(p_1) > 0$ if $p_1 > 0$. So, if $p_1 > 0$, then every buyer with value less than $v(p_1)$ will reach the the second period information set. If $p_1 = 0$, then $v(p_1) = 0$, then the only buyer who reaches the second period information set are zero value buyer. But if we consider a perturbation of this strategy, where buyer with value value v BUYS with probability $1-\epsilon$ and WAITS with probability ϵ if $v \geq v(p)$ and WAITS with probability $1-\epsilon$ and BUYS with probability ϵ otherwise, then we reach **all** information sets with positive measure probability. In particular, even at $p_1 = 0$, we may have some buyers with probability greater than zero with positive probability. The limit of the beliefs induced by these strategies can be shown to be the beliefs induced by cutoff strategies. This in turn will show sequential equilibrium.

19 Repeated Games

19.1 Basic Ideas - The Repeated Prisoner's Dilemma

Consider the Prisoners' Dilemma (PD) game in Table 28. Recall that a dominant strategy equilibrium of this game is (L_1, L_2), and it is the unique Nash equilibrium of the game.

	L_2	R_2
L_1	2,2	6,1
R_1	1,6	5,5

Table 28: Prisoner's Dilemma

Now, suppose the game is played twice with the actions at the end of every stage is observed by all the players, and the payoff of a player at the end of the game is the sum of payoff at the end of each stage. The game can be represented in extensive form now. A subgame perfect equilibrium of this extensive form game requires that the players play a Nash equilibrium in the second stage, and they play a Nash equilibrium of the entire game. Since the unique Nash equilibrium of the game is (L_1, L_2), the players will play (L_1, L_2) in second stage in any subgame pefect equilibrium. Given this, the players now know that they will get a payoff of 1 in the second stage. So, the we can add $(1, 1)$ to the payoff matrix in the first stage, and then compute a Nash equilibrium. This still gives a unique Nash equilibrium of (L_1, L_2). Hence, the outcome of this game in a subgame perfect equilibrium is (L_1, L_2).

This argument can be generalized. Let $G = (N, \{A_i\}_{i \in N}, \{u_i\}_{i \in N})$ denote a strategic-form game of complete information. The game G is called the **stage game** of the repeated game.

DEFINITION **31** *Given a stage game G, let $G(T)$ denote the* **finitely repeated game** *in which G is played T times with actions taken by of all players in the preceding stages observed before the play in the next stage, and payoffs of $G(T)$ are simply the sum of payoffs in all T stages.*

Our arguments earlier lead to the following proposition (without formally defining notions of equilibrium).

PROPOSITION **2** *If the stage game G has a unique Nash equilibrium, then for any finite repetition of G, the repeated game $G(T)$ has a unique subgame perfect outcome: the Nash equilibrium of the stage game G is played in every stage.*

There are two important assumptions here: (a) the stage game has a unique Nash equilibrium and (b) the stage game is repeated finite number of times. We will see that if either of the two assumptions are not present then it is possible for players to get better payoffs.

We now modify the PD game by introducing a new strategy for every player. The new PD game is shown in Table 29. There are two Nash equilibria of this game: (L_1, L_2) and (R_1, R_2).

Now, suppose the stage game in Table 29 is repeated twice. Then, using the arguments earlier, we can say that in every stage playing either of the Nash equilibria is subgame perfect. But, we will show that there exists a subgame perfect equilibrium in which (M_1, M_2) is played in the first stage.

Consider the following strategy of the players: if (M_1, M_2) is played in the first stage, then play (R_1, R_2) in the second stage; if any other outcome happens in the first stage, then play

	L_1	M_1	R_1
L_2	1,1	5,0	0,0
M_2	0,5	4,4	0,0
R_2	0,0	0,0	3,3

Table 29: A Game with Multiple Nash Equilibrium

(L_1, L_2) in the second stage. This means, in the first stage of the game, the players are looking at a payoff table as in Table 30, where second stage payoff $(3,3)$ is added to (M_1, M_2) and second stage payoff $(1,1)$ is added to all other strategy profiles. The addition of different payoffs to different strategy profiles changes the equilibria of this game. Now, we have three pure strategy Nash equilibria in Table 30: (L_1, L_2), (M_1, M_2), and (R_1, R_2). Hence, $((M_1, M_2), (R_1, R_2))$ constitute a subgame perefect equilibrium of this repeated game. Thus, existence of multiple Nash equilibrium in the stage game allowed us to achieve cooperation in the fist stage of the game. Notice that (M_1, M_2) is not a Nash equilibrium of the stage game.

	L_1	M_1	R_1
L_2	2,2	6,1	1,1
M_2	1,6	7,7	1,1
R_2	1,1	1,1	4,4

Table 30: Analyzing Payoffs of First Stage

This is part of a general argument: if G is a static game of complete information with multiple Nash equilibria, there may be subgame perfect outcomes of the finitely repeated game $G(T)$ in which for any stage $t < T$, the outcome in stage t is not a Nash equilibrium.

