7.1 Extensive Games with Perfect Information

An extensive game is a detailed description of the sequential structure of the decision problems encountered by the players in strategic situation. There is perfect information in such a game if each player, when making any decision, is perfectly informed of all the events that have previously occurred.

7.1.1 Definitions

Definition   An **extensive game with perfect information** $\langle N, H, P, U_i \rangle$ has the following components:

- A set of N players

- A set H of sequences (finite or infinite). Each element of H is a **history**; each component of a history is an **action** taken by a player.

- P is the **player function**, $P(h)$ being the player who takes an action after the history h.

- Payoff function $U_i, i \in N$

After any history h player $P(h)$ chooses an action from the set $A(h) = \{a : (h, a) \in H\}$. The empty history is the starting point of the game.

Example

Two players want two identical objects. One of them propose an allocation which the other either accepts or rejects. Both players are reasonable.
In this representation each node corresponds to a history and any edge corresponds to an action.
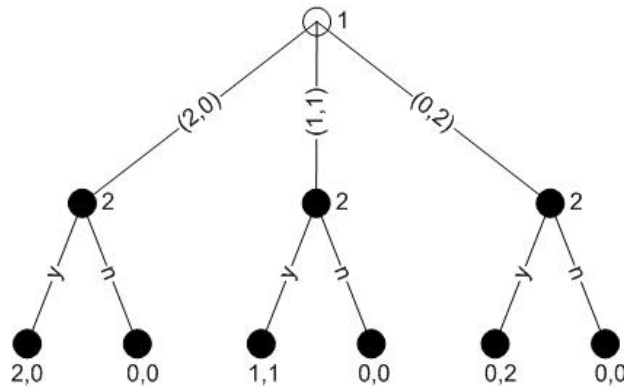
1

Figure 7.1: An extensive game, allocating two identical objects between two people

- $H = \{\emptyset, (2,0), (1,1), (0,2), ((0,2),y), ((2,0),n), ((1,1),y), ((1,1),n), ((0,2),y), ((0,2),n)\}$

- $P(\emptyset) = 1$ and $P((2,0)) = P((1,1)) = P((0,2)) = 2$

7.1.2   Strategy

Definition   A **strategy of player** $i \in N$ in an extensive game $\langle N, H, P, U_i \rangle$ is a function that assigns an action in $A(h)$ to each history $h \in H$ for which $P(h) = i$

A strategy specifies the action chosen by a player for *every* history after which it is her turn to move, *even for histories that , if the strategy is followed, are never reached.*

Example

In the game shown in figure 7.2, the strategies of the first player $S_1 = \{AE, AF, BE, BF\}$, i.e. her strategy specifies an action after the history (A,C), even if she chooses B at the beginning of the game.
One can transform an extensive game with perfect information to a normal game by setting all the possible histories as the possible choices for a normal game.

7.1.3   Nash Equilibrium

Definition   A **Nash equilibrium of an extensive game with perfect information** $\langle N, H, P, U_i \rangle$ is a strategy profile $s^* = (s_i)_{i \in N}$ such that for every player $i \in N$ and for every
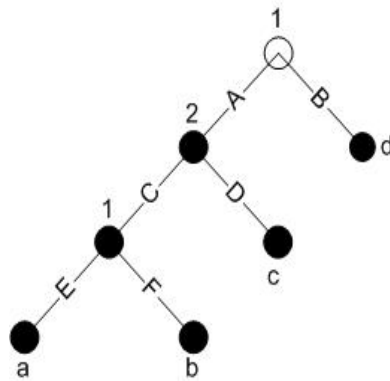
Figure 7.2: An extensive game in which player 1 moves before and after player 2

strategy s we have $U_i(s^*) \geq U_i(s)$
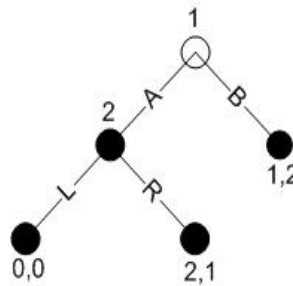
Example

Figure 7.3: Two players extensive game

The game has two Nash equilibria: (A, R) and (B, L) with payoff $(2, 1)$ and $(1, 2)$. The strategy profile (B, L) is a Nash equilibrium because given that player 2 chooses L, it is optimal for player 1 to choose B at the beginning. (B, R) is not a Nash equilibrium since then player one prefer to choose A. Player 2's choice L is a "threat" if player 1 chooses A. If player 2 chooses R, then player 1 prefer A since her payoff increases.

