Undergraduate Game Theory Lecture Notes

By

Omer Tamuz

California Institute of Technology

2018

Acknowledgments

These lecture notes are partially adapted from Osborne and Rubinstein [29], Maschler, Solan and Zamir [23], lecture notes by Federico Echenique, and slides by Daron Acemoglu and Asu Ozdaglar. I am indebted to Seo Young (Silvia) Kim and Zhuofang Li for their help in finding and correcting many errors. Any comments or suggestions are welcome.

Contents

Disclaimer

This a not a textbook. These are lecture notes.

1 What is a Game?

A game is a mathematical model of a *strategic interaction*. We will be studying a wide variety of games, but all of them will have the following common elements.

- *Players*. We often think of players as people, but sometimes they model businesses, teams, political parties, countries, etc.
- *Choices*. Players have to make a choice or multiple choices between different actions. A player's *strategy* is her rule for choosing actions.
- *Outcomes*. When the players are done choosing, an outcome is realized and the game ends. This outcome depends on the choices. Examples of outcomes include "player 1 wins," "Flora gets a dollar and Miles gets two dollars," or "a nuclear war starts and everyone dies."
- *Preferences*. Players have preferences over outcomes. For example, Flora may prefer the outcome "Flora gets two dollars and Miles gets nothing" over the outcome "Miles gets a dollar and Flora gets nothing." Miles may have the opposite preference.

Two important features make a game strategic: first, the fact that outcomes are determined by everyone's actions, rather than by the actions of just one player. Second, that players have different preferences. This creates tensions, which make games interesting.

Games differ in many aspects.

- *Timing*. Do players choose once (e.g., rock-paper-scissors), or again and again over time (e.g., chess)? In the latter case, does the game eventually end, or does it continue forever? Do they choose simultaneously, or in turn?
- *Observations*. Can players observe each other's choices?
- *Uncertainty*. Is the outcome random, or is it a deterministic function of the players' actions? Do some players have information that the others do not?

It is important to note that a game does not specify what the players actually do, but only what their options are and what the consequences are. Unlike an equation which (maybe) has a unique solution, the answer in games is much less clear cut. A *solution concept* is a way to think about what they players might decide to do. It is not part of the description of the game, and different solution concepts can yield different predictions for the same game.

2 Finite extensive form games with perfect information

We will start by studying a simple family of games, which includes many that are indeed games in the layperson meaning of the word. In these games players take turns making moves, all players observe all past moves, nothing is random, and the game ends after some fixed number of moves or less. We will present some examples and then define this class of games formally.

2.1 Tic-Tac-Toe

Two people play the following game. A three-by-three square grid is drawn on a piece of paper. The first player marks a square with an "x", then the second player marks a square from those left with an "o", etc. The winner is the first player to have marks that form either a row, a column or a diagonal.

Does the first player have a strategy that assures that she wins? What about the second player?

2.2 The Sweet Fifteen Game

Two people play the following game. There are cards on the table numbered one through nine, facing up, and arranged in a square. The first player marks a card with an "x", then the second player marks a card from those left with an "o", etc. The winner is the first player to have three cards (out of the three or more that they have picked) that sum to exactly fifteen.

Does the first player have a strategy that assures that she wins? What about the second player?

2.3 Chess

We assume the students are familiar with chess, but the details of the game will, in fact, not be important. We will choose the following (non-standard) rules for the ending of chess: the game ends either by the capturing of a king, in which case the capturing side wins and the other loses, or else in a draw, which happens when there a player has no legal moves, or more than 100 turns have elapsed.

As such, this games has the following features:

- There are two players, white and black.
- There are (at most) 100 times periods.
- In each time period one of the players chooses an action. This action is observed by the other player.
- The sequence of actions taken by the players so far determines what actions the active player is allowed to take.

- Every sequence of alternating actions eventually ends with either a draw, or one of the players winning.

We say that white can *force a victory* if, for any moves that black chooses, white can choose moves that will end in its victory. Zermelo showed in 1913 [34] that in the game of chess, as described above, one of the following three holds:

- White can force a victory.
- Black can force a victory.
- Both white and black can force a draw.

We will prove this later.

Exercise 2.1. *The same theorem applies to tic-tac-toe. Which of the three holds there?*

2.4 Definition of finite extensive form games with perfect information

In general, an extensive form game (with perfect information) G is a tuple $G = (N, A, H, P, \{u_i\}_{i\in N})$ where

1. N is a finite set of players.

2. A is a finite set of actions.

3. H is a finite set of allowed histories. This is a set of sequences of elements of A such that if $h \in H$ then every prefix of h is also in H. each $h = (a_1, a_2, \ldots, a_n)$ is a series of allowed legal moves in the game.

4. The set of terminal histories $Z \subseteq H$ is the set of sequences in H that are not subsequences of others in H. Thus Z is the set of histories at which the game ends. Note that we can specify H by specifying Z; H is the set of subsequences of sequences in Z.

5. P is a function from $H \setminus Z$ to N. When $P(h) = i$ then it is player i's turn to play after history h.

6. For each player $i \in N$, u_i is a function from the terminal histories to $\mathbb{R}$. The number $u_i(h)$ is the utility that player i assigns to the terminal history h. Players are assumed to prefer higher utilities. Note that the numbers themselves do not matter (for now); only their ordering matters.

We denote by $A(h)$ the actions available to player $P(h)$ after history h:

$$A(h) = \{a \in A : ha \in H\}.$$

A *strategy* for player i is a map σ_i from the set $\{h \in H : P(h) = i\}$ of histories h at which $P(h) = i$ to the set of actions A. A *strategy profile* $s = \{s_i\}_{i\in N}$ constitutes a strategy for each player. Given a stratgy profile we know how players are going to play, and we denote by $h(s)$ the *path of play*, i.e., the history that is realized. We also denote by $u_i(s)$ the utility that player i recieves under this history.

2.5 The ultimatum game

In the ultimatum game player 1 makes an offer $a \in \{0, 1, 2, 3, 4\}$ to player 2. Player 2 either accepts or rejects. If player 2 accepts then she receives a dollars and player 1 receives $4 - a$ dollars. If 2 rejects then both get nothing. This is how this game can be written in extensive form:

1. $N = \{1, 2\}$.

2. $A = \{0, 1, 2, 3, 4, a, r\}$.

3. $Z = \{0a, 1a, 2a, 3a, 4a, 0r, 1r, 2r, 3r, 4r\}$.

4. $O = \{(0,0),(0,4),(1,3),(2,2),(3,1),(4,0)\}$. Each pair corresponds to what players 1 receives and what player 2 receives.

5. For $b \in \{0,1,2,3,4\}$, $u_1(ba) = 4 - b$, $u_2(ba) = b$ and $u_1(br) = u_2(br) = 0$.

6. $P(\emptyset) = 1$, $P(0) = P(1) = P(2) = P(3) = P(4) = 2$.

A strategy for player 1 is just a choice among $\{0,1,2,3,4\}$. A strategy for player 2 is a map from $\{0,1,2,3,4\}$ to $\{a,r\}$: player 2's strategy describes whether or not she accepts or rejects any given offer.

Remark 2.2. *A common mistake is to think that a strategy of player 2 is just to choose among $\{a,r\}$. But actually a strategy is a complete contingency plan, where an action is chosen for every possible history in which the player has to move.*

2.6 Equilibria

Given a strategy profile $s = \{s_i\}_{i\in N}$, we denote by (s_{-i}, s'_i) the strategy profile in which i's strategy is changed from s_i to s'_i and the rest remain the same.

A strategy profile s^* is a *Nash equilibrium* if for all $i \in N$ and strategy s_i of player i it holds that

$$u_i(s^*_{-i}, s_i) \le u_i(s^*).$$

When s is the equilibrium $h(s)$ is also known as the *equilibrium* path associated with s.

Example: in the ultimatum game, consider the strategy profile in which player 1 offers 3, and player 2 accepts 3 or 4 and rejects 0, 1 or 2. It is easy to check that this is an equilibrium profile.

2.7 The centipede game

In the centipede game there are n time periods and 2 players. The players alternate in turns, and at each turn each player can either stop (S) or continue (C), except at the last turn, where they must stop. Now, there is a piggy bank which initially has in it 2 dollars. In the beginning of each turn, this amount doubles. If a player decides to stop (which she must do in period n), she is awarded three fourth of what's in the bank, and the other player is awarded the remainder. If a player decides to continue, the amount in the bank doubles. Hence, in period m, a player is awarded $\frac{3}{4} \cdot 2 \cdot 2^m$ if she decided to stop, and the other player is given $\frac{1}{4} \cdot 2 \cdot 2^m$.

	$m=1$	$m=2$	$m=3$	$m=4$	$m=5$	$m=6$	$m=7$	$m=8$	$m=9$	
player 1	3	2	12	8	48	32	192	128	768	512
player 2	1	6	4	24	16	96	64	384	256	1536

Exercise 2.3. *Define the centipede game formally, for $n = 4$. How many strategies does each player have? Make sure you understand Remark 2.2 before answering this.*

Exercise 2.4. *Show that the strategy profile in which both players play S in every time period is an equilibrium.*

Theorem 2.5. *In every Nash equilibrium, player 1 plays S in the first period.*

Proof. Assume by contradiction that player 1 plays C in the first period under some equilibrium s. Then there is some period $m > 1$ in which S is played for the first time on the equilibrium path. It follows that the player who played C in the previous period is awarded $\frac{1}{4}\cdot 2\cdot 2^m$. But she could have been awarded $\frac{3}{4}\cdot 2\cdot 2^{m-1} = \frac{3}{2}\frac{1}{4}\cdot 2\cdot 2^m$ by playing S in the previous period, and therefore s is not an equilibrium. □

2.8 Subgames and subgame perfect equilibria

A *subgame* of a game $G = (N, A, H, P, \{u_i\}_{i\in N})$ is a game that starts after a given finite history $h \in H$. Formally, the subgame $G(h)$ associated with $h = (h_1, \ldots, h_n) \in H$ is $G(h) = (N, A, H_h, P, \{u_i\}_{i\in N})$, where

$$H_h = \{(a_1, a_2, \ldots) : (h_1, \ldots, h_n, a_1, a_2, \ldots) \in H\}.$$

The functions P and u_i are as before, just restricted to the appropriate subdomains.

A strategy s of G can likewise used to define a strategy s_h of $G(h)$. We will drop the h subscripts whenever this does not create (too much) confusion.

A *subgame perfect equilibrium* of G is a strategy profile s^* such that for every subgame $G(h)$ it holds that s^* (more precisely, its restriction to H_h) is a Nash equilibrium of $G(h)$. We will prove Kuhn's Theorem, which states that every finite extensive form game with perfect information has a subgame perfect equilibrium. We will then show that Zermelo's Theorem follows from Kuhn's.

As an example, consider the following Cold War game played between the USA and the USSR. First, the USSR decides whether or not to station missiles in Cuba. If it does not, the game ends with utility 0 for all. If it does, the USA has to decide if to do nothing, in which case the utility is 1 for the USSR and -1 for the USA, or to start a nuclear war, in which case the utility is -1,000,000 for all.

Exercise 2.6. *Find two equilibria for this game, one of which is subgame perfect, and one which is not.*

Exercise 2.7. *Find two equilibria of the ultimatum game, one of which is subgame perfect, and one which is not.*

An important property of finite horizon games is the *one deviation property*. Before introducing it we make the following definition.

Let s be a strategy profile. We say that s'_i is a *profitable deviation* from s for player i at history h if s'_i is a strategy for the subgame G such that

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

Note that a strategy profile has no profitable deviations if and only if it is a subgame perfect equilibrium.

Theorem 2.8 (The one deviation principle)**.** *Let $G = (N, A, H, P, \{u_i\}_{i\in N})$ be a finite extensive form game with perfect information. Let s be a strategy profile that is* not *a subgame perfect equilibrium. There there exists some history h and a profitable deviation $\bar{s}_i$ for player $i = P(h)$ in the subgame $G(h)$ such that $\bar{s}_i(k) = s_i(k)$ for all $k \neq h$.*

Proof. Let s be a strategy profile that is *not* a subgame perfect equilibrium. Then there is a subgame $G(h)$ and a strategy s'_i for player $i = P(h)$ such that s'_i is a profitable deviation for i in $G(h)$. Denote $s' = (s_{-i}, s'_i)$, and note that $u_i(s') > u_i(s)$. Let h be a history that is maximal in length among all histories with this property. Let $\bar{s}_i$ be given by $\bar{s}_i(k) = s_i(k)$ for all $k \neq h$, and $\bar{s}_i(h) = s'_i(h)$. By the maximal depth property of h we have that $\bar{s}_i$ is still a profitable deviation, since otherwise i would have a profitable deviation in some subgame of $G(h)$. We thus have that $\bar{s}_i$ is a profitable deviation for $G(h)$ that differs from s_i in just one history. $\square$

2.9 The dollar auction

Two players participate in an auction for a dollar bill. Player 1 acts in the odd periods, and player 2 in the even periods. Both players start with a zero bid. In each period the playing player can either stay or quit. If she quits the other player gets the bill, both pay the highest they have bid so far, and the game ends. If she stays, she must bid 10 cents higher than the other player's last bid (except in the first period, when she must bid 5 cents) and the game continues. If one of the bids exceeds 100 dollars the game ends, the person who made the highest bid gets the dollar, and both pay the highest they have bid so far. So assuming both players stay, in the first period the first player bids 5 cents. In the second period the second player bids 15 cents. In the third period the first player bids 25 cents, etc.

Exercise 2.9. *Does this game have equilibria? Subgame perfect equilibria?*

2.10 Backward induction, Kuhn's Theorem and a proof of Zermelo's Theorem

Let $G = (N, A, H, P, \{u_i\}_{i \in N})$ be an extensive form game with perfect information. Recall that $A(\emptyset)$ is the set of allowed initial actions for player $i = P(\emptyset)$. For each $b \in A(\emptyset)$, let $s^{G(b)}$ be some strategy profile for the subgame $G(b)$. Given some $a \in A(\emptyset)$, we denote by s^a the strategy profile for G in which player $i = P(\emptyset)$ chooses the initial action a, and for each action $b \in A(\emptyset)$ the subgame $G(b)$ is played according to $s^{G(b)}$. That is, $s^a_i(\emptyset) = a$ and for every player j, $b \in A(\emptyset)$ and $bh \in H \setminus Z$, $s^a_j(bh) = s^{G(b)}_j(h)$.

Lemma 2.10 (Backward Induction)**.** *Let $G = (N, A, H, P, \{u_i\}_{i \in N})$ be a finite extensive form game with perfect information. Assume that for each $b \in A(\emptyset)$ the subgame $G(b)$ has a subgame perfect equilibrium $s^{G(b)}$. Let $i = P(\emptyset)$ and let a be a maximizer over $A(\emptyset)$ of $u_i(s^{G(a)})$. Then s^a is a subgame perfect equilibrium of G.*

Proof. By the one deviation principle, we only need to check that s^a does not have deviations that differ at a single history. So let s differ from s^a at a single history h.

If h is the empty history then $s = s^{G(b)}$ for $b = s_i(\emptyset)$. In this case $u_i(s^a) > u_i(s) = u_i(s^{G(b)})$, by the definition of a as the maximizer of $u_i(s^{G(a)})$.

Otherwise, h is equal to bh' for some $b \in A(\emptyset)$ and $h' \in H_b$, and $u_i(s) = u_i(s)$. But since s^a is a subgame perfect equilibrium when restricted to $G(b)$ there are no profitable deviations, and the proof is complete. □

Kuhn [22] proved the following theorem.

Theorem 2.11 (Kuhn, 1953)**.** *Every finite extensive form game with perfect information has a subgame perfect equilibrium.*

Given a game G with allowed histories H, denote by $\ell(G)$ the maximal length of any history in H.

Proof of Theorem 2.11. We prove the claim by induction on $\ell(G)$. For $\ell(G) = 0$ the claim is immediate, since the trivial strategy profile is an equilibrium, and there are no proper subgames. Assume we have proved the claim for all games G with $\ell(G) < n$.

Let $\ell(G) = n$, and denote $i = P(\emptyset)$. For each $b \in A(\emptyset)$, let $s^{G(b)}$ be some subgame perfect equilibrium of $G(b)$. These exist by our inductive assumption, as $\ell(G(b)) < n$.

Let $a^* \in A(\emptyset)$ be a maximizer of $u_i(s^{a^*})$. Then by the Backward Induction Lemma s^{a^*} is a subgame perfect equilibrium of G, and our proof is concluded. □

Given Kuhn's Theorem, Zermelo's Theorem, as stated below, admits a simple proof.

Theorem 2.12 (Zermelo)**.** *Let G be a finite extensive form game with two players and where $u_1 = -u_2$ and $u_1(h) \in \{-1, 0, 1\}$. Then exactly one of the following three hold:*

1. *There exists a stragegy s^*_1 for player 1 such that $u_1(s^*_1, s_2) = 1$ for all strategies s_2 of player 2.*

2. *There exists a stragegy s_2^* for player 2 such that $u_2(s_1, s_2^*) = 1$ for all strategies s_1 of player 2.*

3. *There exist strategies s_1^*, s_2^* for players 1 and 2 such that $u_1(s_1^*, s_2) \geq 0$ and $u_2(s_1, s_2^*) \geq 0$ for all strategies s_1, s_2 of players 1 and 2.*

Proof. Let s^* be a subgame perfect equilibrium of any finite extensive form game with two players and where $u_1 = -u_2$ and $u_1(h) \in \{-1, 0, 1\}$. Consider these three cases. If $u_1(s^*) = 1$ then for any s_B

$$u_2(s_1^*, s_2) \leq u_2(s^*) = -1.$$

But $u_2 \geq -1$, and so $u_2(s_1^*, s_2) = -1$. That is, player 1 can force victory by playing s_1^*. The same argument shows that if $u_2(s^*) = 1$ then black can force victory. Finally, if $u_1(s^*) = 0$ then for any s_2

$$u_2(s_1^*, s_2) \leq u_2(s^*) = 0,$$

so $u_2(s_1^*, s_2)$ is either 0 or -1. By the same argument $u_1(s_1, s_2^*)$ is either 0 or -1 for any s_1, and so we have proven the claim. $\square$

3 Strategic form games

3.1 Definition

A game in strategic form (or normal form) is a tuple $G = (N, \{S_i\}_{i\in N}, \{u_i\}_{i\in\mathbb{N}})$ where

- N is the set of players.
- S_i is the set of actions or strategies available to player i. We denote by $S = \prod_i S_i$ the set of strategy profiles.
- The function $u_i \colon S \to \mathbb{R}$ is player i's utility (or payoff) for each strategy profile.