19.2 A Formal Model of Infinitely Repeated Games

Let $G \equiv (N, \{A_i\}_{i \in N}, \{u_i\}_{i \in N})$ be a strategic form game. When we repeat such a stage game G, we will assume that players observe all the actions taken in each period. At any period, let a^t denote the action profile chosen by players. The sequence of actions profile $(a^1, \ldots, a^{t-1})$ that leads to current period will be called the history of period t.

An **infinitely repeated game** of G is defined by $G^\infty \equiv (G, H, \{u_i^*\}_{i \in N})$, where

- $H = \cup_{t=1}^{\infty} A^t$ are the set of all possible histories, with $A^1 \equiv \emptyset$ denoting the null history, A^t denoting the possible histories till period t, and A^∞ denoting all infinite length histories.

- $u_i^* : A^\infty \to \mathbb{R}_+$ for every $i \in N$ is a utility function that assigns every infinite history a payoff for Player i.

A history is terminal if and only if it is infinite. Note that an infinitely repeated game is a special type of infinite extensive game.

Strategies in a Repeated Game.

What is a strategy of a player in an infinitely repeated game? Remember, a strategy needs to assign an action for every *possible* situation. This means that we need to assign an action at every period for every possible history. Thus, strategy of Player i is a collection of infinite maps $\{s_i^t\}_{t=1}^\infty$, where

$$s_i^t : A^t \to A_i.$$

Since a strategy seems to be a really complicated (infinite) object here, it is difficult to imagine it. One easy way to think of a strategy is a machine (or automaton). The machine for Player i has the following components.

- A set Q_i of **states**.
- An element $q_i^0 \in Q_i$, indicating the initial state.
- A function $f_i : Q_i \to A_i$ that assigns an action to every state.
- A transition function $\tau_i : Q_i \times A \to Q_i$ that assigns a state for every state and every action profile.

States represent situations that Player i cares about. We give an example showing how a strategy in Prisoner's Dilemma can be modeled as a machine. The strategy we consider is called a **trigger** strategy. It chooses the cooperate action C as long as the history consists of all players choosing C. Else, it chooses D. We only care about two states here: whether everyone chosen C in the past or not. We will denote this as $\mathcal{C}$ and $\mathcal{D}$ respectively. Since we want to choose C in the first period, we set $q_i^0 := \mathcal{C}$. Now, $f_i(\mathcal{C}) = C$ and $f_i(\mathcal{D}) = D$. The transition function looks as follows:

$$\tau_i(\mathcal{C}, (C, C)) = \mathcal{C}, \tau(\mathcal{X}, (X, Y)) = \mathcal{D} \text{ if } (\mathcal{X}, (X, Y)) \neq (\mathcal{C}, (C, C)).$$

This is an example of a strategy which is relatively simple. Note that the number of states here is finite. As one can see that we can construct strategies that care about more number of states (possibly infinite). For our purposes, the kinds of strategies that we will use will require machines with finite state space.

Payoffs in Repeated Games.

Fix a strategy profile of players $s \equiv (s_1, \ldots, s_n)$. This strategy profile leads to outcomes in each stage/period. Denote by v_i^t, the payoff due to this strategy profile in period t. So, agent i has an infinite stream of payoffs $\{v_i^t\}_{t=1}^{\infty}$ from this strategy profile. Similarly, if there is another strategy profile s', then it will generate an infinite stream of payoffs $\{w_i^t\}_{t=1}^{\infty}$. As a result, if Player i has to compare outcomes of two strategy profiles, it compares two infinite streams of payoffs: $\{v_i^t\}_{t=1}^{\infty}$ and $\{w_i^t\}_{t=1}^{\infty}$.

There are many ways to make this comparison. The most standard way is to use a *discounted criterion*. In this way, we have a discount factor $\delta \in (0,1)$ which is same for all the players. Player i attaches a payoff equal to

$$\sum_{t=1}^{\infty} \delta^{t-1} v_i^t,$$

to the payoff stream $\{v_i^t\}_{t=1}^{\infty}$. For instance, if there is a payoff stream that generates payoffs $v \equiv (1,1,1,\ldots)$, then the payoff from this stream is $1(1+\delta+\delta^2+\ldots) = \frac{1}{1-\delta}$. Note that even though the payoff is 1 in each period, we get a higher payoff overall. It is often convenient to assign a payoff of

$$(1-\delta)\sum_{t=1}^{\infty} \delta^{t-1} v_i^t,$$

to the payoff stream $\{v_i^t\}_{t=1}^{\infty}$. This normalizes the payoff and makes it easy to compare it with the stage game payoff. Note that comparisons across two infinite stream of payoffs still remain the same.