7.1.4   Subgame perfect Equilibrium

Definition   The **subgame of the extensive game with perfect information** $\Gamma =$

$\langle N, H, P, U_i \rangle$ that follows the history h is the extensive game $\Gamma(h) = \langle N, H|_h, P|_h, U_i|_h \rangle$ where $H|_h$ is the set of sequences $h\prime$ of actions for which $(h, h\prime) \in H$.

Definition A **subgame perfect equilibrium of an extensive game with perfect information** $\langle N, H, P, U_i \rangle$ is a strategy profile s^* such that for every player $i \in N$ and every history $h \in H$ for which $P(h) = i$ we have $U_i(s^*|_h) \geq U_i(s|_h)$ for every strategy s_i of player i in the subgame $\Gamma(h)$

Lemma 7.1 *The strategy profile s^* is a subgame perfect equilibrium if and only if for every player $i \in N$ and every history $h \in H$ for which $P(h) = i$ and for every $a_i \in A_i(h)$ exists $U_i(s^*|_h) \geq U_i(s^*_{-i}|_h, s_i)$ such that s_i differs from $s^*_i|_h$ only in the action a_i after the history h.*

    Proof: If s^* is a subgame perfect equilibrium then it satisfies the condition. Now suppose there is a history h which player $P(h)$ should change her action. Let h be the longest history as above. For $P(h) = i$ she can change to $a_i \in A_i(h)$ and increases her payoff. Thus s^* is not a subgame perfect equilibrium. □

Theorem 7.2 *Every extensive game with perfect information has a subgame perfect equilibrium.*

    Proof: We will use a backwards induction procedure. We start with the leaves and walk up through the tree. In every vertex we choose the best action (Best Response). By Lemma **??** this profile is a subgame perfect equilibrium. □

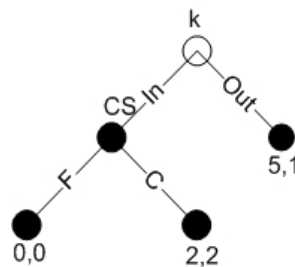**The Chain-Store Game**

Figure 7.4: Player's choices in city k in the chain-store game

    A chain-store (player CS) has branches in K cities. In each city k there is a single competitor, player k. In each period one of the potential competitors decides whether or

not to compete. If player k decides to compete then the chain-store can either fight (F) or cooperate (C). If challenged in any given city the chain-store prefers to cooperate rather then fight, but obtains the highest payoff if there is no entry. Each potential competitor is better off staying out than entering and being fought, but obtains the highest payoff when it enters and the chain-store is cooperative.

The game has a multitude of **Nash equilibria**: (Out, F) or (In, C).

The game has a unique **subgame perfect equilibrium**: every challenger chooses In and the chain-store always chooses C.

7.2 Repeated Games

The idea behind repeated games is that if we let the players play the same game a couple of times, they could get to different equilibria than those of a Nash Equilibrium of a one single round game. For example, we would like to achieve cooperation in the Prisoner's Dilemma game.

7.2.1 Finitely Repeated Games

Lets look again at the Prisoner's Dilemma game:

	C	D
C	$(3,3)$	$(0,4)$
D	$(4,0)$	$(1,1)$

Claim 7.3 *In a repeated game of T steps, where T is a final number, the only Nash Equilibrium is to play (D,D) in all T steps.*

Proof: In the last step, both players will play D, since otherwise at least one of the players would want to change her decision in order to improve her benefit. Now, using induction, if both players played the last i steps (D,D), then the same reason will hold for the $i-1$ step. $\square$

We shall look now at a modified Prisoner's Dilemma game:

Claim 7.4 *In the finite T steps modified game, there is a subgame perfect equilibrium for which the outcome is (C,C) in every step but the last three, in which it is (D,D).*

Proof: The strategy chosen by the first player should be to play $T-3$ times C and then the last 3 times to play D. However, if the second player has played differently than this

	C	D	E
C	$(3,3)$	$(0,4)$	$(0,0)$
D	$(4,0)$	$(1,1)$	$(0,0)$
E	$(0,0)$	$(0,0)$	$(\frac{1}{2},\frac{1}{2})$

strategy, we will play E for the rest of the steps. Since we stop cooperating at the $T-2$ step, its enough to see if the other player has played differently at the $T-3$ step. Here are the two possible outcomes starting from the $T-3$ step:

Playing according to the strategy will yield (C,C) (D,D) (D,D) (D,D). The total payoff if these steps for the second player is $3+1+1+1=6$.

If the second player will change her strategy, the best moves that can be made are (C,D) (E,E) (E,E) (E,E). The total payoff for the second player is $4+\frac{1}{2}+\frac{1}{2}+\frac{1}{2}=5\frac{1}{2}$.