We will assume that players have the obvious preferences over utility: more is prefered to less. We say that G is finite if N is finite and S is finite.

3.2 Nash equilibria

Given a strategy profile s, a *profitable deviation* for player i is a strategy t_i such that

$$u_i(s_{-i}, t_i) > u_i(s_{-i}, s_i).$$

A strategy profile s is a *Nash equilibrium* if no player has a profitable deviation. These are also called *pure* Nash equilibria, for reasons that we will see later. They are often just called equilibria.

3.3 Classical examples

- **Extensive form game with perfect information.** Let $G = (N, A, H, P, \{u_i\}_{i\in N})$ be an extensive form game with perfect information, where, instead of the usual outcomes and preferences, each player has a utility function $u_i \colon Z \to \mathbb{R}$ that assigns her a utility at each terminal node. Let G' be the strategic form game given by $G' = (N', \{S_i\}_{i\in N}, \{u_i\}_{i\in\mathbb{N}})$, where
  - $N' = N$.
  - S_i is the set of G-strategies of player i.
  - For every $s \in S$, $u_i(s)$ is the utility player i gets in G at the terminal node at which the game arrive when players play the strategy profile s.

  We have thus done nothing more than having written the same game in a different form. Note, however, that not every game in strategic form can be written as an extensive form game with perfect information.

  Exercise 3.1. *Show that $s \in S$ is a Nash equilibrium of G iff it is a Nash equilibrium of G'.*

Note that a disadvantage of the strategic form is that there is no natural way to define subgames or subgame perfect equilibria.

- **Matching pennies.** In this game, and in the next few, there will be two players: a row player (R) and a column player (C). We will represent the game as a payoff matrix, showing for each strategy profile $s = (s_R, s_C)$ the payoffs $u_R(s), u_C(s)$ of the row player and the column player.

	H	T
H	1,0	0,1
T	0,1	1,0

In this game each player has to choose either heads (H) or tails (T). The row player wants the choices to match, while the row player wants them to mismatch.

Exercise 3.2. *Show that matching pennies has no pure Nash equilibria.*

- **Prisoners' dilemma.**

Two prisoners are faced with a dilemma. A crime was committed in the prison, and they are the only two who could have done it. Each prisoner has to make a choice between testifying against the other (and thus betraying the other) and keeping her mouth shut. In the former case we say that the prisoner defected (i.e., betrayed the other), and in the latter she cooperated (with the other prisoner, not with the police).

If both cooperate (i.e., keep their mouths shut), they will have to serve the remainder of their sentences, which are 2 years each. If both defect (i.e., agree to testify against each other), each will serve 3 years. If one defects and the other cooperates, the defector will be released immediately, and the cooperator will serve 10 years for the crime.

Assuming that a player's utility is minus the number of years served, the payoff matrix is the following.

	D	C
D	$-3,-3$	$0,-10$
C	$-10,0$	$-2,-2$

Exercise 3.3. *Show that the unique pure Nash equilibrium is* (D,D).

- **Split or steal.**

Split or steal is a game played in a British game show called *golden balls*: https://www.youtube.com and this one: https://www.youtube.com/watch?v=S0qjK3TWZE8.

	St	Sp
St	0,0	1,0
Sp	0,1	1/2,1/2

Exercise 3.4. *What are the equilibria of this game?*

- **Stag hunt.**

  Two friends go together on a hunt. Each has to decide whether to try and catch a hare—which she can do by herself—or to cooperate on trying to catch a stag. The payoff matrix is the following.

	S	H
S	2,2	0,1
H	1,0	1,1

Exercise 3.5. *What are the Nash equilibria?*

- **Battle of the sexes**[1]**.** Adam and Steve are a married couple. They have to decide whether to spend the night at the monster truck show (M) or at the opera (O). Adam (the row player) prefers the truck show, while Steve prefers the opera. Both would rather go out then do nothing, which is what would happen if they could not agree. Their payoff matrix is the following.

	M	O
M	2,1	0,0
O	0,0	1,2

Exercise 3.6. *Find all the equilibria of this game.*

- **Bertrand competition.** There are $n > 1$ companies that are selling the same product. There are 1,000 customers who want to buy one product each, but will not buy if the price is more than \$1,000. Each company has to set a price, which we will assume has to be one of $\{1, 2, \ldots, 1000\}$.

  All customers buy from the company that set the lowest price. If more than one has set the lowest price then the revenue is split evenly between these companies. Therefore, if m companies chose the lowest price p, then the payoff to each of these companies is $1000 \cdot p/m$, and the payoff to the rest is zero.

Exercise 3.7. *Find all the equilibria of this game.*

Thought experiment: what can the members of the cartel do (in real life) to ensure that all set the price to \$1,000?

- **The regulated cartel.** This is as in the previous game, but now in addition each company that sets a price p has to pay a tax of p/n.

[1]The name of this game seems archaic in this day and age, but we will keep it, as it is standard in the literature.

Exercise 3.8. *Find the equilibria of this game.*

- **Cournot.** Consider a finite set of companies N that are selling the same product. Each has to decide on a quantity q_i to manufacture. The function $D(p)$ describes the demand as a function of price, and we set denote its inverse by $P = D^{-1}$. Let $c(q)$ be the cost of manufacturing, as a function of quantity. Assuming prices are set so that all the manufactured goods are sold, company i's payoff is

$$u_i(q_1,\ldots,q_n) = q_i \cdot P\left(\sum_i q_i\right) - c(q_i).$$

Exercise 3.9. *Let $P(q) = A - B \cdot q$, let $c(q) = C \cdot q$ (so that the cost per unit is independent of the amount manufactured), and assume $A > C > 0$, $B > 0$. Show that there is a Nash equilibrium in which all companies set the same quantity q, and calculate this q.*

- **Public goods.** Each of four players has to choose an amount in $q_i \in [0,1]$ to contribute to a public goods project. The total amount contributed $q = \sum_i q_i$ is multiplied by 2 to yield the total return $r = 2q$. This return is distributed evenly among the players. Hence player i's utility is $r/4 - q_i$.

Exercise 3.10. *Find all equilibria of this game.*

Exercise 3.11. *Now arrange the players in some order and have each one choose after seeing the choices of her predecessors. What are the possible utilities in equilibria of this game? In subgame perfect equilibria?*

- **Voter turnout.** There is a finite set of voters $N = N^a \cup N^b$ with $N^a \cap N^b = \emptyset$ and $|N^a| \geq |N^b|$. Each voter chooses one of the three actions $\{a, b, 0\}$: either vote for candidate a, for candidate b, or abstain. After the vote the winner is decided according to the majority, or is chosen randomly if the vote was tied.

  Each voter in N^a gets utility 1 if a wins, utility 0 if b wins, and utility 1/2 if the winner is chosen at random. The same hold for voters in N^b, with the roles of a and b revered. Additionally, each voter pays a cost $c \in (0, 1/2)$ if they do not abstain.

Exercise 3.12. *Find all equilibria of this game. In the following cases: (1) N^a and N^b are of equal size and (2) N^a is larger than N^b.*

3.4 Dominated strategies

A strategy s_i of player i in $G = (N, \{S_i\}_{i \in N}, \{u_i\}_{i \in \mathbb{N}})$ is *strictly dominated* (or just dominated) if there exists another strategy t_i such that, for all choices of strategy s_{-i} of the other players it holds that

$$u_i(s_{-i}, t_i) > u(s_{-i}, s_i).$$

That is, regardless of what the other players do, t_i is a better choice for i than s_i.

We say that s_i is *weakly dominated* if there exists another strategy t_i such that for all s_{-i}

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

and furthermore for some s_{-i}

$$u_i(s_{-i}, t_i) > u_i(s_{-i}, s_i).$$

Exercise 3.13. *Does matching pennies have strictly dominated strategies? Weakly dominated strategies? How about the prisoners' dilemma? The regulated cartel? Voter turnout?*

3.5 Repeated elimination of dominated strategies

It seems unreasonable that a reasonable person would choose a strictly dominated strategy, because she has an obviously better choice. Surprisingly, taking this reasoning to its logical conclusion leads to predictions that sharply contradict observed human behavior.

Consider the regulated cartel game G. It is easy to see that \$1,000 is a dominated strategy; if all other companies choose \$1,000 then \$999 is a better strategy. If the lowest price $p_{\min}$ \$999, the clearly playing \$1,000 better. And if it is lower than \$999, then still \$999 is a better strategy, since then the tax is smaller. It is likewise easy to check that \$1,000 is the only dominated strategy.

Since no reasonable player would choose \$1,000, it is natural to define a new game G' which is identical to G, except that the strategy space of every player no longer includes \$1,000. Indeed, assuming that we are interested in equilibria, the following theorem guarantees that analyzing G' is equivalent to analyzing G.

Theorem 3.14. *Let $G = (N, \{S_i\}_{i \in N}, \{u_i\}_{i \in \mathbb{N}})$, let $d_j \in S_j$ be a dominated strategy of player j, and let $G' = (N, \{S'_i\}_{i \in N}, \{u'_i\}_{i \in \mathbb{N}})$, where*

$$S'_i = \begin{cases} S_i & \text{for } i \neq j \\ S_j \setminus \{d_j\} & \text{for } i = j \end{cases},$$

and u'_i is the restriction of u_i to S'.

Then every Nash equilibrium $s \in S$ of G is in S'. Furthermore $s \in S'$ is a Nash equilibrium of G if and only if it is a Nash equilibrium of G'.

The proof of this Theorem is straightforward and is left as an exercise.

Now, as before, it is easy to see that \$999 is a dominated strategy for G', and to therefore remove it from the set of strategies and arrive at a new game, G''. Indeed, if we repeat this process, we will arrive at a game in which the single strategy is \$1. However, in experiments this is very rarely the chosen strategy[2].

The following theorem of Gilboa, Kalai and Zemel [17] shows that the order of elimination of dominated strategies does not matter, as long as one continues eliminating until there are no more dominated strategies. Note that this is not true for weakly dominated strategies [17].

Theorem 3.15 (Gilboa, Kalai and Zemel, 1990)**.** *Fix a finite game G, and let G_1 be a game that*

- *is the result of repeated elimination of dominated strategies from G, and*
- *has no dominated strategies.*

Let G_2 be a game with the same properties. Then $G_1 = G_2$.

3.6 Dominant strategies

A strategy s_i of player i in $G = (N, \{S_i\}_{i \in N}, \{u_i\}_{i \in \mathbb{N}})$ is *strictly dominant* if for every other strategy t_i it holds that

$$u_i(s_{-i}, s_i) > u_i(s_{-i}, t_i).$$

That is, regardless of what the other players do, s_i is a better choice for i than t_i.

Exercise 3.16. *Which of the games above have a strictly dominant strategy?*

The proof of the following theorem is straightforward.

Theorem 3.17. *If player i has a strictly dominant strategy d_i then, for every pure Nash equilibrium s^* it holds that $s^*_i = d_i$.*

[2]Except in populations of game theory students.

3.7 Mixed equilibria and Nash's Theorem

Given a finite set X, denote by ΔX set of probability distributions over X.

Let $G = (N, \{S_i\}, \{u_i\})$ be a finite game. A *mixed strategy* σ_i is an element of ΔS_i. Given a *mixed strategy profile* $\sigma = (\sigma_1, \ldots, \sigma_n)$, we overload notation and let σ be the element of ΔS given by the product $\prod_i \sigma_i$. That is, σ is the distribution over $\prod_i S_i$ in which we pick independently from each S_i, with distribution σ_i.

Define a game $\hat{G}(N, \{\hat{S}_i\}, \{\hat{u}_i\})$ as follows:

- $\hat{S}_i = \Delta S_i$.
- $\hat{u}_i(\sigma) = \mathbb{E}_{s\sim\sigma}[u_i(s)]$.

That is, $\hat{G}$ is a game whose strategies are the mixed strategies of G, and whose utilities are the expected utilities of G, taken with respect to the given mixed strategies. A pure Nash equilibrium of $\hat{G}$ is called a *mixed Nash equilibrium* of G. That is, a mixed strategy profile $\sigma \in \prod_i \Delta S_i$ is a mixed Nash equilibrium if no player can improve her expected utility by deviating to another mixed strategy. We will often just say "Nash equilibrium" when referring to mixed equilibria. We will use u_i to also mean $\hat{u}_i$; that is, we will extend u_i from a function $S \to \mathbb{R}$ to a function $\Delta S \to \mathbb{R}$.

Fix any σ_{-i}, and note that as a function $\Delta S_i \to \mathbb{R}$, $u_i(\sigma_{-i}, \cdot)$ is linear. That is,

$$u_i(\sigma_{-i}, \alpha\sigma_i + (1-\alpha)\tau_i) = \alpha u_i(\sigma_{-i}, \sigma_i) + (1-\alpha)u_i(\sigma_{-i}, \tau_i).$$

Nash's celebrated theorem [27] states that every finite game has a mixed equilibrium.

Theorem 3.18 (Nash, 1950)**.** *Every finite game has a mixed Nash equilibrium.*

To prove Nash's Theorem we will need Brouwer's Fixed Point Theorem.

Theorem 3.19 (Brouwer's Fixed Point Theorem)**.** *Let X be a compact convex subset of $\mathbb{R}^d$. Let $T\colon X \to X$ be continuous. Then T has a fixed point. I.e., there exists an $x \in X$ such that $T(x) = x$.*

Corollary: if you are in a room and hold a map of the room horizontally, then there is a point in the map that is exactly above the point it represents.

Exercise 3.20. *Prove Brouwer's Theorem for the case that X is a convex compact subset of $\mathbb{R}$.*

Exercise 3.21. *Show that both the compactness and the convexity assumptions are necessary.*

Brouwer's Theorem has a simple proof for the case that T is affine, i.e., in the case that

$$T(\alpha x + (1-\alpha)y) = \alpha T(x) + (1-\alpha)T(y)$$

for all $0 \le \alpha \le 1$.

To give some intuition for this proof consider first the case that T has a *periodic point*. That is, if we denote by T^n the n-fold composition of T with itself[3] then there is some x and $n \geq 1$ such that $T^n(x) = x$. In this case

$$z = \frac{1}{n}\left(x + T(x) + T^2(x) + \cdots + T^{n-1}(x)\right)$$

is a fixed point of T, by the affinity property of T.

For the more general case that T has no periodic points, pick any $x \in X$. Let

$$x_n = \frac{1}{n}\left(x + T(x) + T^2(x) + \cdots + T^{n-1}(x)\right).$$

Because T is affine,

$$T(x_n) = \frac{1}{n}\left(T(x) + T^2(x) + T^3(x) + \cdots + T^n(x)\right).$$

Adding and subtracting $\frac{1}{n}x$ yields

$$\begin{aligned} T(x_n) &= \frac{1}{n}\left(x + T(x) + T^2(x) + \cdots + T^{n-1}(x)\right) + \frac{1}{n}T^n(x) - \frac{1}{n}x \\ &= x_n + \frac{1}{n}T^n(x) - \frac{1}{n}x. \end{aligned}$$

Hence

$$\lim_n T(x_n) - x_n = \lim_n \frac{1}{n}T^n(x) - \frac{1}{n}x = 0.$$

Since X is compact the sequence $(x_n)_n$ has a subsequence $(x_{n_k})_k$ that converges to some $z \in X$. Along this subsequence we also have that $\lim_k T(x_{n_k}) - x_{n_k} = 0$, and so, since T is continuous, $T(z) - z = 0$.

3.8 Proof of Nash's Theorem

Consider a "lazy player" who, given that all players are currently playing a mixed strategy profile σ, has utility for playing some mixed strategy σ'_i which is given by

$$g_i^\sigma(\sigma'_i) = u_i(\sigma_{-i}, \sigma'_i) - \lambda \|\sigma'_i - \sigma_i\|_2^2$$

for some $\lambda > 0$. That is, her utility has an extra addend which is lower the further away the new strategy is from her current strategy. Analyzing what happens when all players are lazy is the key to the following proof of Nash's Theorem.

[3] E.g., $T^3(x) = T(T(T(x)))$.

Proof of Theorem 3.18. Let $G = (N, \{S_i\}, \{u_i\})$ be a finite game. Let $f\colon \prod_i \Delta S_i \to \prod_i \Delta S_i$ be given by

$$[f(\sigma)]_i = \operatorname*{argmax}_{\sigma'_i \in \Delta S_i} g_i^{\sigma}(\sigma'_i).$$

It is straightforward to show that f is continuous. This maximum is unique, since g is strictly concave and ΔS_i is a convex set.

Since we can think of $\prod_i \Delta S_i$ as a convex subset of $\mathbb{R}^{|S|}$, f has a fixed point σ^*, by Brouwer's Theorem. We will show that σ^* is a mixed Nash equilibrium.