Obviously, discounting puts different weights on payoffs of different periods. Particularly, future is valued less than present. Note that changes in payoff in a single period may matter in the discounting criteria. To see this, compare $v \equiv (1,1,\ldots)$ and $w \equiv (1+\epsilon, 1-\epsilon, 1-\epsilon, \ldots)$, where $\epsilon \in (0,1)$. Payoff from v is 1 and payoff from w is $(1+\epsilon)(1-\delta)+(1-\epsilon)\delta = 1+\epsilon-2\epsilon\delta = 1+\epsilon(1-2\delta)$. This is greater than 1 if and only if $\delta > \frac{1}{2}$. [4]

Similarly, look at the payoff streams $v \equiv (1,-1,0,0,\ldots)$ and $w \equiv (0,0,0,\ldots)$. The payoff from w is zero but the payoff from v is $(1-\delta)^2$. Hence, for any $\delta \in (0,1)$, v is

[4]Sometimes, discounting is interpreted differently. A discount δ means that the stage game continues to next period with probability δ.

preferred to w. However, consider the stream $v' \equiv (-1, 1, 0, 0, \ldots)$. This generates a payoff of $(1-\delta)(-1+\delta) = -(1-\delta)^2$. Hence, v' is worse than w. This shows that the discounting puts more emphasis on current payoffs than future payoffs.

This is contrasted in the following two streams of payoffs $v \equiv (0, 0, 0, \ldots, 1, 1, 1, \ldots)$ and $w \equiv (1, 0, 0, \ldots)$. The payoff stream v has M zeros and then all 1s. The payoff from v is δ^M and from w is $(1-\delta)$. For every δ, there is a M such that w is preferred to v. But for a fixed M, we can find δ close to 1 such that v is preferred to w.

Given a strategy profile, $s \equiv (s_1, \ldots, s_n)$, we get a unique stream of action profiles $\{a^t\}_{t=1}^{\infty}$ associated with this strategy profile. Note how this action profile is obtained - first, each player i plays $a_i^1 := s_1^1(\emptyset)$. Having generated the action profiles $h^t \equiv (a^1, \ldots, a^{t-1})$, player i plays $a_i^t \equiv s_i^t(h^t)$. From this, we can compute the utility of Player i as

$$u_i^*(s) := (1-\delta) \sum_{t=1}^{\infty} \delta^{t-1} u_i(a^t).$$

Having defined strategies and payoffs, we are now ready to define the equilibrium concepts for repeated games.

DEFINITION **32** *A strategy profile $s \equiv (s_1, \ldots, s_n)$ is a* **Nash equilibrium** *of the infinitely repeated game G^∞ if for every $i \in N$, for every s_i', we have*

$$u_i^*(s_i, s_{-i}) \geq u_i^*(s_i', s_{-i}).$$

A strategy profile s is a **subgame perfect equilibrium** *if its restriction from any period t is a Nash equilibrium of the subgame starting from that period.*

19.3 Folk Theorems: Illustrations

There are two interesting take-aways from the results of repeated games. First, repeated games allow for a large set of payoffs to be achieved in Nash and subgame perfect equilibrium. Such theorems are called Folk Theorems. The second take-away is the kind of strategies that support such equilibrium payoffs. Such strategies are very common in many social interactions. To be able to establish folk theorems using such common real-life strategies give a strong foundation for such results.

We will now illustrate the basic idea behind the folk theorems using the Prisoner's Dilemma example - see Table 31. We first show that there are subgame perfect equilibria where cooperation can be achieved.

	L_2	R_2
L_1	1,1	-1,2
R_1	2,-1	0,0

Table 31: Prisoner's Dilemma

PROPOSITION **3** *Suppose $\delta \geq \frac{1}{2}$. Then, there is a subgame perfect equilibrium in the Prisoner's Dilemma game (Table 31), where both the players play (L_1, L_2) in every period.*

Proof: We describe the following strategy. Each player i follows L_i if the history consists of both players playing (L_1, L_2). If the history is different from (L_1, L_2) play in each period in the past, i plays R_i. The strategy stated here is called a *trigger strategy.* Fix Player 1 and assume that Player 2 is following the trigger strategy stated in the Proposition. We show that following the trigger strategy is optimal for Player 1. We need to consider two types of subgames.

CASE 1. We consider a subgame where the history so far has been (L_1, L_2). In that case, following L_1 gives Player 1 a payoff of 1. Playing R_1 in some periods has the following consequence. In the first period he plays R_1 he gets a payoff of 2 since Player 2 plays L_2. But in subsequent periods Player 2 plays R_2. So, he gets a maximum payoff of 0. As a result, his payoff is less than $(1-\delta)\big(1+\delta+\ldots+\delta^{t-1}+2\delta^t\big)$, where t is the first period from this subgame where he deviates. Remember the truthful payoff stream is $(1, 1, 1, \ldots)$. The deviated payoff stream payoff is less than the payoff stream $(1, 1, \ldots, 2, 0, 0, 0, \ldots)$. Then, it is sufficient to compare the payoff streams $(1, 1, 1, \ldots)$ and $(2, 0, 0, \ldots)$. The later one gives a payoff of $2(1-\delta)$. But $\delta \geq \frac{1}{2}$ implies that $1 \geq (1-\delta)2$. Hence, no deviation is profitable in this subgame.