As we can see playing differently than the stated strategy will yield less profit for the deviating player. Thus it is best to play the proposed strategy by both players. $\square$

The average payoff in this game is $(3(T-3)+3)/T$ which is $3-\frac{6}{T}$. This payoff is close to 3 which is the payoff of repeated cooperation.

7.2.2   Infinitely Repeated Games

There are several ways to look at the payoff of a player in an infinitely repeated game, a game that is repeated an infinite number of steps. We shall look at an N players game G with a payoff function $\vec{u}$, where u^i is the payoff function of player i. We define u^i_t as the payoff of player i at step t.

Definition The **average payoff** of a game G is the limit of the average payoff of the first T steps:
$\frac{1}{T}(\sum_{t=1}^{T} u^i_t) \rightarrow_{T\to\infty} \overline{u}^i$

Definition The **finite payoff** of a game G is the sum of the payoff of the first H steps of the game: $\sum_{t=1}^{H} u^i_t$

Definition The **discount payoff** of a game G is the weighted sum of the payoff of the steps of the game: $\sum_{t=1}^{\infty} u^i_t$

In the rest of the document when we refer to the payoff of an infinitely repeated game, we shall mean an average payoff $\overline{u}^i$.

Definition The **payoff profile** of an infinitely repeated game G is the payoff vector $\vec{w}$, where w_i is the payoff of player i. A payoff profile $\vec{w}$ is feasible if there are β_a for each

outcome $a \in A$, $K = \sum_{a \in A} \beta_a$, such that $\vec{w} = \sum_{a \in A} \frac{\beta_a}{K} \vec{u}(a)$.

Definition The **minimax payoff** of player i in a single step is: $v_i = \min_{a_{-i} \in A_{-i}} \max_{a_i \in A_i} u^i(a_{-i}, a_i)$

Claim 7.5 *In every Nash Equilibrium of a single game, the payoff of player i is at least v_i.*

    Proof: If a player has a smaller payoff than v_i then by the definition of the minimax payment, there is a different strategy that she can play in order to profit at least v_i.   □

Definition A payoff profile $\vec{w}$ is **enforceable** if $\forall_{i \in N} v_i \leq w_i$. A payoff profile is **strictly enforceable** if $\forall_{i \in N} v_i < w_i$.

Theorem 7.6 *Every feasible enforceable payoff profile $\vec{w}$ in an infinitely repeated game G is a Nash Equilibrium with an average payoff.*

    Proof: We will describe a strategy that is Nash Equilibrium with the payoff $\vec{w}$.
Since $\vec{w}$ is feasible there are β_a for each $a \in A$, $K = \sum_{a \in A} \beta_a$, such that $\vec{w} = \sum_{a \in A} \frac{\beta_a}{K} \vec{u}(a)$.
We shall assume that $\forall_{a \in A} \beta_a \in \mathbf{N}$.
The strategy of each player is to play cycles of K steps, going over all the possible outcomes $a \in A$ in an ordered list and playing her outcome in a β_a times. If player i deviates from this strategy, the rest of the players will change to a new strategy P_{-i} that enforces the payoff of player i to be at most the minimax payoff v_i.
Thus, if a player i deviates from the main strategy, her payoff will be v_i. which is not better than her payoff in $\vec{w}$. Because each deviation will not improve the payoff of player i, $\vec{w}$ is a Nash Equilibrium.   □

Theorem 7.7 *Every feasible strictly enforceable payoff profile $\vec{w}$ in an infinitely repeated game G has a Subgame Perfect Equilibrium with an average payoff.*

    Proof: We will describe a strategy that is Subgame Perfect Equilibrium with the payoff $\vec{w}$.
We shall use the same cyclic strategy as in the previous theorem, where all the players play the outcome a for β_a steps. If a player deviates from the strategy, the other players will punish her but only in a finite number of steps. At the end of the punishing steps all players will resume to play the cyclic strategy.
More specifically, at the end of each K steps cycle, the players will check if one of the players has deviated from the cyclic strategy. If a player, let say player j, has indeed played differently, the other players will play, for m^* steps, the minimax strategy P_{-j} that will enforce a payoff of at most v_j for player j, where m^* is chosen as follows:
We mark player j's strategy in each step t of the last cycle as $a_j'^t$. The maximal payoff benefit

for player j out of the steps in the cycle is $g^* = \max_{a^t}[u^j(a^t_{-j}, a'^t_j) - u^j(a^t)]$, where $a^t \in A$ is the expected outcome in step t of the K steps. We would like to get $Kg^* + m^* v_j < m^* w_j$ in order to make the punishment worthwhile. However, since $\vec{w}$ is strictly enforceable, we know that $v_j < w_j$ and so there exist m^* such that $0 < \frac{Kg^*}{w_j - v_j} < m^*$.