Assume by contradiction that there exists a player i and a $\sigma_i \in \Delta S_i$ such that

$$\delta := u_i(\sigma^*_{-i}, \sigma_i) - u_i(\sigma^*) > 0.$$

Given any $\varepsilon > 0$, let $\tau_i^\varepsilon = (1-\varepsilon)\cdot\sigma_i^* + \varepsilon\cdot\sigma_i$. Then

$$g_i^{\sigma^*}(\tau_i^\varepsilon) = u_i(\sigma^*_{-i}, \tau_i^\varepsilon) - \lambda\|\tau_i^\varepsilon - \sigma_i^*\|_2^2. \tag{3.1}$$

Now, by the linearity u_i,

$$\begin{aligned} u_i(\sigma^*_{-i}, \tau_i^\varepsilon) &= u_i(\sigma^*_{-i}, (1-\varepsilon)\cdot\sigma_i^* + \varepsilon\cdot\sigma_i) \\ &= (1-\varepsilon)\cdot u_i(\sigma^*) + \varepsilon\cdot u_i(\sigma^*_{-i}, \sigma_i) \\ &= u_i(\sigma^*) + \varepsilon\cdot\delta. \end{aligned}$$

By the definition of τ_i^ε

$$\|\tau_i^\varepsilon - \sigma_i^*\|_2^2 = \|\varepsilon\cdot(\sigma_i - \sigma_i^*)\|_2^2 = \varepsilon^2\|\sigma_i - \sigma_i^*\|_2^2.$$

Plugging these expressions back into (3.1) we get

$$g_i^{\sigma^*}(\tau_i^\varepsilon) = u_i(\sigma^*) + \varepsilon\cdot\delta - \varepsilon^2\lambda\|\sigma_i - \sigma_i^*\|_2^2,$$

which, for ε small enough, is greater than $u_i(\sigma^*)$. Hence

$$\sigma_i^* \neq \operatorname*{argmax}_{\sigma'_i \in \Delta S_i} g_i^{\sigma^*}(\sigma'_i),$$

and σ^* is not a fixed point of f - contradiction. □

3.9 Best responses

Let $G = (N, \{S_i\}, \{u_i\})$ be a finite game, and let σ be a mixed strategy profile in G. We say that $s_i \in S_i$ is a *best response* to σ_{-i} if, for all $t_i \in S_i$,

$$u_i(\sigma_{-i}, s_i) \geq u_i(\sigma_{-i}, t_i).$$

This notion can be naturally extended to mixed strategies.

The following proposition is helpful for understanding mixed equilibria.

Proposition 3.22. *Let $G = (N, \{S_i\}, \{u_i\})$ be a finite game, and let σ^* be a mixed Nash equilibrium of G. Then any s_i in the support of σ^*_i (i.e., any s_i to which σ^*_i assigns positive probability) is a best response to σ^*_{-i}.*

Proof. Suppose $s_i \in S_i$ is in the support of σ^*_i, but is not a best response to σ^*_{-i}, and let $t_i \in S_i$ be some best response to σ^*_{-i}. We will prove the claim by showing that t_i is a profitable deviation for i.

Let $C = u_i(\sigma^*_{-i}, t_i)$. Then $u_i(\sigma^*_{-i}, r_i) \le C$ for any $r_i \in S_i$, and $u_i(\sigma^*_{-i}, s_i) < C$. It follows that

$$\begin{aligned} u_i(\sigma^*) &= \sum_{r_i \in S_i} \sigma^*_i(r_i) u_i(\sigma^*_{-i}, r_i) \\ &< C, \end{aligned}$$

and so t_i is indeed a profitable deviation, since it yields utility C for i. □

It follows that if σ^* is an equilibrium then $u_i(\sigma^*_{-i}, s_i)$ is the same for every s_i in the support of σ^*_i. That is, i is indifferent between all the pure strategies in the support of her mixed strategy.

The following claim is also useful for calculating mixed equilibria. In its proof we use the fact that if the expectation of a random variable X is lower than x then with positive probability $X < x$.

Proposition 3.23. *Let $G = (N, \{S_i\}, \{u_i\})$ be a finite game, and let σ be a mixed strategy profile of G. Then σ is a mixed equilibrium if and only if, for every player i, every s_i in the support of σ_i is a best response to σ_{-i}.*

Proof. Assume that σ is a mixed equilibrium. Then the claim follows by Proposition 3.22.

Assume now that σ is not a mixed equilibrium. Then there exists a player i and a strategy σ'_i such that $u_i(\sigma_{-i}, \sigma'_i) > u_i(\sigma)$. Hence there exists a strategy s'_i such that $u_i(\sigma_{-i}, s'_i) > u_i(\sigma)$. Hence there is some t_i in the support of σ_i such that $u_i(\sigma_{-i}, s'_i) > u_i(\sigma_{-i}, t_i)$, and we have proved the claim. □

3.10 Trembling hand perfect equilibria

3.10.1 Motivating example

Recall the Cuban missile crisis game, but this time in strategic form. There are two players: the US and the USSR. The USSR has to decide whether to station missiles in Cuba (M) or not (N). If the USSR chooses N, the US does nothing. If the USSR chooses M, the US has to decide if to start a nuclear war (W) or peace (P). The payoff matrix in normal form is the following:

	W	P
M	$-1000,-1000$	$1,-1$
N	$0,0$	$0,0$

There are two equilibria to this game: (M,P) and (N,W). One way to eliminate the (N,W) equilibrium is to consider the extensive form of this game, and to note that it is not subgame perfect. Another approach would be to consider that each player might make a mistake with some very small probability. Assume that the players play (N,W), but that the USSR makes a mistake with some very small probability $\varepsilon > 0$ and plays M. Then the utility for the US for playing P is $-\varepsilon$, whereas the utility for playing W is -1000ε. Hence for any $\varepsilon > 0$, W is not a best response for N.

Consider now the equilibrium (M,P), and assume again that USSR errs with some probability $\varepsilon > 0$ and plays N. Then the utility for the US for playing W is $-1000\cdot(1-\varepsilon)$, while the utility for playing P is $-(1-\varepsilon)$, and so under any such perturbation P is a still a best response. From the point of view of the USSR, if the US errs and plays W with probability ε, the utility for playing M is $-1000\varepsilon+(1-\varepsilon) = 1-1001\varepsilon$, while the utility for playing N is 0. Thus, for ε small enough, M is still a best response.

In a *trembling hand perfect equilibrium* we require each strategy to be a best response if it still is a best response for some arbitrarily small error. The formal definition follows.

3.10.2 Definition and results

Let $G = (N,\{S_i\},\{u_i\})$ be a finite game, and let σ be a mixed strategy profile in G. We say that σ_i is completely mixed if it assigns positive probability to each $s_i \in S_i$.

A mixed strategy profile σ is called a *trembling hand perfect equilibrium* [32] if there exists a sequence of completely mixed strategy profiles $\{\sigma^k\}_{k\in\mathbb{N}}$ that converge to σ and such that, for each $i \in N$ and $k \in \mathbb{N}$, σ_i is a best response to σ^k_{-i}.

An equivalent definition is that σ is *not* a trembling hand perfect equilibrium if there is some $\varepsilon > 0$ such that for every completely mixed strategy τ with $|\tau-\sigma| < \varepsilon$ there is a player i such that σ_i is not a best response to τ.

Exercise 3.24. *Prove that these definitions are equivalent.*

We state without proof that every finite game has a trembling hand perfect equilibrium.

Exercise 3.25. *Show that every trembling hand perfect equilibrium is a Nash equilibrium.*

Claim 3.26. *Let s_i be a weakly dominated strategy of player i. Then $\sigma_i(s_i) = 0$ in any trembling hand perfect equilibrium σ.*

Proof. Let $\{\sigma^k\}_{k \in \mathbb{N}}$ be a sequence of completely mixed strategy profiles that converge to σ, and such that σ_i is a best response to σ^k_{-i} for all i and k.

Since s_i is weakly dominated then there exists an s_{-i} and a t_i such that

$$u_i(s_{-i}, s_i) < u_i(s_{-i}, t_i).$$

Since $\sigma^k(s_{-i}, s_i) > 0$, and since $u_i(s'_{-i}, s_i) \le u_i(s'_{-i}, t_i)$ for all s'_{-i}, it follows that s_i is not a best response to σ^k_{-i}, and thus $\sigma_i(s_i) = 0$. □

In fact, for finite two player games, if σ is a mixed Nash equilibrium in which every strategy is not weakly dominated (as a mixed strategy), then σ is trembling hand perfect. In general, this is not true for games with more than two players.

3.11 Correlated equilibria

3.11.1 Motivating example

Aumann [4] introduced the notion of a *correlated equilibrium*. Consider the following game, which is usually called "Chicken". There are two drivers who arrive at the same time to an intersection. Each one would like to drive on (strategy D) rather than yield (strategy Y), but if both drive then they run the risk of damaging the cars. If both yield time is wasted, but no egos are hurt. The payoff matrix is the following.

	Y	D
Y	$3,3$	$0,5$
D	$5,0$	$-4,-4$

This game has three Nash equilibria: two pure ((Y,D) and (D,Y)) and one mixed, in which each player drives with probability 1/3.

Exercise 3.27. *Show that this is indeed a mixed equilibrium.*

The players' expected utilities in these equilibria are $(0,5)$, $(5,0)$ and $(2,2)$.

A natural way to resolve this conflict is by the installation of a traffic light which would instruct each player whether to yield or drive. For example, the light could choose uniformly at random from (Y,D) and (D,Y). It is easy to convince oneself that a player has no incentive to disobey the traffic light, assuming that the other player is obeying it. The players' utilities in this case become $(2.5,2.5)$.

One could imagine a traffic light that chooses from $\{(Y,Y),(Y,D),(D,Y)\}$, where the first option is chosen with probability p and the second and third are each chosen with probability $(1-p)/2$. Now, given that a player is instructed to drive, she knows that the other player has been instructed to yield, and so, if we again assume that the other player is obedient, she has no reason to yield.

Given that a player has been instructed to yield, she knows that the other player has been told to yield with conditional probability $p_Y = p/(p+(1-p)/2)$ and to drive with conditional probability $p_D = ((1-p)/2)/(p+(1-p)/2)$. Therefore, her utility for yielding is $3p_Y$, while her utility for driving is $5p_Y - 4p_D$. It thus follows that she is not better off disobeying, as long as $3p_Y \geq 5p_Y - 4p_D$. A simple calculation shows that this is satisfied as long as $p \leq 1/2$.

Now, each player's expected utility is $3p + 5(1-p)/2$. Therefore, if we choose $p = 1/2$, the players' expected utilities are $(2.75,2.75)$. In this equilibrium the sum of the players' expected utilities is larger than in any Nash equilibrium.

3.11.2 Definition

We now generalize and formalize this idea. Let $G = (N,\{S_i\}_{i\in N},\{u_i\}_{i\in N})$ be a finite game. A distribution $\mu \in \Delta S$ is a *correlated equilibrium* if for every player i and every $s_i,t_i \in S_i$ it holds that

$$\sum_{s_{-i}} \mu(s_{-i}, s_i) u_i(s_{-i}, s_i) \geq \sum_{s_{-i}} \mu(s_{-i}, s_i) u_i(s_{-i}, t_i). \tag{3.2}$$

Player i's expected utility under a correlated equilibrium μ is simply $\mathbb{E}_{s\sim\mu}[u_i(s)]$.

Note that for given i and s_i, the condition in (3.2) is closed (i.e., if each of a converging sequence $\{\mu_n\}$ of distributions satisfies it then so does $\lim_n \mu_n$). Note also that if μ_1 and μ_2 satisfy (3.2) then so does any convex combination of μ_1 and μ_2. These observations immediately imply the following claim.

Claim 3.28. *The set of correlated equilibria is a compact, convex subset of ΔS.*

An advantage of correlated equilibria is that they are easier to calculate than Nash equilibria, since they are simply the set of solutions to a linear program. It is even easy to find a correlated equilibrium that maximizes (for example) the sum of the players' expected utilities, or indeed any linear combination of their utilities. Finding Nash equilibria can, on the other hand, be a difficult problem [18].

3.12 Zero-sum games

3.12.1 Motivating example

Consider the following game [28]. We write here only the utilities of the row player. The utilities of the column player are always minus those of the row player.

	J	1	2	3
J	$+1$	-1	-1	-1
1	-1	-1	$+1$	$+1$
2	-1	$+1$	-1	$+1$
3	-1	$+1$	$+1$	-1

Consider the point of view of the row player, and assume that she uses the mixed strategy σ_R in which J is played with some probability q and each of 1, 2 and 3 is played with probability $(1-q)/3$. Then

$$u_R(\sigma_R, J) = 2q-1 \quad \text{and} \quad u_R(\sigma_R,1) = u_R(\sigma_R,2) = u_R(\sigma_R,3) = 1/3-4q/3.$$

If we assume that the column player knows what q is and would do what is best for her (which here happens to be what is worst for the row player), then the row player would like to choose q that maximizes the minimum of $\{2q-1, 1/3-4q/3\}$. This happens when $q = 2/5$, in which case the row player's utility is $-1/5$, regardless of which (mixed!) strategy the column player chooses.

The same reasoning can be used to show that the column player will also choose $q = 2/5$. Her expected utility will be be $+1/5$. Note that this strategy profile (with both using $q = 2/5$) is a mixed Nash equilibrium, since both players are indifferent between all strategies.

3.12.2 Definition and results

A two player game $G = (\{1,2\}, \{S_i\}, \{u_i\})$ is called *zero-sum* if $u_1 + u_2 = 0$. For such games we drop the subscript on the utility functions and use $u := u_1$.

A general normal form game may have *cooperative* components (see, e.g., the battle of the sexes game), in the sense that moving from one strategy profile to another can benefit both players. However, zero-sum games are *competitive*: whatever one player gains the other loses. Hence a player may want to prepare herself for the worst possible outcome, namely one in which, given her own strategy, her opponent will choose the strategy yielding her the minimal utility. Hence an interesting quantity for a strategy σ_1 for player 1 is the guaranteed utility u_g when playing σ_1:

$$u_g(\sigma_1) = \min_{\sigma_2 \in \Delta S_2} u(\sigma_1, \sigma_2).$$

Continuing with this line of reasoning, player 1 will choose a strategy that maximizes her guaranteed utility. She would thus choose an action in[4]

$$\operatorname*{argmax}_{\sigma_1 \in \Delta S_1} u_g(\sigma_1) = \operatorname*{argmax}_{\sigma_1 \in \Delta S_1} \min_{\sigma_2 \in \Delta S_2} u(\sigma_1, \sigma_2).$$

[4]The argmax of an expression is in general a set, rather than a single value.

Any such strategy is called a *maxmin* strategy for player 1. It gives her the best possible guaranteed utility, which is

$$\max_{\sigma_1 \in \Delta S_1} u_g(\sigma_1) = \max_{\sigma_1 \in \Delta S_1} \min_{\sigma_2 \in \Delta S_2} u(\sigma_1, \sigma_2).$$

The next theorem shows that maxmin strategies and mixed Nash equilibria are closely related in zero-sum games.

Theorem 3.29 (Borel, 1921, von Neumann, 1928)**.** *Let G be a finite zero-sum game.*

1. *In every mixed Nash equilibrium σ^* each strategy is a maxmin strategy for that player.*
2. *There is a number $v \in \mathbb{R}$ such that*

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

and

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

The quantity v is called the *value* of G. It follows that $u(\sigma^*) = v$ for any equilibrium σ^*.

4 Knowledge and belief

4.1 Knowledge

4.1.1 The hats riddle

The hats riddle is a motivating example for studying knowledge. Consider n players, each of which is wearing a hat that is either red (r) or blue (b). The players each observe the others' hats, but do not observe their own.

An outside observer announces in the presence of all the players that "At least one of you has a red hat." They now play the following (non-strategic) game: a clock is set to ring every minute. At each ring, anyone who knows the color of their hat announces it, in which case the game ends. Otherwise the game continues.

Exercise 4.1. *What happens?*

We will formally analyze what transpires, after introducing some concepts and notation.

4.1.2 Knowledge spaces

Consider a situation in which there is uncertainty regarding one or more variables of interest. The set of all possible combinations of values of these variables is called the set of *states of the world*, so that knowing the state of the world is equivalent to knowing all there is to know and thus having no uncertainty.

We refer to a player's *type* as the information (or the type of information) this player has regarding the state of the world. Thus, for example, one player may know one variable of interest, another player may know another variable, and a third could know the sum of the first two variables. A fourth could just know whether the first variable is non-zero.

Formally, a *knowledge space*[5] is a tuple $(N, \Omega, \{T_i\}_{i \in N}, \{t_i\}_{i \in N})$ where

- N is a set of players.
- Ω is the space of the *states of the world*. We will assume here that Ω is finite.[6]
- T_i the space of possible types of player i.
- $t_i : \Omega \to T_i$ is player i's *private signal* or *type*.

We define the information partition function $P_i : \Omega \to 2^\Omega$ by

$$P_i(\omega) = \{\omega' : t_i(\omega') = t_i(\omega)\}.$$

That is, $P_i(\omega)$ is the set of states of the world for which player i has the same type as she does in ω. It is easy to show that $\omega \in P_i(\omega)$. The set $\{P_i(\omega)\}_{\omega \in \Omega}$ is easily seen to be a partition

[5] This terminology is non-standard.

[6] And treat infinite cases informally, without delving into measurability issues.

of Ω, and is usually called i's *information partition*. A knowledge space can thus be given as $(N,\Omega,\{P_i\}_{i\in N})$.

As an example, let Ω be the set of all possible outcomes of tossing a pair of dice:

$$\Omega = \{(k,\ell),: k,\ell = 1,\dots,6\}.$$

Let $N = \{1,2\}$, let $T_1 = \{1,2,\dots,6\}$ with

$$t_1(\omega) = t_1(k,\ell) = k$$

and let $T_2 = \{2,\dots,12\}$ with

$$t_2(\omega) = t_2(k,\ell) = k+\ell.$$

So player 1 sees the first die, and player 2 knows their sum.