CASE 2. We consider a subgame where the history involves action profiles other than (L_1, L_2). In that case, Player 2 is repeatedly playing R_2 in this subgame. But if Player 2 is playing R_2, Player 2 gets a payoff stream of $(0, 0, \ldots)$ by Playing R_1 in every period but gets a payoff stream where in every period he gets payoff less than or equal to 0 by playing some other strategy.

Hence, the specified strategy is a Nash equilibrium in this subgame. ■

19.4 Nash Folk Theorem

The trigger strategies used in Proposition 3 can be used to establish a general result about what payoffs can be achieved in a Nash equilibrium of G^∞.

The important payoff for folk theorems is the minmax value. Define the **minmax value** of player i in the stage game G as

$$\underline{v}_i = \min_{a_{-i}} \max_{a_i} u_i(a_i, a_{-i}),$$

where (a_i, a_{-i}) denotes an action profile of the stage game. [5] This is the minimum payoff player i can be held to by its opponents (using pure actions), given that he plays best response to the action profile a_{-i}. Let $u_i(\underline{a}_i, \underline{a}_{-i}) = \underline{v}_i$ for player i. Then, we call $\underline{a}^i = (\underline{a}_i, \underline{a}_{-i})$ the **minmax action profile** against player i. Notice that this includes an action for Player i also.

The reason minmax values are important is the following lemma.

Lemma **10** *Player i's payoff is at least $\underline{v}_i$ in any pure action Nash equilibrium of the stage game G and the infinitely repeated game G^∞, regardless of the value of δ.*

Proof: Let a be a Nash equilibrium of the stage game. Then for every $i \in N$,

$$u_i(a) = \max_{a_i} u_i(a_i, a_{-i}) \geq \min_{a'_{-i}} \max_{a_i} u_i(a_i, a'_{-i}) = \underline{v}_i.$$

Hence, Player i's payoff is at least $\underline{v}_i$ in any Nash equilibrium of the stage game.

Now, suppose player i plays a best response to the actions of other players in each period of G^∞. This guarantees him $\underline{v}_i$ in every period irrespective of the strategy played by other players. Hence, a player i is guaranteed of a payoff of $\underline{v}_i$ by this strategy in G^∞. So, any strategy that does not guarantee $\underline{v}_i$ will have a deviation where Player i just best responds to the actions of other players in every period. ■

Hence, Player i is guaranteed to get at least $\underline{v}_i$ payoff in any pure action Nash equilibrium of the repeated game.

Definition **33** *A payoff profile $v = (v_1, \ldots, v_n)$ is* **strictly enforceable** *if for every $i \in N$, we have $v_i > \underline{v}_i$.*

[5] The minmax and maxmin payoff of a player can be quite different. Please construct examples to see that the minmax is different from maxmin. In the early parts of the lectures, I used $\underline{v}_i$ to denote the maxmin payoff of Player i in a strategic form game, but here I use it for minmax payoff of Player i. I apologize for this confusion.

The minmax

We now give a weaker version of Folk Theorem.

THEOREM **21 (Pure Nash Folk Theorem)** *Suppose v is a strictly enforceable payoff profile and there exists an action profile a in the stage game G such that $u_i(a) = v_i$ for all $i \in N$. Then, there exists a $\underline{\delta}$, such that for all $\delta \geq \underline{\delta}$, there is a Nash equilibrium of G^∞ with discount δ where a is played in every period.*

Proof: Suppose v is a strictly enforceable feasible payoff profile and there exists an action profile a in the stage game G such that $u_i(a) = v_i$ for all $i \in N$. Consider the following strategy. It is described by three states: (a) normal state (b) i-punishment state, and (c) more-punishment state. The initial state is normal state. In normal state, the strategy recommends playing a_i to each Player i. Now, we inductively define the states at every history.

Consider Player j. If the state is normal and every player $i \in N$ plays a_i in a period, then the state remains normal in the next period. If the state is normal and a **unique** player $i \in N$ **does not play** a_i (here i can be equal to j), then the state becomes i-punishment. If the state is normal and **more than one player** in N **does not play** a_i, then state becomes more-punishment.

The strategy for Player j requires him to play a_j in normal state; play the action $\underline{a}_j^i$, corresponding to the minmax action profile against Player i, in i-punishment state, and play some fixed action (does not matter which one) in more-punishment state.

The strategy is shown in Table 34.