Playing the punishment strategy for m^* steps will yield a strictly smaller payoff for player j than playing the cyclic strategy without deviation. $\qquad\square$

7.3   Bounded Rationality

We have seen that in a repeated Prisoner's Dilemma game of N rounds there is no coopera-tion. A way to circumvent this problem is to assume that the players have limited resources. This kind of player is said to have **Bounded Rationality**.

A Bounded Rationality player is an automata with:

S - the state set.

A - the actions.

$\delta : S \times A \to S$ - the state transfer function.

$f : S \to A$ - the actions function.

S_0 - the starting state.

We assume that the automata is deterministic and that in each state only one action is chosen. A stochastic strategy is to randomly choose one automata from a set of deterministic automatas.

7.3.1   Tit for Tat

The Tit for Tat strategy (TfT) for the repeated Prisoner's Dilemma game will play in the next round what the opponent played in the last round (see figure **??**).

Theorem 7.8 *The TfT strategy is a Nash Equilibrium if the opponents has at most $N - 1$ states while the game is of N rounds.*

Proof: Any diversion of the opponent from the cooperation action C for $k > 0$ rounds, playing D, will yield a profit of $4 + 1(k - 1) + 0\delta$ for the opponent. On the other hand, if a cooperation is kept, the opponent's profit is $3k + 3\delta$. Where δ is 1 when the diversion is not at the end of the game (i.e. there are more rounds afterwards) and 0 at the end of the game. If δ is 1, the next round the opponent will play C, she will profit 0, while profiting 3
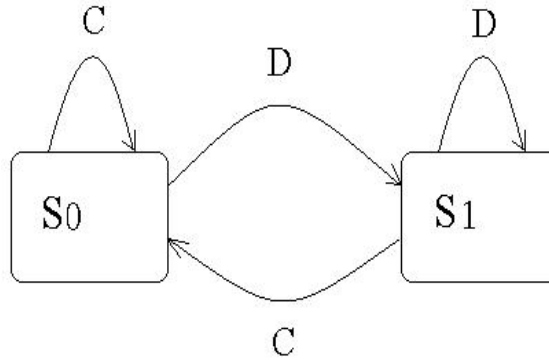
Figure 7.5: The automata of the Tit for Tat strategy

if a cooperation was maintained. This means that in the middle of the game the opponent will always prefer to cooperate since it has a bigger profit, $3k + 3 > 3 + k$. The problem is at the end of the game. While $k > 1$ it is still more profitable to cooperate, since $3k > 3 + k$. However when k is 1, it is better to defect, meaning that the only time that it is better to defect is in the last round.

If the opponent has an automata of at most $N - 1$ states then for any action series of length $N - 1$ the automata will return to an already visited state, arriving at a cycle. If until then the automata did not play D, it will never play D. However, if it did play D in one of the first $N - 1$ rounds the opponent will gain less than playing a full cooperation. Thus, the only Nash Equilibrium for at most $N - 1$ states automata is to play C constantly.

Since this logic is true for any kind of at most $N - 1$ states automata, it is also true for a stochastic strategy over a set of such automatas. □

7.3.2 Bounded Prisoner's Dilemma

We have seen an example for a simple bounded rationality strategy for the Prisoner's Dilemma game that will yield a cooperation in some conditions. The next step is to analyze any general bounded rationality strategy for that game, described by a final automata, and find what are the conditions that will lead to a cooperation between two players using these strategies in an N rounds game.

Theorem 7.9 *If, in a Repeated Prisoner's Dilemma with N rounds, both players have an automata with at least 2^{N-1} states then the only Equilibrium is the one is which both players play (D,D) in all rounds.*

Proof: Given an opponent, that optionally can play stochastically, it is possible to play an optimal strategy as follows:

We shall build a game history tree of depth N. At each node we shall calculate the distribution of the optimal action using dynamic programming, starting from the leafs and up the tree. Based on the profits of all the possible paths from a specific node to the leafs, we can choose the best response at every node.

The chosen optimal startegy can be encoded in a full binary tree of depth $N-1$, describing the first $N-1$ rounds, and one node for the last round (any optimal automata plays D in the last stage for every history of nonzero probability), summing to 2^{N-1} states.

As this is the optimal unrestricted strategy, the only Equilibrium is to play (D,D) at each round, as shown earlier.   □

Theorem 7.10 *For a Repeated Prisoner's Dilemma with N rounds, when both players have automatas with at most $2^{\epsilon_1 N}$ states (when it is possible to change to a different automata with a related size boundary $2^{\epsilon_2 N}$), there exists an Equilibrium with a profit of $3 - \epsilon_3$.*