As another example, consider the knowledge space associated to the hats riddle, before it is declared that at least one player has a red hat. Here, $\Omega = \{r,b\}^n$, $T_i = \{r,b\}^{n-1}$ and $t_i(\omega) = \omega_{-i}$.

4.1.3 Knowledge

Let $(N,\Omega,\{T_i\}_{i\in N},\{t_i\}_{i\in N})$ be a knowledge space. Let $A \in 2^\Omega$ be an event. We say that player i *knows* A at ω if $P_i(\omega) \subseteq A$, or, equivalently, if $t_i(\omega') = t_i(\omega)$ implies that $\omega' \in A$. Intuitively, when ω occurs, i does not know that - she only knows that $P_i(\omega)$ occurred. She thus knows that A occurred only if $P_i(\omega)$ is contained in A.

Given an event $A \in 2^\Omega$, let K_iA be the set of states of the world in which i knows A:

$$K_iA = \{\omega : P_i(\omega) \subseteq A\}.$$

Equivalently,

$$K_iA = \bigcup\{P_i(\omega) : P_i(\omega) \subseteq A\}.$$

Theorem 4.2 (Kripke's S5 system).

1. *$K_i\Omega = \Omega$. A player knows that some state of the world has occurred.*
2. *$K_iA \cap K_iB = K_i(A\cap B)$. A player knows A and a player knows B if and only if she knows A and B.*
3. Axiom of knowledge. *$K_iA \subseteq A$. If a player knows A then A has indeed occurred.*
4. Axiom of positive introspection. *$K_iK_iA = K_iA$. If a player knows A then she knows that she knows A.*
5. Axiom of negative introspection. *$(K_iA)^c = K_i((K_iA)^c)$. If a player does not know A then she knows that she does not know A.*

Proof. 1. This follows immediately from the definition.

2.

$$\begin{aligned}K_iA \cap K_iB &= \{\omega : P_i(\omega) \subseteq A\} \cap \{\omega : P_i(\omega) \subseteq B\}\\ &= \{\omega : P_i(\omega) \subseteq A, P_i(\omega) \subseteq B\}\\ &= \{\omega : P_i(\omega) \subseteq A \cap B\}\\ &= K_i(A \cap B).\end{aligned}$$

3. By definition, if $\omega \in K_iA$ then $P_i(\omega) \subseteq A$. Since $\omega \in P_i(\omega)$, it follows that $\omega \in A$. Hence $K_iA \subseteq A$.

4. By the previous part we have that $K_iK_iA \subseteq K_iA$.

   To see the other direction, let $\omega \in K_iA$, so that $P_i(\omega) \subseteq A$. Choose any $\omega' \in P_i(\omega)$. Hence $P_i(\omega') = P_i(\omega)$, and it follows that $\omega' \in K_iA$. Since ω' was an arbitrary element of $P_i(\omega)$, we have shown that $P_i(\omega) \subseteq K_iA$. Hence, by definition, $\omega \in K_iK_iA$.

5. The proof is similar to that of the previous part.

□

Interestingly, if a map $L: 2^\Omega \to 2^\Omega$ satisfies the Kripke S5 system axioms, then it is the knowledge operator for some type: there exists a type space T and a function $t: \Omega \to T$ such that L is equal to the associated knowledge operator given by

$$KA = \{\omega : P(\omega) \subseteq A\},$$

where $P(\omega) = t^{-1}(t(\omega))$.

4.1.4 Knowledge in terms of self-evident event algebras

An event $A \subset 2^\Omega$ is said to be *self-evident* to player i if $K_iA = A$. Let Σ_i be the collection of self evident events.

Claim 4.3. *1. A is self-evident if and only if it is the union of partitions elements $P_i(\omega)$.*

2. Σ_i is an algebra*:*

   (a) $\Omega \in \Sigma_i$.

   (b) If $S \in \Sigma_i$ then $S^c \in \Sigma_i$.

   (c) If $A, B \in \Sigma_i$ then $A \cup B \in \Sigma_i$.

As an alternative to specifying the partition elements $P_i(\omega)$, one can specify the (self-evident) event algebras Σ_i. We can recover $P_i(\omega)$ from Σ_i by

$$P_i(\omega) = \cap\{S \in \Sigma_i : \omega \in S\}.$$

Accordingly, we will occasionally define knowledge spaces using these event algebras only. That is, a knowledge space will be given by a tuple $(N, \Omega, \{\Sigma_i\}_{i \in N})$.

Proposition 4.4. *Let $A \in 2^{\Omega}$. Then $K_iA \in \Sigma_i$, and in particular*

$$K_iA = \cup\{S \in \Sigma_i : S \subseteq A\}.$$

Since any event algebra is closed to unions, it follows that K_iA is the largest element of Σ_i that is contained in A:

$$K_iA = \max\{S \in \Sigma_i : S \subseteq A\}.$$

Hence K_iA is always self-evident. This makes the proof of the axioms of introspection and negative introspection immediate.

It is important to note that we can take this to be a definition of K_i. The advantage of this definition is that it is entirely in terms of our newly defined algebras $\{\Sigma_i\}$.

Exercise 4.5. *Given a knowledge space $B = (N, \Omega, \{\Sigma_i\}_{i\in N})$ find a knowledge space $B' = (N, \Omega, \{T_i\}_{i\in N}, \{t_i\}_{i\in N})$ such that each Σ_i is the algebra generated by player i's information partition in B'.*

4.1.5 Common knowledge

Let $(N,\Omega,\{T_i\}_{i\in N},\{t_i\}_{i\in N})$ be a finite knowledge space. An event A is said to be *common knowledge* at $\omega\in\Omega$ if for any sequence $i_1,i_2,\ldots,i_k\in N$ it holds that

$$\omega\in K_{i_1}K_{i_2}\cdots K_{i_k}A.$$

We will give two alternative ways of thinking about common knowledge events. In order to introduce these we will need a few definitions.

Let Σ,Π be two sub-algebras of some algebra. We say that Σ is a *refinement* of Π if $\Pi\subseteq\Sigma$. In this case we also say that Π is a *coarsening* of Σ. In terms of the partitions that generate these algebras, this is the same as saying that each of the information sets that generates Π is a union of the corresponding sets associated with Σ. In terms of knowledge, this means that if agents i and j have algebras Π and Σ then agent j is better informed than agent i in the sense that if i knows A at ω then so does j.

The *meet* of two algebras is $\Sigma_1,\Sigma_2\subseteq\Sigma$ is the finest sub-algebra of Σ that is a coarsening of each Σ_i. Their *join* is the coarsest sub-algebra of Σ that is a refinement of each Σ_i.

Exercise 4.6. *Show that the meet of Σ_1 and Σ_2 is their intersection $\Sigma_1\cap\Sigma_2$. Show that their join is the algebra generated by their union.*

In terms of information, knowing the join is the same as having the information in both. Knowing the meet captures something more complicated, which we will now explore. Let $\Sigma_C=\cap_i\Sigma_i$ be the meet of the player's algebras.

Claim 4.7. *The following are equivalent:*

1. $C\in\Sigma_C$.
2. $K_iC=C$ *for all* $i\in N$.

Proof. We note first that $K_iA=A$ iff $A\in\Sigma_i$. This follows from Proposition 4.4.

Hence $K_iC=C$ for all i iff $C\in\Sigma_i$ for all i iff $C\in\Sigma_C$, as $\Sigma_c=\cap_i\Sigma_i$. □

Recall that, by Proposition 4.4,

$$K_iA=\cup\{S\in\Sigma_i : S\subseteq A\}=\max\{S\in\Sigma_i : S\subseteq A\}.$$

Analogously we define

$$K_CA=\cup\{S\in\Sigma_C : S\subseteq A\}=\max\{S\in\Sigma_C : S\subseteq A\}.$$

As we show in Theorem 4.8 below, K_CA is the set of states of the world in which A is common knowledge.

Another way of thinking about common knowledge is through the undirected graph G_C, whose vertices are the elements of Ω, and where there is an edge between $\omega,\omega'\in\Omega$ if there exists a player i such that $P_i(\omega)=P_i(\omega')$. Let $C(\omega)$ be the *connected component* of ω; that is the maximal set of vertices in the graph for which there exists a path from ω. Here, a path from ω to ω' is a sequence of edges $(\omega_1,\omega_2),(\omega_2,\omega_3),(\omega_3,\omega_4),\ldots,(\omega_{k-1},\omega_k)$ such that $\omega_1=\omega$ and $\omega_k=\omega'$.

Theorem 4.8. *Given a finite knowledge space $(N, \Omega, \{\Sigma_i\}_{i \in N})$, a state of the world $\omega \in \Omega$ and an event A, the following are equivalent.*

1. *A is common knowledge at ω.*
2. *$C(\omega) \subseteq A$.*
3. *$\omega \in K_C A$.*

Proof. We first show that (1) implies (2). Choose an arbitrary $\omega' \in C(\omega)$; we will show that $\omega' \in A$. Since $\omega' \in C(\omega)$, there exists a path $(\omega_1, \omega_2), (\omega_2, \omega_3), \ldots, (\omega_{k-1}, \omega_k)$ such that $\omega_1 = \omega$ and $\omega_k = \omega'$. There exists therefore a sequence of players $i_1, \ldots, i_{k-1}$ such that $P_{i_j}(\omega_j) = P_{i_j}(\omega_{j+1})$.

Now, by the common knowledge assumption

$$\omega = \omega_1 \in K_{i_1} K_{i_2} \cdots K_{i_{k-1}} A.$$

Hence

$$P_{i_1}(\omega_1) \subseteq K_{i_2} \cdots K_{i_{k-1}} A$$

and in particular

$$\omega_2 \in K_{i_2} \cdots K_{i_{k-1}} A,$$

since $\omega_2 \in P_{i_1}(\omega_1)$. Applying this argument inductively yields $\omega' = \omega_k \in A$, and thus we have shown that $C(\omega) \subseteq A$.

To show that (2) implies (3) we first prove that $C(\omega) \in \Sigma_C$. To this end it suffices to show that $C(\omega) \in \Sigma_i$ for all $i \in N$. Let $\omega' \in C(\omega)$. Then, by the definition of the graph G, $P_i(\omega') \subseteq C(\omega)$. Hence $C(\omega)$ is a union of sets of the form $P_i(\omega')$, and thus it is an element of Σ_i.

Now, assuming (2), $\omega \in C(\omega) \subseteq A$. Therefore, by the definition of K_C, $\omega \in K_C A$.

Finally, we show that (3) implies (1). Because $\omega \in K_C$ then there is some $C \in \Sigma_C$ such that $\omega \in C$ and $C \subseteq A$. Let $i_1, i_2, \ldots, i_k$ be an arbitrary sequence of players. We first prove that

$$\omega \in K_{i_1} K_{i_2} \cdots K_{i_k} C.$$

Indeed, by Claim 4.7,

$$K_{i_1} K_{i_2} \cdots K_{i_k} C = C,$$

and since $\omega \in C$, we have shown that

$$\omega \in K_{i_1} K_{i_2} \cdots K_{i_k} C.$$

Since $C \subseteq A$,

$$K_{i_1} K_{i_2} \cdots K_{i_k} C \subseteq K_{i_1} K_{i_2} \cdots K_{i_k} A$$

and thus we also have that

$$\omega \in K_{i_1} K_{i_2} \cdots K_{i_k} A.$$

$\square$

4.1.6 Back to the hats riddle

Recall the knowledge space associated to the hats riddle (before the declaration that there is at least one red hat): $\Omega = \{r,b\}^n$, $T_i = \{r,b\}^{n-1}$ and $t_i(\omega) = \omega_{-i}$.

The information partition functions are

$$P_i(\omega) = \{(\omega_{-i}, r), (\omega_{-i}, b)\}.$$

Thus the set of edges of the graph G_C is exactly given by

$$E = \{P_i(\omega)\}_{i,\omega} = \left\{\{(\omega_{-i}, r), (\omega_{-i}, b)\} : \omega_{-i} \in \{r,b\}^{n-1}\right\}.$$

Claim 4.9. $\Sigma_C = \{\emptyset, \Omega\}$.

Proof. It is easy to see that for every $\omega, \omega' \in \Omega$ there is a sequence of edges (corresponding to the coordinates in which they differ) that forms a path from ω to ω' in G_C. Thus $C(\omega) = \Omega$, and the events that can be common knowledge are Ω and $\emptyset$. □

In particular in no state is it common knowledge that there is at least one red hat, even when everyone has one.

4.2 Beliefs

A *belief* of a player over a finite space Ω is a probability measure on Ω.

4.2.1 A motivating example

Beliefs are useful for modeling uncertainty. For example, consider a monopoly seller who would like to sell a single object to a single buyer. The value of the product to the buyer is v. The sellers sets a price p; hence her strategy space is $\mathbb{R}$. The buyer learns p and then decides to either buy or not to buy. Thus her strategies can each be characterized by a set $B \subset \mathbb{R}$ of the prices in which she decides to buy.

If the buyer buys then her utility is $v-p$, and the seller's utility is p. Otherwise both have zero utility. Note that this is an (infinite) extensive form game with perfect information.

Exercise 4.10. *Prove that if the seller knows v then a subgame perfect equilibrium is for the seller to choose $p = v$ and for the buyer to buy if and only if $p \leq v$.*

We would like to model the case in which the seller does not know v exactly. One way to do this is to assume that the seller has some *belief* μ over v. That is, she believes that v is chosen at random from μ. Given this belief, she now has to choose a price p. The seller first learns v. She then learns p and has to decide whether or not to buy. In a subgame perfect equilibrium, she will buy whenever $v \geq p$, or whenever $v > p$. The buyer's utility is now the expected price that the seller paid, given that she bought, and where expectations are taken with respect to μ. The buyer's utility is, as before, $v - p$ if she buys and zero otherwise.

Assume that the seller's belief can be represented by a probability distribution function $f : [0,\infty) \to \mathbb{R}$, so that the probability that $v \in [a,b)$ is $\int_a^b f(v)\mathrm{d}v$. In this case it will not matter whether the buyer's strategy is to buy whenever $p \leq v$ or whenever $p < v$; the following calculation will be the same in both cases, since the seller's belief is that probability that $p = v$ is zero. Then the seller's utility is

$$u_S(p) = p \cdot \int_p^\infty f(v)\mathrm{d}v,$$

which we can write as $p \cdot (1 - F(p))$, where $F(x) = \int_0^x f(x)\mathrm{d}x$ is the cumulative distribution function associated with f. To find the seller's best response we need to find a maximum of u_S. Its derivative with respect to p is

$$\frac{\mathrm{d}u_S}{\mathrm{d}p} = 1 - F(p) - p \cdot f(p).$$

Hence in any maximum it will be the case that

$$p = \frac{1 - F(p)}{f(p)}.$$

It is easy to show that if this equation has a single solution then this solution is a maximum. In this case p is called the *monopoly price*.

Exercise 4.11. *Let*

$$f(v) = \begin{cases} 0 & \text{if } v < v_0 \\ \frac{v_0}{v^2} & \text{if } v \geq v_0 \end{cases}$$

for some v_0. Calculate $u_S(p)$. Under the seller's belief, what is the buyer's expected utility as a function of p?

Exercise 4.12. *Let $f(v) = \frac{1}{v_0} \cdot e^{-v/v_0}$ for some $v_0 > 0$. Calculate the monopoly price. Under the seller's belief, what is the buyer's expected utility as a function of p?*

4.2.2 Belief spaces

A *belief space* is a tuple $(N, \Omega, \{\mu_i\}_{i\in N}, \{T_i\}_{i\in N}, \{t_i\}_{i\in N})$ where

- $(N, \Omega, \{T_i\}_{i\in N}, \{t_i\}_{i\in N})$ is a knowledge space.
- μ_i is player i's belief over Ω.

We denote by $\mathbb{P}_i[A|t_i]$ the probability, under μ_i, of the event A, conditioned on t_i. When the underlying knowledge space is given in terms of algebras, we write this as $\mathbb{P}_i[A|\Sigma_i]$

When Ω is finite, this is given by

$$\mathbb{P}_i[A|\Sigma_i](\omega) = \frac{\mu_i(A \cap P_i(\omega))}{\mu_i(P_i(\omega))}. \tag{4.1}$$

Note that $\mathbb{P}_i[A|\Sigma_i]$ is a *function* from Ω to $\mathbb{R}$.

If the different μ_i's are equal then we say that the players have a *common prior*. This will be our standing assumption.

4.3 Rationalizability

Rationalizability tries to formalize the idea of *common knowledge of rationality*. Let $G = (N, \{S_i\}_{i\in N}, \{u_i\}_{i\in N})$ be a normal form game, with $S = \prod_i S_i$, as usual. Let $\Omega = S$, let $T_i = S_i$ and let $t_i(s) = s_i$. Consider belief spaces of the form $(N, \Omega, \{\mu_i\}_{i\in N}, \{T_i\}_{i\in N}, \{t_i\}_{i\in N})$. What this captures is a situation in which each player has a belief regrading the strategies used by the others, and knows her own strategy.

We assume that all measures μ_i are product measures. That is, if $s = (s_1, \dots, s_n)$ is chosen from μ_i, then $(s_1, \dots, s_n)$ are independent.

We say that μ_i is best responding if conditioned on t_i, $t_i(s)$ maximizes expected utility, where expectations are taken according to μ_i, conditioned on $t_i(s)$. Note that if $\mu_i = \mu$ for all i, then μ is a Nash equilibrium if and only if all agents are best responding. In this case, however, we do not require that all agents have the same prior belief.

We say that μ_i is k-rationalizable if it is best responding, and if for every $j \neq i$ there is a μ'_j that is $k-1$-rationalizable, and such that the distribution of s_j under μ'_j is identical to the distribution of s_j under μ_i. We define the set of 0-rationalizable μ_i to be all possible μ_i. Finally, we say that μ_i is rationalizable if it is k-rationalizable for all k.