Predecessor state	Action profile observed	Current state	Recommended action
Normal	a	Normal	a_j
Normal	(a'_i, a_{-i})	i-punishment	$\underline{a}_j^i$
Normal	$(a'_S, a_{N \setminus S})$ with $\lvert S\rvert > 1$	more-punishment	Any fixed action
i-punishment	a'	i-punishment	$\underline{a}_j^i$
more punishment	a'	more punishment	Any fixed action

Table 32: Trigger strategy for Nash folk theorem

To see this strategy profile can be sustained in Nash equilibrium, first observe that the payoff from equilibrium is $v_i \equiv (u_i(a))$ for Player i. Suppose all the other players except i follows the prescribed strategy. Let the **best response** to a_{-i} give Player i a payoff $\bar{v}_i$ in the stage game G:

$$\bar{v}_i = \max_{a'_i \in A_i} u_i(a'_i, a_{-i}).$$

If Player i deviates, then he gets a maximum payoff of $\bar{v}_i$. This maximum payoff he gets in the first period he deviates and thereafter he is punished, and hence, gets a payoff less than or equal to $\underline{v}_i$. Hence, if he deviates in period t, his **maximum possible payoff** from deviation is (the original payoff can be less than this):

$$(1-\delta)\big(v_i + \delta v_i + \ldots + \delta^{t-1}\bar{v}_i + \delta^t \underline{v}_i + \delta^{t+1}\underline{v}_i + \ldots\big)$$

For deviation to be not profitable, we need to ensure that

$$v_i \geq (1-\delta)\big(v_i + \delta v_i + \ldots + \delta^{t-1}\bar{v}_i + \delta^t \underline{v}_i + \delta^{t+1}\underline{v}_i + \ldots\big).$$

Expanding the LHS, we get

$$(1-\delta)\big(v_i + \delta v_i + \delta^2 v_i + \ldots\big).$$

Canceling common terms in expanded LHS and RHS, we need to ensure that

$$\delta^{t-1}\bar{v}_i + \delta^t \underline{v}_i + \delta^{t+1}\underline{v}_i + \ldots \leq \delta^{t-1} v_i + \delta^t v_i + \delta^{t+1} v_i + \ldots.$$

This means, we need to ensure that $\bar{v}_i(1-\delta) + \delta \underline{v}_i \leq v_i$.

This is equivalent to ensuring

$$\delta \geq \frac{\bar{v}_i - v_i}{\bar{v}_i - \underline{v}_i}.$$

Define

$$\underline{\delta} := \frac{\bar{v}_i - v_i}{\bar{v}_i - \underline{v}_i}.$$

Note that by assumption $\bar{v}_i > v_i > \underline{v}_i$. Hence, $\underline{\delta} \in (0,1)$. This proves the claim. ■

The exact version of folk theorems involve use of mixed behavior strategies by players.

One of the issues with the Nash folk theorem is the strategies required to sustain the Nash equilibrium is very extreme - it requires you to punish the deviant for infinite number of periods. This may not be a reasonable threat. For instance, consider the game in Table 33. Theorem 21 says that (T, L) is achievable in Nash equilibrium of G^∞ for sufficiently patient players as long as the Column player can punish deviations by action R. This will hurt the Row player but the Column player is also badly hurt. This motivates the next set of results that require subgame perfect equilibrium - even punishments need to happen in equilibrium.

	L	R
T	6,6	0,-100
B	7,1	0,-100

Table 33: A Stage game

19.5 The One-Shot Deviation Principle

The one-shot deviation principle is a useful tool in the repeated games setting. Two strategies s_i and s'_i are one-shot deviations of each other if they differ from each other by actions chosen at one period for one history, i.e., $s_i^t(h^t) \neq \bar{s}_i^t(h^t)$ but $s_i^{t'}(h^{t'}) = \bar{s}_i^{t'}(h^{t'})$ for all $(t', h^{t'}) \neq (t, h^t)$. The one-shot deviation principle says that, fixing Player i and strategies s_{-i} of other players, if strategy s_i of Player i is optimal over all strategies $\bar{s}_i$ that are one-shot deviations from s_i, then it is optimal over all strategies.

To see why the one-shot deviation principle is true, consider Player i by fixing the strategies of other players at s_{-i}. Suppose strategy s_i is optimal over all one-shot deviations. Suppose another strategy s'_i differs from s_i at *finite* set of decision vertices (i.e., periods and histories). Then, we go the last period t where s_i and s'_i differ at some history h^t. In this subgame, s_i and s'_i differ from each other by one-shot deviation. Hence, s'_i cannot be profitable in this subgame. So, all the gains from s'_i must be occurring before this period. So, we restore s'_i to s_i in all histories in this period. We inductively repeat this procedure to reach a stage where s_i and s'_i are one-shot deviations. This is the same argument we have done for the backward induction procedure. Indeed, we did not use any specifics of repeated games in this argument.

The difference here is that s_i and s'_i can differ from each other at infinite number of decision vertices. Here, the discounted criteria of repeated games rescue us. Suppose strategy s_i is suboptimal. Then, there is some history h^t after which Player i can make a sequence of different moves than those prescribed by s_i. If the number of such different moves is finite, the previous argument applies. Else, let γ be the gain of Player i from this deviation, which starts in period t at history h^t. Let M be the best conceivable one-period *gain* in payoff to Player i by deviating from s_i. Choose a period $s > t$ such that $\delta^{s-t}M < \frac{\gamma}{2}$ - note that since M is finite and $\delta \in (0, 1)$, we can find such a s. Note that $\delta^{s-t}M$ is the maximum possible payoff gain from period s onwards - here, instead of multiplying the payoff in period t by δ^{t-1}, we multiply by 1 as if the game started from period t. Thus, gain from period s onwards cannot be more than $\frac{\gamma}{2}$. So, gain from period t to s must be at least $\frac{\gamma}{2}$. So, s_i can be modified at finite decision vertices such that we get a new strategy s''_i that is better than

s_i. Moreover, s_i'' differs from s_i at finite histories. But this contradicts our earlier argument that one-shot deviation principle guarantees deviations at finite decision histories.

19.6 Perfect Folk Theorem - Reversion to Nash

To make punishments credible, we must require Nash equilibrium at every subgame. This is the main motivation for using subgame perfect equilibrium. For every history, players must be playing Nash equilibrium actions. The following is quite immediate.

Proposition **4** *Suppose a is a Nash equilibrium of G. Then playing a at every period for every history is a subgame perfect equilibrium of G^∞.*

Proof: This follows from the one-shot deviation principle. If this strategy is not subgame perfect equilibrium, then there is some history h^t at which a Player i has a one-shot deviation, where he plays a_i'. But the payoff from such a deviation only differs from the the prescribed strategy by $u_i(a_i, a_{-i}) - u_i(a_i', a_{-i})$, which is positive because a is a Nash equilibrium. This completes the proof. ■

Now, denote by v_i^* the worst payoff of Player i over all Nash equilibria action profiles in G. Also, denote the corresponding Nash equilibrium profile as $a^{*,i}$. We are now ready to state a mild version of the perfect folk theorem.

Theorem **22 (Pure Perfect Folk Theorem with Nash Reversion)** *Suppose a is any action profile such that $u_i(a) > v_i^*$ for all $i \in N$. Then, there exists a $\underline{\delta} \in (0,1)$ such that for all $\delta \in (\underline{\delta}, 1)$, there is a subgame perfect equilibrium of G^∞ where a is played in every period on equilibrium path.*

Proof: We describe a strategy that is a subgame perfect equilibrium. It is described by three states: (a) normal state (b) i-punishment state, and (c) more-punishment state. The initial state is normal state. In normal state, the strategy recommends playing a_i to each Player i. Now, we inductively define the states at every history.

Consider Player j. If the state is normal and every player $i \in N$ plays a_i, then the state remains normal. If the state is normal and a **unique** player $i \in N$ **does not play** a_i (here i can be equal to j), then the state becomes i-punishment. If the state is normal and **more than one player** in N **does not play** a_i, then state becomes more-punishment.

The strategy for Player j requires him to play a_j in normal state, play the action $a_j^{*,i}$ corresponding to the **worst Nash equilibrium profile** of Player i (giving Player i a maximum payoff of v_i^*) for i-punishment state, and play an action corresponding some **fixed Nash equilibrium** (does not matter which one) in more-punishment state.

The strategy is shown in Table 34.

Predecessor state	Action profile observed	Current state	Recommended action
Normal	a	Normal	a_j
Normal	(a'_i, a_{-i})	i-punishment	Worst Nash for Player i - $a_j^{*,i}$
Normal	$(a'_S, a_{N\setminus S})$ with $\lvert S\rvert > 1$	more-punishment	Any fixed Nash action
i-punishment	a'	i-punishment	Worst Nash for Player i - $a_j^{*,i}$
more punishment	a'	more punishment	Any fixed Nash action

Table 34: Trigger strategy for perfect folk theorem

In any history which is either a i-punishment state or a more-punishment state, the strategy recommends playing a Nash equilibrium. By Proposition 4, this is a Nash equilibrium of this subgame.