Exercise 4.13. *What is the relation to Nash equilibria?*

Exercise 4.14. *What is the relation to iterated removal of strictly dominated strategies?*

Exercise 4.15. *Does this formally capture "common knowledge of rationality"?*

4.4 Agreeing to disagree

Let $(N, \Omega, \{\mu_i\}_{i \in N}, \{\Sigma_i\}_{i \in N})$ be a finite belief space with common priors, so that $\mu_i = \mu$ for all i. Given a random variable $X \colon \Omega \to \mathbb{R}$, we denote by $\mathbb{E}[X]$ the expectation of X according to μ. The expectation of X conditioned on player i's information at ω is

$$\mathbb{E}[X|\Sigma_i](\omega) = \frac{\sum_{\omega' \in P_i(\omega)} \mu(\omega') X(\omega')}{\sum_{\omega' \in P_i(\omega)} \mu(\omega')}$$

Aumann [3] proved the following theorem.

Theorem 4.16 (Aumann's Agreement Theorem, 1976)**.** *Let $X \colon \Omega \to \mathbb{R}$ be a random variable. Suppose that at ω_0 it is common knowledge that $\mathbb{E}[X|\Sigma_i] = q_i$ for $i = 1, \ldots, n$ and some $(q_1, q_2, \ldots, q_n) \in \mathbb{R}^n$. Then $q_1 = q_2 = \cdots = q_n$.*

Note that we implicitly assume that the conditional expectations $\mathbb{E}[X|\Sigma_i]$ are well defined everywhere. Before proving the Theorem we will recall the law of total expectation. Let S be an element of an algebra $\Pi \subseteq 2^{\Omega}$, and let $X \colon \Omega \to \mathbb{R}$ be a random variable. Then

$$\mathbb{E}[X|S] = \mathbb{E}[\mathbb{E}[X|\Pi]|S]. \tag{4.2}$$

Exercise 4.17. *Prove the law of total expectation for a finite probability space. Hint: write C as a disjoint union of elements of the partition that generates Π: $C = \cup_j P_j$, with $P_j \in \Pi$ being elements with no proper, non-empty subsets in Π.*

Proof of Theorem 4.16. Let A be the event that for $i = 1, \ldots, n$ it holds that $\mathbb{E}[X|\Sigma_i] = q_i$.

By the common knowledge hypothesis there is a $C \in \Sigma_C = \cap_i \Sigma_i$ such that $\omega_0 \in C \subset A$. Hence $\mathbb{E}[X|\Sigma_i](\omega) = q_i$ for all $\omega \in C$. Thus

$$\mathbb{E}[X|C] = \mathbb{E}[\mathbb{E}[X|\Sigma_i]|C] = \mathbb{E}[q_i|C] = q_i,$$

and q_i is independent of i. $\square$

In his paper, Aumann stated this theorem for the case that X is the indicator of some event:

Corollary 4.18. *If two players have common priors over a finite space, and it is common knowledge that their posteriors for some event are q_1 and q_2, then $q_1 = q_2$.*

4.5 No trade

Milgrom and Stokey [24] apply Aumann's theorem to show that rational agents with common priors can never agree to trade. Here we give a theorem that is less general than their original.

Consider two economic agents. The first one has an indivisible good that she might be interested to sell to the second. This good can be sold tomorrow at an auction for an unknown price that ranges between \$0 and \$1,000, in integer increments. Let Ω be some subset of $T_1 \times T_2 \times T$, and let the common prior be some μ. Here $T = \{0, \$1, \ldots, \$1000\}$ represents the auction price of the good, and $T_1 \times T_2$ is some finite set that describes many possible events that may influence the price. Accordingly, μ is not a product measure, so that conditioning on different $(\omega_1, \omega_2) \in T_1 \times T_2$ yields different conditional distributions on T. Let the players' types be given by $t_i(\omega_1, \omega_2, \omega_3) = \omega_i$ for $i = 1, 2$. Let $P : \Omega \to T$ be the auction price $p(\omega_1, \omega_2, \omega_3) = \omega_3$.

We denote by Q_1 player 1's conditional expected price:

$$Q_1 = \mathbb{E}[P|t_1]$$

Analogously,

$$Q_2 = \mathbb{E}[P|t_2].$$

For example, let $T_1 = T_2 = T$ and μ be the uniform distribution over $\{(\omega_1, \omega_2, \omega_3) : \omega_3 = \omega_1 \text{ or } \omega_3 = \omega_2\}$. Equivalently, pick ω_3 uniformly at random, and then w.p. 1/2 set $\omega_1 = k$ and choose ω_2 uniformly at random, and w.p. 1/2 set $\omega_2 = k$ and choose ω_1 uniformly at random. Then

$$Q_i = \frac{1}{2}\left(\$500 + t_i\right).$$

Thus in any state ω in which $t_1 < t_2$ we have, for $q = \frac{1}{2}(t_1 + t_2)$, that $Q_1(\omega) < q < Q_2(\omega)$, so both players expect a positive return for trading the good for q dollars. However, note that the players do not know that the other player also has positive expectation. What if they knew that the other person is willing to trade?

Theorem 4.19 (Milgrom and Stokey, 1982)**.** *If at some $\omega \in \Omega$ it is common knowledge that $Q_1 \le q \le Q_2$ then $Q_1(\omega) = q = Q_2(\omega)$.*

Proof. Let A be the event that $Q_1 \le q \le Q_2$, i.e., A is the event that $\mathbb{E}[P|\Sigma_1] \le q \le \mathbb{E}[P|\Sigma_2]$. Then A is common knowledge at ω, and so there is an event $C \in \Sigma_C$ with $\omega \in C \subset A$. Hence $C \in \Sigma_1$, and so, by (4.2),

$$\mathbb{E}[P|C] = \mathbb{E}[\mathbb{E}[P|\Sigma_1]|C] \le \mathbb{E}[q|C] = q.$$

By the same argument we also get that

$$\mathbb{E}[P|C] = \mathbb{E}[\mathbb{E}[P|\Sigma_2]|C] \ge \mathbb{E}[q|C] = q.$$

Hence these are both equalities, and thus, on C (and therefor at ω), $\mathbb{E}[P|\Sigma_1] = \mathbb{E}[P|\Sigma_2] = q$. ☐

Exercise 4.20. *Construct an example in which, for some $\omega \in \Omega$, $Q_1(\omega) < q$ and $Q_2(\omega) > q$, player 1 knows that $Q_2(\omega) > q$ and player 2 knows that $Q_1(\omega) < q$.*

4.6 Reaching common knowledge

Geanakoplos and Polemarchakis [15] show that repeatedly communicating posteriors leads agents to convergence to a common posterior, which is then common knowledge. We state this theorem in somewhat greater generality than in the original paper, requiring a more abstract mathematical formulation.

Theorem 4.21 (Geanakoplos and Polemarchakis, 1982)**.** *Let $(\Omega, \Sigma, \mathbb{P})$ be a probability space, and fix $A \in \Sigma$. Let X_1 and X_2 be two random variables on this space. Denote $Q_1^0 = \mathbb{P}[A|X_1]$ and $Q_2^0 = \mathbb{P}[A|X_2]$. For $t \in \{1, 2, \dots\}$ let*

$$Q_1^t = \mathbb{P}\left[A \middle| X_1, Q_2^0, Q_2^1, \dots, Q_2^{t-1}\right]$$

and

$$Q_2^t = \mathbb{P}\left[A \middle| X_2, Q_1^0, Q_1^1, \dots, Q_1^{t-1}\right].$$

Then $\lim_t Q_1^t$ and $\lim_t Q_2^t$ almost surely exist and are equal.

We first explain why this Theorem holds for finite belief spaces.

Proof for finite Ω. Assuming Ω is finite, let Σ_1^t be the algebra describing the knowledge of player 1 at time t: this is the set of unions of subsets of Ω on which the random variable $(X_1, Q_2^1, \dots, Q_2^{t-1})$ is constant. Define $\{\Sigma_2^t\}$ analogously. Then

$$Q_1^t = \mathbb{P}\left[A \middle| \Sigma_1^t\right]$$

and

$$Q_2^t = \mathbb{P}\left[A \middle| \Sigma_2^t\right].$$

Note that each sequence $(\Sigma_i^1, \Sigma_i^2, \dots)$ is increasing, in the sense that Σ_i^{t+1} contains Σ_i^t. Since Ω is finite, both must stabilize at some T. At this T, by definition, Q_1^T and Q_2^T are constant on each partition $P_1^T(\omega)$ and on each partition $P_2^T(\omega)$. Thus, at every ω, player 1 knows (Q_1^T, Q_2^T), and thus it is common knowledge that she does. The same holds for player 2. Hence the players posteriors are common knowledge at time T, and thus must be identical, by Theorem 4.16. □

To prove this theorem in its full generality we will need the following classical result in probability.

Theorem 4.22 (Lévy's Zero-One Law)**.** *Let $\{\Pi_n\}$ be a filtration, and let X be a bounded random variable. Let Π_∞ be the sigma-algebra generated by $\cup_n \Pi_n$. Then $\lim_n \mathbb{E}[X|\Pi_n]$ exists for almost every $\omega \in \Omega$, and is equal to $\mathbb{E}[X|\Pi_\infty]$.*

Proof of Theorem 4.21. Note that by Lévy's zero-one law (Theorem 4.22,

$$Q_1^\infty := \lim_t Q_1^t = \mathbb{P}\left[A \middle| X_1, Q_2^0, Q_2^1, \ldots\right]$$

and an analogous statement holds for $\lim_t Q_2^t$. Let Σ_1 be the sigma-algebra generated by $\{X_1, Q_2^0, Q_2^1, \ldots\}$, so that

$$Q_1^\infty = \mathbb{P}[A|\Sigma_1].$$

Since Q_2^∞ is Σ_1-measurable, it follows that

$$Q_1^\infty = \mathbb{P}[A|\Sigma_1] = \mathbb{E}\left[\mathbb{P}\left[A \middle| Q_2^\infty\right] \middle| \Sigma_1\right] = \mathbb{E}\left[Q_2^\infty \middle| \Sigma_1\right] = Q_2^\infty.$$

$\square$

4.7 Bayesian games

The buyer-seller game described in Section 4.2.1 is an example of a *Bayesian game*. In these games there is uncertainty over the payoffs to the players in given pure strategy profiles.

A Bayesian game is a tuple $G = (N, \{A_i\}_{i\in N}, \Omega, \{T_i\}_{i\in N}, \{t_i\}_{i\in N}, \{\mu_i\}_{i\in N}, \{u_i\}_{i\in N})$ where

- $(N, \Omega, \{T_i\}_{i\in N}, \{t_i\}_{i\in N}, \{\mu_i\}_{i\in N})$ is a belief space.
- A_i is the set of *actions* of player i.
- $u_i \colon A \times \Omega \to \mathbb{R}$ is player i's utility function.

The set of *pure strategies* of player i is the set of functions from T_i to A_i. That is, a strategy of a player is a choice of action, given her private signal realization. Given a strategy profile $(s_1, \dots, s_n)$, player i's expected utility is

$$\mathbb{E}_{\mu_i,s}[u_i] = \int_\Omega u_i(s_1(t_1(\omega)), \dots, s_n(t_n(\omega)), \omega)\mathrm{d}\mu_i(\omega).$$

A *Bayes-Nash equilibrium* is a strategy profile in which no player can improve her expected utility by changing her strategy. That is, for any player i and strategy s_i' it holds that

$$\mathbb{E}_{\mu_i,s}[u_i] \geq \mathbb{E}_{\mu_i,(s_{-i},s_i')}[u_i]. \tag{4.3}$$

An alternative definition of a Bayes-Nash equilibrium is a strategy profile in which, for each player i and each type $\tau_i \in T_i$ it holds that

$$\mathbb{E}_{\mu_i,s}[u_i | t_i = \tau_i] = \int_\Omega u_i(s_1(t_1(\omega)), \dots, s_i(\tau_i), \dots, s_n(t_n(\omega)), \omega)\mathrm{d}\mu_i(\omega | t_i = \tau_i) \tag{4.4}$$

cannot be improved:

$$\mathbb{E}_{\mu_i,s}[u_i | t_i = \tau_i] \geq \mathbb{E}_{\mu_i,(s_{-i},s_i')}[u_i | t_i = \tau_i]$$

for all s_i' and $\tau_i \in T_i$.

This is not an equivalent definition, but the second is stronger than the first.

Proposition 4.23. (4.4) *implies* (4.3).

Proof. Let s satisfy (4.4). Then by the law of total expectation

$$\begin{aligned}
\mathbb{E}_{\mu_i,s}[u_i] &= \mathbb{E}_{\mu_i,s}\left[\mathbb{E}_{\mu_i,s}[u_i | t_i]\right] \\
&\geq \mathbb{E}_{\mu_i,s}\left[\mathbb{E}_{\mu_i,(s_{-i},s_i')}[u_i | t_i]\right] \\
&= \mathbb{E}_{\mu_i,(s_{-i},s_i')}[u_i].
\end{aligned}$$

□

Conversely, if Ω is finite and there is a common prior $\mu = \mu_1 = \cdots = \mu_n$ then, if s satisfies (4.3) then there is an s' that satisfies (4.4), and such that the probability (under μ) that $s_i(t_i) \neq s_i'(t_i)$ is zero.

Exercise 4.24. *Find a finite Bayesian game that has a (pure) strategy profile that satisfies* (4.3) *but does not have one that satisfies* (4.4).

5 Auctions

Auctions have been used throughout history to buy and sell goods. They are still today very important in many markets, including on-line markets that run massive computerized auctions.

5.1 Classical auctions

In this section we will consider n players, each of which have a fixed valuation v_i for some item that is being auctioned. We assume that each v_i is a non-negative integer. Furthermore, to avoid having to deal with tie-breaking, we assume that each $v_i = i+1 \mod n$. We also assume without loss of generality that $v_1 > v_2 > \cdots > v_n$.

If it is agreed that player i buys the item for some price p then that player's utility is $v_i - p$. If a player does not buy then she pays nothing and her utility is zero.

We will consider a number of possible auctions.

5.1.1 First price, sealed bid auction

In this auction each player submits a bid b_i, which has to be a non-negative integer, congruent to $i \mod n$. Note that this means that a player cannot ever bid her valuation (which is congruent to $i+1 \mod n$), but can bid one less than her valuation[7].

For example, consider the case that $n = 2$. Then possible valuations are $v_1 = 10$ and $v_2 = 5$, and b_1 must be odd while b_2 must be even.

The bids $b = (b_1, \ldots, b_n)$ are submitted simultaneously. The player who submitted the highest bid $b_{\max}(b) = \max_i b_i$ buys the item, paying $b_{\max}$.

Hence player i's utility for strategy profile b is given by

$$u_i(b) = \begin{cases} v_i - b_i & \text{if } b_i = b_{\max}(b) \\ 0 & \text{otherwise} \end{cases}.$$

We now analyze this game. We first note that $b_i = v_i - 1$ guarantees utility at least 0. Next, we note that any $b_i > v_i$ is weakly dominated by $b_i = v_i - 1$, since it guarantees utility at most 0, but can result in negative utility if $b_i = b_{\max}$. Furthermore, it is impossible that in a pure equilibrium the winner of the auction bid more than v_i, since then she could increase her utility by lowering her bid to $v_i - 1$.

Assume that b^* is an equilibrium.

Claim 5.1. *Player 1 wins the auction: $b_1^* = b_{\max}^*$.*

Proof. Assume by contradiction that player $i > 1$ wins the auction. As we noted above, $b_i^* \le v_i - 1$. Hence $b_{\max}^* = b_i^* \le v_i - 1 < v_1 - 1$. Hence player 1 could improve her utility to 1 by bidding $v_1 - 1$ and winning the auction. $\square$

[7]Think of this as one cent less.

We have thus shown that in any equilibrium the first player wins. It thus remains to show that one exists.

Claim 5.2. *Let b_1^* be the smallest allowed bid*[8] *that is larger than $v_2 - 1$. Let $b_2 = v_2 - 1$. For $i > 2$ (if there are more than 2 players) let b_i^* be any allowed bid that is less than v_2. Then b^* is an equilibrium.*

Exercise 5.3. *Prove Claim 5.2.*

We note a few facts about this equilibrium.

- The item was allocated to the player who values it the most.
- The player who won did not base her bid on her own valuation, but on the other players' valuations, and in particular on the second highest one.

Note that other equilibria exist. For example, if $n = 2$ and $v_1 = 10$ and $v_2 = 5$ then $b_1 = 9$ and $b_2 = 8$ is again an equilibrium. Player 2 gets zero payoff, but can only decrease her utility by raising her price and winning the auction. Player 1 gets positive utility (1), but cannot improve it by lowering her bid.

5.1.2 Second price, sealed bid auction

In this auction each player again submits a bid b_i, which this time has to be a non-negative integer, congruent to $i+1 \mod n$; that is, it *can* be equal to v_i. Again, the player who submitted the highest bid $b_{\max}$ wins. However, in this case she does not pay her bid, but rather the second highest bid b_{nd}. Hence

$$u_i(b) = \begin{cases} v_i - b_{\text{nd}} & \text{if } b_i = b_{\max}(b) \\ 0 & \text{otherwise} \end{cases}.$$

As in the first price auction, any $b_i > v_i$ is weakly dominated by $b_i = v_i$: if $b_i = v_i$ is a winning bid then so is $b_i > v_i$, and the same price is paid. If $b_i = v_i$ is not a winning bid then $b_i > v_i$ might still result in a loss (and the same zero utility), or else would result in a win but a negative utility, since the price paid would be more than v_i.

Moreover, in this auction any $b_i < v_i$ is also weakly dominated by $b_i = v_i$. To see this, let b' be the highest bid of the rest of the players. If $b' > v_i$ then in either bid the player losses the auction, and so both strategies yield zero. If $b' < v_i$ then bidding $b_i < v_i$ may either cause the loss of the auction and utility zero (if $b_i < b'$) or otherwise gaining $v_i - b'$. But bidding $b_i = v_i$ guarantees utility $v_i - b'$.

Hence $b_i = v_i$ is a weakly dominant strategy, and so this is an equilibrium. Auctions in which bidding your valuation is weakly dominant are called *strategy proof* or sometimes *truthful*.

Note that in this equilibrium the item is allocated to the player who values it the most, as in the first price auction. However, the player based her bid on her own valuation, independently of the other player's valuations.

[8] Congruent to $1 \mod n$.

5.1.3 English auction

This auction is an extensive form game with complete information. The players take turns, starting with player 1, then player 2 and so on up to player n, and then player 1 again etc. Each player can, at her turn, either leave the auction or stay in. Once a player has left she must choose to leave in all the subsequent turns.

The auction ends when all players but one have left the auction. If this happens at round t then the player left wins the auction and pays $t-1$.

Claim 5.4. *There is a subgame perfect equilibrium of this game in which each player i stays until period $t = v_i$ and leaves once $t > v_i$.*

Exercise 5.5. *Prove Claim 5.4.*

Exercise 5.6. *What is the relation between this English auction and the second price auction?*

5.1.4 Social welfare

Imagine that the person running the auction is also a player in the game. Her utility is simply the payment she receives; she has no utility for the auctioned object. Then the social welfare, which we will for now define to be the sum of all the players' utilities, is equal to the utility of the winner — her value minus her payment — plus the utility of the losers (which is zero), plus the utility of the auctioneer, which is equal to the payment. This sum is the value of the object to the winner. Hence social welfare is maximized when the winner is a person who values the object most.

5.2 Bayesian auctions

In this section we will consider auctions in which the players do not know the others' valuations exactly. Specifically, the auctions will be Bayesian games with common priors.

We will again have n players. Each player's type will be her valuation v_i, and the players will have some common prior $\mathbb{P}$ over $(v_1,\dots,v_n)$. Formally, the belief space will be $((\mathbb{R}^+)^n,\Sigma,\mathbb{P})$, where Σ is the Borel sigma-algebra, and $\mathbb{P}$ is some probability distribution. Player i's type t_i is given by $t_i(v_1,\dots,v_n) = v_i$.

As before, if a player does not win the auction she has utility zero. Otherwise, assuming she pays a price p, she has utility $v_i - p$. Note that the players' utilities indeed depend on their types in these Bayesian games.

5.2.1 Second price, sealed bid auction

As before, the players will submit bids b_i. In this case we do not restrict the bids, and can allow them to take any value in $\mathbb{R}$. Formally, a pure strategy of a player in this game is a measurable function $b_i : \mathbb{R}^+ \to \mathbb{R}$, assigning a bid to each possible type or valuation.

As before, the player with the highest bid $b_{\max}$ wins and pays the second highest bin b_{nd}. Note that despite the fact that two valuations can be never be the same, it still may be the case that two players choose the same bid. For example, the strategy profile could be such that all players always bid 1. Accordingly, we assume that there is some tie-breaking mechanism (e.g., choose at random from all those with the highest bid), but it will not play a role in our analysis.

Proposition 5.7. *For any joint distribution $\mathbb{P}$, it is weakly dominant for each player to choose $b_i^*(v_i) = v_i$.*

The proof of this is identical to the one in the non-Bayesian case.

As an example, consider the case that the valuations are picked i.i.d. from some non-atomic distribution with cumulative distribution function F. Then $b_i^*(v_i) = v_i$ is the unique Bayes-Nash equilibrium.

Assume that there are two players. Player 1 wins the auction if she has the highest valuation. Conditioning on her valuations v_1, her probability of winning is therefore $F(v_1)$. If she wins then she expects to pay $\mathbb{E}_F[v_2 | v_2 < v_1]$.

5.2.2 First price, sealed bid auction

In this auction, as in the classical one, each player will submit a bid b_i, and the player with the highest bid $b_{\max}$ will win and pay $b_{\max}$. We assume here that the valuations are picked i.i.d. from some non-atomic distribution with cumulative distribution function F with derivative f. To simplify our calculations we will assume that there are only two players; the general case is almost identical.

Exercise 5.8. *Show that $b_i^*(v_i) = v_i$ is* not *an equilibrium of this auction.*

We will try to construct a Bayes-Nash equilibrium with the following properties:

- Symmetry: there is a function $b : \mathbb{R}^+ \to \mathbb{R}^+$ such that $b_i(v_i) = b(v_i)$ for all players.
- Monotony and differentiability: b is monotone increasing and differentiable.

Thus, to construct such an equilibrium we assume all players play b, and try to calculate b assuming it is some player's (say player 1's) best response.

Assume then that player 2 plays $b_2(v_2) = b(v_2)$. Fix v_1, and assume that player 1 bids b_1. Denote by G the cumulative distribution function of $b(v_2)$, and let g be its derivative. Note that we can write G and g in terms of F and f:

$$F(v) = G(b(v)) \tag{5.1}$$

and

$$f(v) = g(b(v)) \cdot b'(v). \tag{5.2}$$

The probability that b_1 is the highest bid is $G(b_1)$. It follows that player 1's expected utility (conditioned on v_1) is

$$u_1(v_1, b_1) = G(b_1) \cdot (v_1 - b_1).$$

Therefore, to maximize expected utility, b_1 must satisfy

$$0 = \frac{\mathrm{d}u_1(v_1, b_1)}{\mathrm{d}b_1} = g(b_1) \cdot (v_1 - b_1) - G(b_1),$$

or

$$G(b_1) = g(b_1) \cdot (v_1 - b_1).$$

Note that $b_1 = v_1$ is a solution only if $G(b_1) = 0$.

Since we are looking for a symmetric equilibrium, we can now plug in $b_1 = b(v)$ to arrive at the condition

$$G(b(v)) = g(b(v)) \cdot (v - b(v)).$$

Translating back to F and f using (5.1) and (5.2) yields

$$F(v) = \frac{f(v)}{b'(v)} \cdot (v - b(v)).$$

Rearranging, we can write

$$F(v)b'(v) + f(v) \cdot b(v) = f(v) \cdot v$$

or

$$\frac{\mathrm{d}}{\mathrm{d}v}[F(v)b(v)] = f(v) \cdot v.$$

Now, clearly $b(0) = 0$ is weakly dominant, and so we will assume that this is indeed the case. We can therefore solve the above expression to arrive at

$$b(v) = \frac{1}{F(v)} \int_0^v f(u) \cdot u \mathrm{d}u.$$

Note that this is equal to $\mathbb{E}_F[v_2 | v_2 < v_1]$, the expectation of v_2, conditioned on v_2 being less than v. It remains to be shown that this strategy profile is indeed a maximum (we only checked the first order condition).

A player's expected utility, conditioned on having valuation v, is simply $F(v)$ (the probability that she wins) times $v - b(v)$. Interestingly, in the second price auction the expected utility is identical: the probability of winning is still $F(v)$, and the expected utility is $v - b(v)$, since, conditioned on winning, the expected payment is the expected valuation of the other player.

5.3 Truthful mechanisms and the revelation principle

The revelation principle is important in mechanism design. The basic idea is due to Gibbard [16], with generalizations by others [19, 26, 10].

In this section we will call Bayesian games of incomplete information *mechanisms*. We will say that a mechanism is *truthful* if $A_i = T_i$ and $\tilde{s}(t_i) = t_i$ is an equilibrium. Note that sometimes this term is used to describe mechanisms in which the same $\tilde{s}$ is weakly dominant.

Theorem 5.9. *Let $G = (N, \{A_i\}_{i\in N}, (\Omega,\Sigma), \{\mu_i\}_{i\in N}, \{T_i\}_{i\in N}, \{t_i\}_{i\in N}, \{u_i\}_{i\in N})$ be a mechanism with an equilibrium s^*. Then there exists a truthful mechanism $G' = (N, \{A'_i\}_{i\in N}, (\Omega,\Sigma), \{\mu_i\}_{i\in N}, \{T_i\}_{i\in N}, \{t_i\}_{i\in}$ such that*

$$\mathbb{E}_{\mu_i,\tilde{s}}\left[u'_i \middle| t_i\right] = \mathbb{E}_{\mu_i,s^*}\left[u_i | t_i\right].$$

That is, for every game and equilibrium one can design a truthful game in which playing truthfully yields the same conditionally expected utilities as in the original game. The idea of the proof is simple: in the new mechanism, the players reveal their types, the mechanism calculates their equilibrium actions, and then implements the original mechanism on those actions.

Proof of Theorem 5.9. Let

$$u'_i(\tau_1,\ldots,\tau_n,\omega) = u_i(s^*_1(\tau_1),\ldots,s^*_n(\tau_n),\omega).$$

Then

$$\begin{aligned}
\mathbb{E}_{\mu_i,\tilde{s}}\left[u'_i \middle| t_i = \tau_i\right] &= \int_\Omega u'_i(\tilde{s}_1(t_1(\omega)),\ldots,\tilde{s}_n(t_n(\omega)),\omega)\mathrm{d}\mu_i(\omega|t_i=\tau_i)\\
&= \int_\Omega u'_i(t_1(\omega),\ldots,t_n(\omega),\omega)\mathrm{d}\mu_i(\omega|t_i)\\
&= \int_\Omega u_i(s^*_1(t_1(\omega)),\ldots,s^*_n(t_n(\omega)),\omega)\mathrm{d}\mu_i(\omega|t_i)\\
&= \mathbb{E}_{\mu_i,s^*}\left[u_i|t_i\right].
\end{aligned}$$

To see that this mechanism is truthful, note that for any player i with type $\tau_i \in T_i$ and action $\tau'_i \in A'_i = T_i$ it holds that the utility for playing τ'_i (instead of τ_i) is

$$\begin{aligned}
&\int_\Omega u'_i(\tilde{s}_1(t_1(\omega)),\ldots,\tau'_i,\ldots,\tilde{s}_n(t_n(\omega)),\omega)\mathrm{d}\mu_i(\omega|t_i=\tau_i)\\
&= \int_\Omega u'_i(t_1(\omega),\ldots,\tau'_i,\ldots,t_n(\omega),\omega)\mathrm{d}\mu_i(\omega|t_i=\tau_i)\\
&= \int_\Omega u_i(s^*_1(t_1(\omega)),\ldots,s^*_i(\tau'_i),\ldots,s^*_n(t_n(\omega)),\omega)\mathrm{d}\mu_i(\omega|t_i=\tau_i).
\end{aligned}$$

But since s^* is an equilibrium this is

$$\le \int_\Omega u_i(s_1^*(t_1(\omega)), \ldots, s_i^*(\tau_i), \ldots, s_n^*(t_n(\omega)), \omega) \mathrm{d}\mu_i(\omega | t_i = \tau_i)$$
$$= \int_\Omega u_i'(\tilde{s}_1(t_1(\omega)), \ldots, \tilde{s}_n(t_n(\omega)), \omega) \mathrm{d}\mu_i(\omega | t_i = \tau_i),$$

which is the utility for playing according to $\tilde{s}_i$. □

It follows that when designing auctions we can assume without loss of generality that the players reveal their types to the auctioneer.

6 Extensive form games with chance moves and imperfect information

6.1 Motivating example: train inspections

Consider a rail system in which there are no physical barriers to boarding a train without a ticket, but where occasionally inspectors check if passengers have tickets.

A ticket to the train costs C. The value of the ride to a passenger is $V > C$. The fine for getting caught on the train without a ticket is F. A passenger can have one of two different types: rich or poor. The rich passenger has enough money in savings to afford the fine. The poor passenger does not, and therefore would have to borrow money to pay the fine, thus paying in addition some amount $I_p > 0$ in interest. The interest for the rich is $I_r = 0$.

The type of the passenger is chosen at random at the beginning of the game, and is rich with probability p and poor with probability $1-p$. The passenger knows her type, but the train company does not. We imagine that p is small, so that most passengers are poor.

The passenger now has to decide whether or not to buy a ticket (B/NB). The company has to decide whether or not to check her for a ticket (C/NC). The cost of such an inspection is ε. We assume that $V < F$. We will see later that we will also need that $p \leq \varepsilon/F$.

The payoff to the passenger is the value of the ride (if she choose to ride), minus any money paid (in fare, fine or interest). The payoff to the company is any money collected, minus the cost of the inspection, if made.

When the passenger has value V for the ride and pays interest I on a loan to pay a fine, the normal form of the game is the following:

	C	NC
B	$V-C, C-\varepsilon$	$V-C, C$
NB	$V-F-I, F-\varepsilon$	$V, 0$

Now, the company does not know I, but believes that, as explained above, it equals I_r with probability p and I_p with probability $1-p$.

We first claim that this game has no pure Nash equilibria. Assume that the company chooses to always inspect. Then the best response of any passenger is to ride and buy. But then inspecting is not a best response for the company.

Assume then that the company does not inspect. Then all passengers best respond by riding and not buying. Now again the company has a profitable deviation, which is to inspect. Hence indeed this game has no pure equilibrium.

We will now build a mixed equilibrium. Assume that the company mixes and inspects with probability σ_c. Then the passenger's utility for riding and buying is $u_p((R,B),\sigma_c) = V-C$. Her utility for riding and not buying is

$$u_p((R,NB),\sigma_c) = \sigma_c(V-F-I) + (1-\sigma_c)V = V - \sigma_c(F+I),$$

and her utility for not riding is zero. Hence if we set $\sigma_c = C/(F+I)$ then the passenger wants to ride, but is indifferent between buying and not buying. Note that it cannot be that both the rich and the poor passenger are indifferent, since they pay different amounts of interest. We will set $\sigma_c = C/(F+I_p)$, so that the poor passenger is indifferent. Under this choice the utility for the rich passenger when riding and not buying is $V - CF/(F+I_p) > V - C$, and so this is her best response.

Assume that the rich passenger rides and does not buy, and that the poor passenger mixes between riding and buying (with probability σ_p) and riding and not buying (with probability $(1-\sigma_p)$). Then the expected utility for the company for inspecting is

$$\begin{aligned} u_c((R,B),\sigma_p,C) &= p(F-\varepsilon) + (1-p)\big[\sigma_p(C-\varepsilon) + (1-\sigma_p)(F-\varepsilon)\big] \\ &= pF - p\varepsilon + (1-p)(C-\varepsilon)\sigma_p + (1-p)(F-\varepsilon) - (1-p)(F-\varepsilon)\sigma_p \\ &= (C-F)(1-p)\sigma_p + F - \varepsilon \end{aligned}$$

The expected utility for not inspecting is

$$u_c((R,B),\sigma_p,NC) = C(1-p)\sigma_p$$

For the company to be indifferent we therefore set

$$\sigma_p = \frac{1-\varepsilon/F}{1-p}.$$

Thus if $p \leq \varepsilon/F$ this will not be larger than one, and we will have our equilibrium. In this equilibrium the rich do not buy a ticket, as they prefer to risk getting fined. The poor are indifferent between not riding and riding without buying. The expected utility for the poor passenger is zero. The expected utility for the rich passenger is $V - CF/(F+I_p) > 0$, and the expected utility for the company is $C(1-\varepsilon/F) \geq 0$.

6.2 Definition

In this section we introduce two new elements to extensive form games: *chance moves* and *imperfect information*. The idea behind chance moves is to model randomness that is introduced to the game by an outside force ("nature") that is not one of the players. Imperfect information models situations where players do not observe everything that happened in the past. We will restrict ourselves to games of *perfect recall*: players will not forget any observations that they made in the past. To simplify matters we will not allow simultaneous moves[9].

In the train inspections game the choice of whether the passenger is rich or poor is an example of a chance move. The fact that the train company decides to inspect without knowing the passenger's decision means that this is a game of imperfect information.

In this section, an extensive form game will be given by $G = (N, A, H, \mathscr{I}, P, \sigma_c, \{u_i\})$ where N, A and H are as in games of perfect information (Section 2.4), Z are again the terminal histories, and

[9] This is (almost) without loss of generality, since simultaneous moves are equivalent to sequential moves that are not observed by others until later.

- $\mathscr{I}$ is a partition of the non-terminal histories $H \setminus Z$ such that, for all $J \in \mathscr{I}$ and $h_1, h_2 \in J$, it holds that $A(h_1) = A(h_2)$. We therefore define can define $A : \mathscr{I} \to A$ by $A(J) = A(h)$ where $h \in J$ is arbitrary. We denote by $J(h)$ the partition element $J \in \mathscr{I}$ that contains $h \in H \setminus Z$.

- $P : \mathscr{I} \to N \cup \{c\}$ assigns to each non-terminal history either a player, or c, indicating a *chance move*. We sometimes think of the chance moves as belonging to a *chance player* c.

  When $P(J) = i$ we say that J is an *information set* of player i. The collection of the information sets of player i is denoted by $\mathscr{I}_i = P^{-1}(i)$ and is called i's *information partition*.

- Let $A_c = \prod_{J \in P^{-1}(c)} A(J)$ be the product of all action sets available to the chance player. σ_c is a product distribution on A_c. That is,

$$\sigma_c = \prod_{J \in P^{-1}(c)} \sigma_c(\cdot|J),$$

  where $\sigma_c(\cdot|J)$ is a probability distribution on $A(J)$, the set of actions available at information set J.

- For each player i, $u_i : Z \to \mathbb{R}$ is her utility for each terminal history.

We will assume that G is a game of *perfect recall*: For each player i and each $h = (a^1, a^2, \ldots, a^n)$ that is in some information set of i, let the *experience* $X(h)$ be the sequence of i's information sets visited by prefixes of h, and the actions i took there. That is, $X(h)$ is the sequence

$$\big((J^1, b^1), (J^2, b^2), \ldots, (J^k, b^k)\big)$$

where each J^m is an element of $\mathscr{I}_i$, each b^m is an element of $A(J^m)$, and $(b^1, \ldots, b^m)$ is the subsequence of h which includes the actions taken by i.