The only complicated history is the one which is in normal state. Fix a Player i and suppose others are following s_{-i}. If Player i follows s_i, then he gets a payoff of $u_i(a)$. By the one-shot deviation principle, we need to check deviations in one history of this subgame. Suppose Player i deviates and plays another action a'_i in some period. He gets a payoff of $u_i(a'_i, a_{-i})$ in this period, but we move to i-punishment state in the subsequent periods. As a result, he gets a payoff of v_i^* after that. Hence, his payoff from deviation is

$$(1-\delta)u_i(a'_i, a_{-i}) + \delta v_i^*.$$

Hence, to be a subgame perfect equilibrium, we will need that

$$u_i(a) \geq (1-\delta)u_i(a'_i, a_{-i}) + \delta v_i^*.$$

This can be assured if we make sure the following holds:

$$u_i(a) \geq (1-\delta)\max_{a''_i \in A_i} u_i(a''_i, a_{-i}) + \delta v_i^*.$$

Denote $\max_{a''_i \in A_i} u_i(a''_i, a_{-i}) = d_i(a_{-i})$. Then, we need to ensure that $u_i(a) \geq (1-\delta)d_i(a_{-i}) + \delta v_i^*$. This is true if

$$\delta \geq \frac{d_i(a_{-i}) - u_i(a)}{d_i(a_{-i}) - v_i^*} = \underline{\delta}.$$

Note that $d_i(a_{-i}) \geq u_i(a) > v_i^*$ ensures that $\underline{\delta} \in [0,1)$. In other words, for $\delta \in [\underline{\delta}, 1)$, the recommended strategy is a subgame perfect equilibrium. This completes the proof. ■

19.7 Exact Versions of the Folk Theorems

Exact version of the Nash folk theorem and perfect folk theorem says that every strictly enforceable *feasible* payoff can be attained as a Nash equilibrium. The same statement is true for subgame perfect equilibrium under some additional conditions of the *feasible* payoff state.

Definition **34** *A payoff profile* $v \equiv (v_1, \ldots, v_n)$ *is* **feasible** *if for every action profile* a *in the stage game* G*, there exists* $\lambda_a \in [0,1]$ *with* $\sum_{a'} \lambda_{a'} = 1$ *and for every* $i \in N$

$$v_i = \sum_{a'} \lambda_{a'} u_i(a').$$

The set of all feasible payoff profiles is denoted as $Conv(V)$. These are payoffs that can be obtained by taking convex combination of different pure action profiles. In particular, if $V = \{v : v = u(a) \; \forall \; a \in A\}$, then $Conv(V)$ is just the convex hull of V - all vectors obtained by taking convex combination of vectors in V.

One way to interpret the feasible payoffs is that these are all the payoffs that can be obtained by playing *correlated strategies*. Correlated strategies require a public randomization device. So, achieving payoffs in $Conv(V)$ requires public randomization. This requires mixed/correlated behavior strategies. A mixed behavior strategy of an agent chooses a mixed action profile at every period. Now, the minmax payoff is determined using mixed action profiles. The problem with mixed actions is that it is difficult to detect deviations. This has led to a wide literature on *monitoring* technologies in repeated games. We give some informal idea about how the folk theorems look.

	L	R
T	3,0	1,-2
B	5,4	-1,6

Table 35: A Stage game

Consider the game in Table 35. We draw its feasible payoff vector in Figure 22. The minmax values of both the players are also shown in Figure 22. It is possible that the number of extreme points of this polytope is less than the number of action profiles. Check for a game with two players and two pure actions with payoffs: $(1,1), (2,2), (3,3), (4,4)$. Here, the feasible payoff vector set is a straight line joining $(1,1)$ and $(4,4)$.

It is clear that any action profile of the stage game leads to a feasible payoff vector. But if the players choose their mixed actions *independently*, then it is possible that some feasible payoff vector may not be attained - this is something we have seen earlier.