Perfect recall means that for each $J \in \mathscr{I}$ and $h_1, h_2 \in J$ it holds that $X(h_1) = X(h_2)$. That is, there is only one possible experience of getting to J, which we can denote by $X(J)$. In particular, in a game of perfect recall each information set is visited at most once along any play path.

6.3 Pure strategies, mixed strategies and behavioral strategies

A pure strategy of player i in G is a map s_i that assigns to each $J \in \mathscr{I}_i$ an action $a \in A(J)$. We can think of s_i as an element of $A_i := \prod_{J \in \mathscr{I}_j} A(J)$. A mixed strategy σ_i of a player in an extensive form game is a distribution over pure strategies. Given a $J \in \mathscr{I}_i$, we denote by $\sigma_i(\cdot|J)$ the distribution over $A(J)$ given by σ_i, conditioned on the experience $X(J)$. That is, $\sigma_i(a|J)$ is the probability, if we choose s_i according to σ_i, that $s_i(J) = a$ conditioned on $s_i(J^1) = b^1, \dots, s_i(J^k) = b^k$, where $((J^1, b^1), (J^2, b^2), \dots, (J^k, b^k))$ is the unique experience that terminates in J. Of course, it could be that this conditional probability is not well defined, in the case that $s_i(J^1) = b^1, \dots, s_i(J^k) = b^k$ occurs with zero probability.

Recall that $A_i = \prod_{J \in \mathscr{I}_i} A(J)$ is the product of all action sets available to player i. A *behavioral strategy* σ_i of player i is a product distribution on A_i:

$$\sigma_i = \prod_{J \in \mathscr{I}_i} \sigma_i(\cdot|J),$$

where $\sigma_i(\cdot|J)$ is a distribution on $A(J)$. Note that σ_c, the chance player's distribution, is a behavioral strategy. Note also that each element of $\prod_{J \in \mathscr{I}_i} A(J)$ can be identified with a function that assigns to each element $J \in \mathscr{I}_i$ an element of $A(J)$. Therefore, by our definition of behavioral strategies, every behavioral strategy is a mixed strategy.

Given a strategy profile σ of either pure, mixed or behavioral strategies (or even a mixture of these), we can define a distribution over the terminal histories Z by choosing a random pure strategy for each player (including the chance player), and following the game path to its terminal history z. A player's utility for σ is $u_i(\sigma) = \mathbb{E}[u_i(z)]$, her expected utility at this randomly picked terminal history.

Proposition 6.1. *Under our assumption of perfect recall, for every mixed (resp., behavioral) strategy σ_i there is a behavioral (resp., mixed) strategy σ'_i such that, for every mixed σ_{-i} it holds that $u_i(\sigma_{-i}, \sigma_i) = u_i(\sigma_{-i}, \sigma'_i)$.*

We prove this proposition for finite games. Note that our definition of behavioral strategies as a special case of mixed strategies is designed for games in which each information set is visited only once. In the general case, behavioral strategies are defined differently: they are simply a distribution on each information set, with the understanding that at each visit a new action is picked independently.

Proof of Proposition 6.1. By our definition of behavioral strategies, every behavioral strategy is a mixed strategy, and so if σ_i is a behavioral strategy we can simply take $\sigma'_i = \sigma_i$.

To see the other direction, let σ_i be a mixed strategy. For $J \in \mathscr{I}_i$ and $a \in A(J)$, let

$$\sigma'_i(a|J) = \sigma_i(a|J),$$

provided the conditioned event of reaching J has positive probability; otherwise let $\sigma'_i(\cdot|J)$ be arbitrary.

Fix σ_{-i} and let $h = (a^1, \dots, a^k)$ be a history. We denote by $\mathbb{P}\left[a^1, \dots, a^k\right]$ the probability that this history is played when using the strategy profile (σ_{-i}, σ_i). Assume by induction that this probability is the same whether we calculate it using σ_i or σ'_i, for all histories of length $< k$.

Note that

$$\mathbb{P}\left[a^1,\ldots,a^k\right] = \mathbb{P}\left[a^k \middle| a^1,\ldots,a^{k-1}\right] \cdot \mathbb{P}\left[a^1,\ldots,a^{k-1}\right].$$

Now, by our inductive assumption $\mathbb{P}\left[a^1,\ldots,a^{k-1}\right]$ takes the same value if we calculate it using σ'_i rather than σ_i. If $h = (a^1,\ldots,a^{k-1})$ is a history in an information set J that does not belong to $\mathscr{I}_i$ then clearly $\mathbb{P}\left[a^k \middle| a^1,\ldots,a^{k-1}\right]$ does not depend on whether we use σ_i or σ'_i, and hence $\mathbb{P}\left[a^1,\ldots,a^k\right]$ does not either.

Otherwise $J \in \mathscr{I}_i$. In this case, let s_i be a strategy that is picked according to σ_i. Then

$$\begin{aligned}
\mathbb{P}\left[a^k \middle| (a^1,\ldots,a^{k-1})\right] &= \mathbb{P}\left[s_i(J) = a^k \middle| a_1,\ldots,a^{k-1}\right] \\
&= \mathbb{P}\left[s_i(J) = a^k \middle| s_i(J^1) = b^1,\ldots,s_i(J^\ell) = b^\ell\right] \\
&= \sigma(a^k | J),
\end{aligned}$$

where $((J^1,b^1),\ldots,(J^\ell,b^\ell))$ is player i's experience at J, the partition element of h. This holds since the other players' choices are independent of i's and hence can be left out. Hence, by our definition of σ'_i, $\mathbb{P}\left[a^k \middle| a^1,\ldots,a^{k-1}\right]$ is the same under σ_i and σ'_i. Therefore the same applies to $\mathbb{P}\left[a^1,\ldots,a^k\right]$, and in particular to any terminal history. Thus the distribution on terminal histories is identical, and hence so are the expected utilities. $\square$

6.4 Belief systems and assessments

Let $G = (N, A, H, \mathscr{I}, P, \sigma_c, \{u_i\})$ be a finite extensive form game. A *belief system* $\{\mu(\cdot|J)\}_{J \in \mathscr{I}}$ is a collection of probability measures, with $\mu(\cdot|J)$ a probability measure over J.

Fix a mixed strategy profile σ, and for a history $h = (a^1, \ldots, a^k)$ denote, as above, by $\mathbb{P}_\sigma[h]$ the probability of the event that the path of play includes h. We likewise denote by $\mathbb{P}_\sigma[J]$ the probability that an information set J is visited. We say that $\mu(\cdot|J)$ is derived from σ if for any history $h \in J$ such that $\mathbb{P}_\sigma[J] > 0$ it holds that $\mu(h|J) = \mathbb{P}_\sigma[h|J]$. For J with $\mathbb{P}_\sigma[J] = 0$, $\mu(\cdot|J)$ can take any value.

An *assessment* is a pair (σ, μ) such that[10] μ is derived from σ.

Recall that we say that σ is *completely mixed* if for all i, $J \in \mathscr{I}_i$ and $a \in A(J)$ it holds that $\sigma_i(a|J) > 0$. That is, in every information set every allowed action has positive probability. In this case $\mathbb{P}_\sigma[J] > 0$ for every information set J, and so there is only one belief system μ that is derived from σ. Hence for completely mixed σ there is only one assessment (σ, μ).

6.5 Sequential rationality and sequential equilibria

In this section we introduce a concept that is a natural adaptation of subgame perfection to games with incomplete information.

Given a strategy profile σ and a belief system μ, we can naturally extend each $\mu(\cdot|J)$ to a distribution over all terminal histories that can be reached from J: given a terminal history $z = (a^1, \ldots, a^k, a^{k+1}, \ldots, a^\ell)$ such that $h = (a^1, \ldots, a^k) \in J$, denote by $J^m := J(a^m)$ the information set to which a^m belongs, and let

$$\mu_\sigma(z|J) = \mu(h|J) \cdot \prod_{m=k}^{\ell-1} \sigma_{P(J^m)}(a^{m+1}|J^m).$$

Note that $\mu_\sigma(z|J)$ is well defined since for any terminal history that passes through J there is, by perfect recall, a unique prefix h that ends in J.

Recall from the previous section that an assessment (σ, μ) induces, for each $J \in \mathscr{I}$, a distribution $\mu(\cdot|J)$ on the terminal histories reachable from J. We say that (σ^*, μ^*) is *sequentially rational* if for each player i and $J \in \mathscr{I}_i$ it holds that

$$\mathbb{E}_{\mu^*_{\sigma^*}(\cdot|J)}[u_i(z)] \geq \mathbb{E}_{\mu^*_\sigma(\cdot|J)}[u_i(z)]$$

for any $\sigma = (\sigma^*_{-i}, \sigma_i)$. Intuitively, sequential rationality corresponds to subgame perfection: there are no profitable deviations for any player at any information set.

Exercise 6.2. *Show that this notion of subgame perfection reduces to the usual notion in the case of games of perfect information.*

Recall that for completely mixed σ there is a unique assessment (σ, μ). However, when σ is not completely mixed, there can be excessive freedom in choosing μ, which can result in

[10] In the literature this requirement is usually not part of the definition.

"strange" belief updates following deviations. For example, consider the three player game in which players 1 and 2 choose among $\{a,b\}$, and player 3 only observes player 1's choice, so that her information sets are J_a and J_b, corresponding to player 1's choice. Player 3 then chooses among $\{a,b\}$. Let $\sigma_1(a) = \sigma_2(a) = \sigma_3(a) = 1$, and let $\mu(aa|J_a) = 1$, $\mu(bb|J_b) = 1$; the rest of the information sets are singletons and therefore have trivial beliefs. The fact that $\mu(bb|J_a) = 1$ means that if player 3 learns that 1 deviated then she assumes that 2 also deviated. See also the game in [33, Figure 4].

A common way to restrict μ is to require *consistency*: We say that an assessment (σ,μ) is consistent if there exists a sequence $\{(\sigma^m,\mu^m)\}_{m\in\mathbb{N}}$ that converges to (σ,μ) and such that each σ^m is completely mixed. We say that (σ,μ) is a *sequential equilibrium* [21] if it is consistent and sequentially rational.

Exercise 6.3. *Show that if G has perfect information then (σ,μ) is a sequential equilibrium iff it is subgame perfect.*

6.6 Trembling hand perfect equilibrium

In extensive form games one could define trembling hand equilibria as simply the trembling hand equilibria of the strategic form game. However, this leads to equilibria that are not subgame perfect (see, e.g., [29, Figure 250.1]).

We therefore define a different notion for extensive form games. To this end, given a behavioral strategy $\sigma_i = \prod_{J\in\mathscr{I}_i}\sigma_i(\cdot|J)$, we think of each decision made at each information set as being made by a different agent. This is sometimes called the *agent form* of the game. Here the set of agents is the set of information sets (except those that belong to the chance player) and the strategy of player $J \in \mathscr{I}_i$ is $\sigma_i(\cdot|J)$.

An *extensive form trembling hand perfect equilibrium* of G is a behavior strategy profile σ that is a trembling hand perfect equilibrium of the agent form of G (in strategic form).

6.7 Perfect Bayesian equilibrium

An assessment that is sequentially rational is said to be a *weak perfect Bayesian equilibrium*. As we mentioned above, this seems to be a problematic notion (see [33, Figure 4]).

The most widely used notion of perfect equilibrium is that of a *perfect Bayesian equilibrium*, due to Fudenberg and Tirole [14]. However, they only define it for a small class of extensive form games with incomplete information, and it is not clear what the (widely used!) definition is for general games. An excellent discussion and two possible notions are presented by Watson [33]. These notions involve two principles:

1. *Bayesian updating.* Whenever possible beliefs are updated using Bayes' rule. This is obviously possible on path, but also sometimes in updates following a deviation but not immediately after it.

2. *Conservation of independence.* A natural assumption is that if a player deviates then the others update their beliefs regarding her strategy / type, but not the other players',

unless necessary. More generally, if two events are independent before a deviation, and if the deviation is (somehow) independent of one, then beliefs regarding the other should not be updated [8].

6.8 Cheap talk

The contents of this section are taken from [5].

Consider a game played between a sender S and a receiver R. There is a state of the world $\omega \in \{H,T\}$, and both players have the common belief that it is chosen from the uniform distribution.

The sender observes ω, sends a message $m(\omega)$ to the receiver. The messages take values in some set M, which, as it turns out, we can assume without loss of generality to be $\{H,T\}$. The receiver then chooses an action $a \in A$.

In a *pure* strategy profile, the sender strategy consists of a choice of the function m. The receiver's strategy s_R is a function $s_R : M \to A$, which is a choice of action for each possible received message.

In equilibrium, the receiver knows the sender's strategy, and, upon receiving the message, updates his belief regarding the state of the world using Bayes' law. The action that he then chooses is optimal, given his belief. Likewise, the sender's expected utility is maximized by sending the message that she sends.

6.8.1 Example 1.

Here $A = \{L,R\}$. When $\omega = H$, the utilities are given by (the first number is always the sender's)

L	R
4,4	0,0

When $\omega = T$, they are

L	R
0,0	4,4

So both want action L when $\omega = H$ and R when $\omega = T$. In this case there are many equilibria:

- *Babbling.* The sender always sends $m(\omega) = H$. The receiver ignores this and chooses L. Likewise, the sender can choose at random (independently of ω) which message to send, which the receiver again ignores.

- *Fully revealing.* The sender sends $m(\omega) = \omega$, and the receiver choose L if the message was H, and R otherwise. By the cooperative nature of this game, neither has incentive to deviate.

6.8.2 Example 2.

Here again $A = \{L,R\}$. When $\omega = H$, the utilities are given by

L	R
0,0	1,4

When $\omega = T$, they are

L	R
0,0	1,−6

A possible interpretation of this game is the following: the sender is a job applicant who knows whether or not she is qualified. The receiver wants to hire only qualified applicants.

Here the sender wants the receiver to choose R, regardless of the state. Without any information, L is the better choice for the receiver. In this case there are only babbling equilibria. This is because if there is any message that the sender can send that will make the receiver choose R, then in equilibrium she must send it. Hence the receiver will ignore this message.

6.8.3 Example 3.

Here $A = \{LL, L, C, R, RR\}$. When $\omega = H$, the utilities are given by

LL	L	C	R	RR
−1,10	1,8	0,5	1,0	−1,−8

When $\omega = T$, they are

LL	L	C	R	RR
−1,−8	1,0	0,5	1,8	−1,10

So the sender wants actions L or R to be taken, and the receiver wants to take LL when $\omega = H$ and RR when $\omega = T$.

A possible interpretation is that the receiver is a voter that can vote for the extreme left or right, for moderate left or right, or to abstain. The "correct" vote depends on the unknown state of the world. The sender is interested in political stability and wants the voter to vote for any of the moderate choices.

Again, we will have a babbling equilibrium. More interestingly, there is a *partially revealing equilibrium*. In this equilibrium the sender chooses $m(\omega) = \omega$ with probability 3/4, and sends the opposite signal with probability 1/4. The receiver chooses L if she got the signal H, and T otherwise. A simple calculation shows that this choice maximizes her expected utility. This is optimal for the sender, since she gets the highest utility possible.

6.9 The Walmart game

In this section (based on [29, Example 239.1]) we describe a game of perfect information, find its unique subgame perfect equilibrium, and discuss why this seems an unsatisfactory description of reality. We then analyze a Bayesian version of the same game, which seems to have a more attractive solution.

6.9.1 Perfect information, one round

A mobster approaches a local grocer and asks for payment in exchange for "protection services". The grocer has two possible courses of action: comply (C) or refuse (R). In the latter case the mobster has two possible responses: punish (P) or forgive (F). The payoff matrix in strategic form is the following, with the mobster as the row player. Here $a > 1$ and $0 < b < 1$.

	C	R
P	$a,0$	$-1,b-1$
F	$a,0$	$0,b$

Alternatively, this can be a model of a game played by Walmart and the same grocer: Walmart opens a store in town. The grocer has two possible courses of action: leave town or stay. In the latter case Walmart has two possible responses: undercut or compete; undercutting refers to the practice of charging below cost in order to bankrupt the competition.

A Nash equilibrium of this game is for the grocer to comply, under threat of punishment by the mobster. However, this is not a credible threat, and in the unique subgame perfect equilibrium of this game the grocer refuses and the mobster forgives.

6.9.2 Perfect information, many rounds

Consider now the case that there are n grocers that the mobster engages, one after the other. Thus this is an $n+1$-player game, with the game between the mobster and each grocer being the one described above. Note that for the last grocer the unique subgame perfect equilibrium is again (R,F), and therefore, by backward induction, in the unique subgame perfect equilibrium of this game each grocer refuses and the mobster forgives, after any history.

Consider the case that the mobster deviates and punishes the first ten grocers who refuse. Under this (unique) subgame perfect equilibrium, the eleventh grocer will not learn the lesson and will again refuse.

6.9.3 Imperfect information, one round

Consider now the case that there are two types of mobsters: the sane mobster (s) and the crazy mobster (c). The former's utility is as described above. The latter derives sadistic pleasure from punishing, and therefore, when playing with a crazy mobster, the utility matrix is the following.

	C	R
P	$a,0$	$0,b-1$
F	$a,0$	$-1,b$

Consider now a Bayesian game in which the mobster is crazy with probability ε and sane with probability $1-\varepsilon$. The mobster observes his type but the grocer does not. They then play the same game.

The mobster does have complete information here, and so the analysis of the subgame perfect equilibrium is still simple. If the grocer refuses, clearly the sane mobster forgives and the crazy mobster punishes. Hence the grocer's expected utility for refusing is $(1-\varepsilon)b + \varepsilon(b-1) = b-\varepsilon$ and his utility for complying is 0. We assume that $\varepsilon < b$, and so we again have that the grocer refuses.