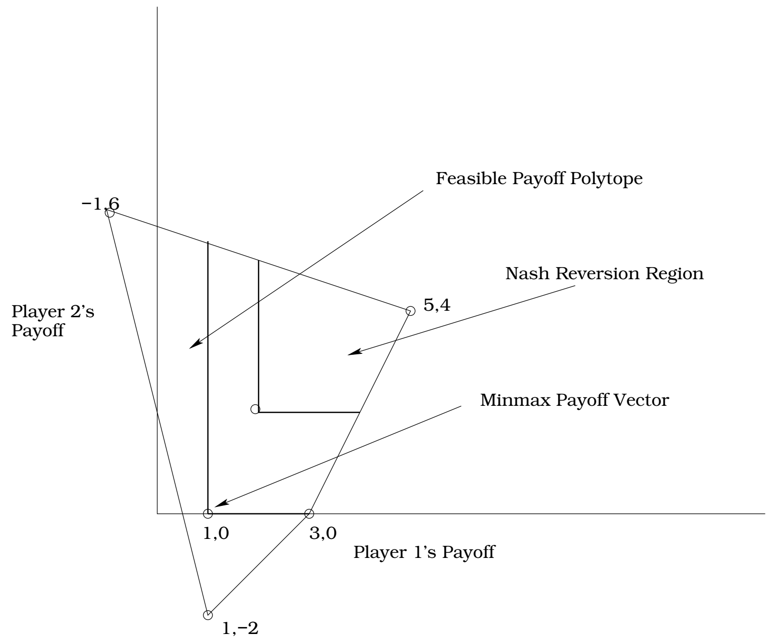

Figure 22: Feasible Payoff Vectors and Minmax Values

For this reason to achieve any payoff in the feasible payoff vector, the players should use *public randomization device*, and everyone observes the outcome of this device, and play a strategy according to this. The public randomization device randomizes amongst the (pure strategy) payoff vectors of the stage game. Based on the payoff vector chosen by the randomizing device, everyone chooses the corresponding strategy. An analogous proof to Theorem 21 and its subgame perfect version using public randomization device can be done to establish the exact folk theorems. They will say that every strictly enforceable feasible payoff can be achieved in Nash and subgame perfect equilibrium. The subgame perfect version of these theorems use more detailed "punishment and reward" strategies and extra technical condition. We state the theorem without a proof - the theorem is due to Fudenberg and Maskin.

THEOREM **23** *Suppose either $Conv(V)$ has dimension n or $n = 2$. Then, for every strictly enforceable feasible payoff vector, there is a discount factor (sufficiently close to 1) such that the infinitely repeated game generates the same payoff vector in a subgame perfect equilibrium.*

The proof of theorem uses a different type of strategy, which we illustrate below using an example. The stage game is shown in Table 36.

Notice that the minmax payoff vector is $(0, 0)$. The unique pure Nash equilibrium is (T, L). Using Theorem 22 is not so useful here. But the exact version of the folk theorem as-

	L	C	R
T	2,2	2,1	0,0
M	1,2	1,1	-1,0
B	0,0	0,-1	-1,-1

Table 36: A Stage game

sures that $(T, L), (T, C), (M, L), (M, C)$ are possible to get in a subgame perfect equilibrium. We show below how (M, C) is possible.

THEOREM **24** *Suppose $\delta \geq \frac{1}{2}$. Then, there is a subgame perfect equilibrium of the infinitely repeated game of the stage game in Table 36 such that (M, C) is played in every period in equilibrium.*

Proof: The strategy used classifies each history in each period as two states: (a) normal state (b) punishment state. A normal state recommends agents to play (M, C) and a punishment state recommends agents to play (B, R). The initial period (with null history) is a normal state.

Now, we can inductively define the state of every history. For every history in period t, there is a history in period $(t-1)$ that leads to this history, called the *predecessor*. If the predecessor is in normal state, and agents play (M, C), the current history (of period t) becomes a normal state. If the predecessor is in punishment state, and agents play (B, R), the current history becomes a normal state. Else, the current history becomes punishment state.

In other words, deviations (both in normal and punishment state) are punished for one period by staying in punishment state.

Hence, we can classify each history as a normal state or punishment state and look at deviations in each of them. Since the game is symmetric, we fix Player 1 without loss of generality and assume that Player 2 follows this strategy. If Player 1 follows the strategy, then he gets a payoff of 1. We consider two types of subgames.

NORMAL STATE. This is a subgame which starts from a normal state history. If the recommendation is followed, then player 1 gets 1. By the one-shot deviation principle, we only need to consider deviation in one period. If Player 2 plays C, then the maximum payoff of Player 1 by deviating is 2 in that period. Since this is a one period deviation, Player 1 follows the strategy from next period onwards. Since the next period will have a punishment history, he will undergo punishment and receive -1, and then normal state prevails, and he gets 1 from there onwards. The total payoff from deviation is thus computed as:

$$(1-\delta)\big(2+\delta(-1)+\delta^2+\delta^3+\dots\big) = (1-\delta)(1-2\delta)+1.$$

Since $\delta \geq \frac{1}{2}$, this expression is less than or equal to 1. Hence, deviation is not profitable.

Punishment state. This is a subgame which starts from a punishment state history. If the recommendation is followed, then Player 1 gets punished in this period and gets (-1), which is followed by normal state that gives 1 in each period. So, the total payoff is

$$(1-\delta)\big(-1+\delta+\delta^2+\dots\big) = 1-2(1-\delta).$$

The one-shot deviation will mean that Player 1 deviates in this period. Best deviation is to play T get 0. But this will result in a punishment in the next period and normal play from there on. Thus, the resulting payoff is

$$(1-\delta)\big(0+\delta(-1)+\delta^2+\delta^3+\dots\big) = 1-(1+2\delta)(1-\delta).$$

Note that since $\delta \geq \frac{1}{2}$, we have $1+2\delta \geq 2$. Hence, deviation is not profitable.

So, we conclude that deviation in any subgame is not profitable. This implies that the recommended strategy is a subgame perfect equilibrium. ■

The proof of the perfect Folk Theorem uses similar ideas but the punishment phase can last for more than one period (this is because the result is for general games). The number of periods the punishments last depend on the parameters of the problem.