6.9.4 Imperfect information, many rounds

We now expand this Bayesian game to a game with n rounds, where, as before, in each round the mobster engages a new grocer. We number the rounds $\{0,1,\ldots,n-1\}$. We associate with each round k the number $b_k = b^{n-k}$.

The mobster's type is picked once in the beginning of the game, and is not observed by any of the grocers. We think of ε as small but fixed, and of n of being very large, so that $b_0 = b^n < \varepsilon$.

We describe an assessment that is sequentially rational, and is in fact a perfect Bayesian equilibrium, according to the definition in Fudenberg and Tirole [14]. In this equilibrium the mobster will (obviously) punish if he is crazy. However, even if he is sane, he will punish (except towards the end), to maintain the grocers' beliefs that he might be crazy. The result is that the grocers all comply, except for the last few.

Note that the mobster observes his own type and all the grocers' actions, and so there is no need to define beliefs for him. The grocers' beliefs are probability distributions over the mobster's type, and so we denote the belief of grocer $P(h)$ after a history h by $\mu(h)$, which we will take to denote the belief that the mobster is crazy (rather than having it denote the entire distribution).

The game will have two phases. In the first phase, the mobster will always punish any grocer that does not comply. This will last k^* rounds, where

$$k^* = \max\{k \in \mathbb{N} : \varepsilon > b_k\}.$$

In the second phase the sane mobster will use a mixed strategy to decide whether to punish or forgive, while the crazy mobster will always punish. Finally, in the last round the sane mobster always forgives (while the crazy still punishes).

The beliefs are defined as follows:

- Initially, $\mu(\emptyset) = \varepsilon$; this must be so if this is indeed to be an assessment (recall that in assessments we require beliefs to be derived from the strategy profile, and in particular from the chance moves).

- When a grocer complies the grocers do not observe any action of the mobster, and thus maintain the same belief. That is, $\mu(hC) = \mu(h)$ for any history h.

- If the mobster ever forgives a refusal then the grocers henceforth believe that he is sane. That is, $\mu(h) = 0$ for any history h that contains F.

- Given a history h in which the mobster has never forgiven, and which ends in round k in which the mobster punished, the belief $\mu(hP) = \max\{\mu(h), b_k\}$.

The strategies σ are defined as follows, where h is a history at round k:

- The grocer refuses if $\mu(h) < b_k$. He complies if $\mu(h) > b_k$. If $\mu = b_k$ then he complies with probability $1/a$ and refuses with probability $1 - 1/a$.

- The mobster has to move only if the grocer refused. If the mobster is crazy then he always punishes. If he is sane then on the last round he always forgives. On other rounds he punishes if $\mu(h) > b_k$. If $\mu(h) \le b_k$ then he punishes with probability p_k and forgives with probability $1 - p_k$, where

$$p_k = \frac{(1 - b_{k+1})\mu(h)}{b_{k+1}(1 - \mu(h))}.$$

We now analyze the game play along the equilibrium path. We assume that $\varepsilon > b^n$, and thus $k^* \ge 1$. Since $\mu(\emptyset) = \varepsilon$, in round 0 we have that $\mu > b_0$, and so the grocer complies. This leaves μ unchanged, and so, as long as the round k is at most k^*, we have that $\mu > b_k$ and the grocers keep complying. This is rational since the mobster (whether sane or not) punishes when $\mu > b_k$. If any of the grocers were to deviate and refuse in this phase, the mobster would punish (again, whether sane or not) and so the grocers would learn nothing and continue complying.

In the second phase, if $\mu(h) = b_k$ then a sane mobster will punish with probability p_k. Hence the probability that a mobster will punish is

$$\begin{aligned}
b_k + (1 - b_k)p_k &= b_k + (1 - b_k)\frac{(1 - b_{k+1})\mu(h)}{b_{k+1}(1 - \mu(h))} \\
&= b_k + (1 - b_k)\frac{(1 - b_{k+1})b_k}{b_{k+1}(1 - b_k)} \\
&= b_k + \frac{(1 - b_{k+1})b_k}{b_{k+1}} \\
&= b_k + (1 - b_{k+1})b \\
&= b.
\end{aligned}$$

It follows that a grocer's utility for complying is $(b-1)b + b(1-b) = 0$, making him indifferent between complying and refusing. It is therefore rational for him to play any mixed strategy. If $\mu(h) < b_k$ then the mobster punishes with probability less than b. In that case the only rational choice of the grocer is to refuse.

In the second phase, we claim that μ at round k is always at most b_k. We prove this by induction; it is true by definition at the first round of the second phase, round k^*.

If the grocer refuses and the mobster punishes at round $k > k^*$ (with probability p_k), and assuming the he has never forgiven before, Bayes' law yields that

$$\begin{aligned}\frac{\mu(hP)}{1-\mu(hP)} &= \frac{\mu(h)}{(1-\mu(h))p_k} \\ &= \frac{\mu(h)}{(1-\mu(h))}\frac{b_{k+1}(1-\mu(h))}{(1-b_{k+1})\mu(h)} \\ &= \frac{b_{k+1}}{1-b_{k+1}}.\end{aligned}$$

Hence $\mu(hP) = b_{k+1}$ at round $k+1$. If the grocer refuses and the mobster forgives then the grocers learn that the mobster is sane and $\mu(hF) = 0$. If the grocer complies then μ remains unchanged since no action is observed, and $\mu(hC) = \mu(h) < b_{k+1}$. Thus in any case it indeed holds that μ is at most b_k. Also, it follows that the grocers' beliefs are derived from σ.

Finally, we note (without proof) that the probabilities that the grocers choose in their mixed strategies make the mobster indifferent, and thus this assessment is indeed sequentially rational.

7 Repeated games

7.1 Definition

Let $G_0 = (N, \{A_i\}, \{u_i\})$ be a strategic form game, and, as usual, let $A = \prod_i A_i$. We will only consider games in which A is finite.

A repeated game G, with base game G_0, and lasting T periods (where T can be either finite or infinity) is a game in which, in each period, the players simultaneously choose an action from G_0, and then observe the other's actions. Their utility is some function of the *stage utilities*: the utilities they get in the base game in each period. Formally, the game is given by $G = (G_0, T, \{v_i\}_{i\in N})$, defined as follows.

- The *action* of player i in period t is denoted by a_i^t. The *action profile* in period t is $a^t = (a_i^t)_{i\in N}$.
- The history of the game at period t is the tuple $(a^1, \ldots, a^t)$ of action profiles played until and including period t. The history of the entire game is $a = (a^1, a^2, \ldots)$.
- The *stage utility* at period t is $u_i(a^t)$.
- A *strategy* s_i for player i is a map that assigns to each history $(a^1, \ldots, a^t)$, with $t < T$, the action a_i^{t+1}.
- For each player i, the *utility* or *payoff* in the game is $v_i = v_i(u_i(a^1), u_i(a^2), \ldots)$. As usual, given a strategy profile s, we denote by $v_i(s)$ the utility for i when the players play s. We will also write $v_i(a)$ to denote the utility player i gets when the history of the game is $a = (a^1, a^2, \ldots)$.

We will consider different choices for v_i below. An equilibrium is, as usual, a strategy profile s^* such that for each i and each s_i it holds that $v_i(s^*) \geq v_i(s_{-i}^*, s_i)$.

We call $(v_1(a), \ldots, v_n(a))$ the *payoff profile* associated with a. Given a strategy profile s of G, we likewise define the payoff profile $(v_1(s), v_2(s), \ldots, v_n(s))$ associated with s to be the payoff profile associated with the path of play generated by s.

A *subgame* of a repeated game G is simply the same game, but started at some later time period $t+1$, after a given history $(a^1, \ldots, a^{t+1})$.

7.2 Payoffs

7.2.1 Finitely repeated games

When T is finite then a natural choice is to let the players' utilities in the repeated game be the sum of their base game utilities in each of the periods:

$$v_i(a^1, \ldots, a^T) = \sum_{t=1}^{T} u_i(a^t).$$

When $T = \infty$ we will consider two types of preference relations: *discounting* and *limit of means*.

7.2.2 Infinitely repeated games: discounting

In *discounting* utilities, we fix some $\delta \in (0,1)$ and let $v_i : Z \to \mathbb{R}$ be given by the discounted sum

$$v_i(a^1, a^2, \ldots) = (1-\delta)\sum_{t=1}^{\infty} \delta^{t-1} u_i(a^t).$$

Note that we chose the normalization $(1-\delta)$ to make v_i a weighted average of the stage utilities. Discounting has the advantage that it is *stationary*: every subgame of G is isomorphic to G, in the sense that given a subgame starting after history $(a^1, \ldots, a^t)$, we can write the utility in the original game as

$$v_i(a^1, \ldots, a^t, a^{t+1}, \ldots) = (1-\delta)\sum_{\tau=1}^{t} \delta^{\tau-1} u_i(a^\tau) + \delta^t v_i(a^{t+1}, a^{t+1}, \ldots).$$

Thus, from the point of view of a player in the beginning of period $t+1$, she is facing the exact same game, except that her utilities will be multiplied by δ^t, and have the constant $(1-\delta)\sum_{\tau=1}^{t} \delta^{\tau-1} u_i(a^\tau)$ (that does not depend on her actions from this point on) added to them.

7.2.3 Infinitely repeated games: limit of means

In the usual definition of "limit of means", a preference relation is defined (rather than utilities), where player i weakly prefers the history a over the history b if

$$\liminf_{\tau} \frac{1}{\tau}\sum_{t=1}^{\tau} u_i(a^t) - \frac{1}{\tau}\sum_{t=1}^{\tau} u_i(b^t) \geq 0. \tag{7.1}$$

Note that, under this definition, a pair $(a^1, a^2, \ldots)$ and $(b^1, b^2, \ldots)$ that differ only in finitely many periods are equivalent. This definition captures preferences of agents who are only interested in the very long term. Note that the choice of the limit inferior here is somewhat arbitrary. We could have chosen the limit superior, for example.

We will take a slightly different approach to limits of means, and use utilities v_i rather than just preferences. It is tempting to define $v_i(a)$ to equal $\liminf_\tau \frac{1}{\tau}\sum_{t=1}^{\tau} u_i(a^t)$. However, this will not give us the same preference relation, since the limit inferior of a difference is not always the same as the difference of the limit inferiors. This can happen when these sequences do not converge.

Nevertheless, it turns out that there does exist a function $v_i(a^1, a^2, \ldots)$ with the property that v_i is always the limit of some subsequence of $\frac{1}{\tau}\sum_{t=1}^{\tau} u_i(a^t)$. In particular, when this limit exists (which will be the case that interests us) then

$$v_i(a^1, a^2, \ldots) = \lim_{\tau \to \infty} \frac{1}{\tau}\sum_{t=1}^{\tau} u_i(a^t).$$

As it turns out, we can furthermore choose these v_i's to be *stationary*. In this case, this means that

$$v_i(a^1, a^2, \dots) = v_i(a^{t+2}, a^{t+3}, \dots).$$

Thus, the utility in every subgame is exactly the same.

7.3 Folk theorems

7.3.1 Examples

What payoffs profiles are achievable in Nash equilibria of infinite repeated games? It turns out that the answer is: more or less all of them. To get some intuition as to how this is done, consider the following example. Let G_0 be the following prisoners' dilemma game:

	D	C
D	1,1	10,0
C	0,10	4,4

Consider the following symmetric strategy profile, called "grim trigger": start with C, and keep on playing C until the other person plays D. Then play D henceforth. It is easy to see that this is an equilibrium under both limiting means and discounting, for δ close enough to one.

A similar example is the following version of the public goods game. In G_0, each of 10 players has to choose a strategy $s_i \in \{0,1\}$, which we think of as the amount of effort they invest in a public goods project. The efforts are summed, doubled, and redistributed evenly, so that

$$u_i = -s_i + \frac{1}{10}\sum_{j=1}^{10} 2 \cdot s_j.$$

As in the prisoners' dilemma, $s_i = 0$ is dominant, since $u_i = -\frac{4}{5}s_i + \frac{1}{10}\sum_{j \neq i} 2 \cdot s_j$, but there is a grim trigger equilibrium in which, on path, all players play 1 and thus get utility 1.

7.3.2 Enforceable and feasible payoffs

We fix an infinitely repeated game G with base game G_0. Define the *minmax payoff* of player i in the base game G_0 as the lowest payoff that the rest of the players can force on i:

$$u_i^{mm} = \min_{a_{-i}} \max_{a_i} u_i(a_{-i}, a_i).$$

Equivalently, this is the payoff that player one can guarantee for herself, regardless of the other player's actions.

We say that a payoff profile $w \in \mathbb{R}^n$ is *enforceable* if $w_i \geq u_i^{mm}$ for all $i \in N$. It is *strictly enforceable* if $w_i > u_i^{mm}$ for all $i \in N$.

Claim 7.1. *Let s^* be a Nash equilibrium of G, under either discounting or limiting means. Then the payoff profile associated with s is enforceable.*

Proof. Since player i can guarantee a stage utility of u_i^{mm}, she can always choose a strategy such that her stage utilities will each be at least u_i^{mm}. Hence under both discounting and limiting means her payoff will be at least u_i^{mm} in G. □

We say that a payoff profile w is *feasible* if it is a convex combination of utilities achievable in G_0. That is, if for all $i \in N$

$$w_i = \sum_{a \in A} \alpha_a \cdot u_i(a)$$

for some $\{\alpha_a\}$ that sum to one. Clearly, every payoff profile in G is feasible.

Exercise 7.2. *Consider the game*

	D	C	F
D	$0,0$	$1,0$	$0,1$
C	$0,1$	$2,2$	$-2,3$
F	$1,0$	$2,-3$	$-2,-2$

Draw a diagram of the feasible and enforceable profiles.

7.4 Nash folk theorems

Aumann and Shapley [6] proved the Nash folk theorem for limit of means. We state it for limiting means.

Theorem 7.3. *For every feasible, enforceable payoff profile w there exists a Nash equilibrium of G with limiting means utilities whose associated payoff profile is w.*

The construction of these equilibria involves *punishing*: players all play some equilibrium, and if anyone deviates the rest punish them.

Proof of Theorem 7.3. Let $w_i = \sum_{a \in A} \alpha_a \cdot u_i(a)$. Let $(a^1, a^2, \ldots)$ be a sequence in A such that, for each $a \in A$, the fraction of periods in which a^t is equal to a tends to α_a, i.e.,

$$\lim_{t \to \infty} \frac{|\{\tau \le t : a^\tau = a\}|}{t} = \alpha_a.$$

Such a sequence exists since (for example) if we choose each a^t independently at random to equal a with probability α_a then with probability one the sequence has this property. If the coefficients α_a are rational then one can simply take a periodic sequence with period equal to the lowest common denominator.

Let s^* be the following strategy profile. For each player i let s^*_i be the strategy in which she chooses a^t_i, unless in some previous period τ some player j did not choose a^τ_j, in which case she chooses a strategy b_i, where

$$b_{-j} \in \operatorname*{argmin}_{a_{-j}} \max_{a_j} u_j(a_{-j}, a_j).$$

Hence the stage utilities of a player i who deviates will be, from the point of deviation on, at most u^{mm}_i. Therefore her utility will be at most u^{mm}_i, since utilities do not depend on any finite set of stage utilities. Since w is enforceable, it follows that $w_i \ge u^{mm}_i$, and so no deviation is profitable, and s^* is an equilibrium. $\square$

A similar proof technique can be used to show the following theorem, which is due to Fudenberg and Maskin [13], with an earlier, weaker version by Friedman [12].

Theorem 7.4. *For every feasible,* strictly *enforceable payoff profile w and $\varepsilon > 0$ there is a $\delta_0 > 0$ such that for all $\delta > \delta_0$ there exists a Nash equilibrium of G with δ-discounting utilities whose associated payoff profile w' satisfies $|w'_i - w_i| < \varepsilon$ for all $i \in N$.*

The discount factor needs to be large enough to make eternal punishment pose more of a loss than can be gained by a single deviation. It also needs to be large enough to allow for the discounted averages to approximate a given convex combination of the base game utilities.

We provide here a proof of a weaker statement.

Theorem 7.5. *Let a be an action profile whose associated payoff profile w is strictly enforceable*[11]*. Then there is a δ_0 such that, for all $\delta > \delta_0$, there exists a Nash equilibrium of G with δ-discounting utilities whose associated payoff profile is also w.*

Proof. Let s^* be the strategy profile defined in the proof of Theorem 7.3, with $\alpha_a = 1$. We will show that for δ close enough to 1 it is also a Nash equilibrium, and that its associated payoff profile is w. In fact, the latter is immediate, since the stage utilities for player i are all equal to w_i on the equilibrium path.

Let M be the largest possible stage utility achievable by any player in one period by any deviation. Then the utility of player i who first deviates in period $\tau + 1$ to some strategy s_i satisfies

$$v_i(s^*_{-i}, s_i) \le (1-\delta)\left(\sum_{t=1}^{\tau} \delta^{t-1} w_i + \delta^\tau M + \sum_{t=\tau+2}^{\infty} \delta^{t-1} u_i^{mm}\right).$$

Hence

$$\begin{aligned} v_i(s^*) - v_i(s^*_{-i}, s_i) &\ge (1-\delta)\left(\delta^\tau (w_i - M) + \sum_{t=\tau+2}^{\infty} \delta^{t-1}(w_i - u_i^{mm})\right) \\ &= (1-\delta)\delta^\tau \left(w_i - M + (w_i - u_i^{mm}) \sum_{t=2}^{\infty} \delta^{t-1}\right) \\ &= (1-\delta)\delta^\tau \left(w_i - M + (w_i - u_i^{mm}) \frac{\delta}{1-\delta}\right). \end{aligned}$$

Now, $\delta/(1-\delta)$ tends to infinity as δ tends to one. Hence the expression in the parentheses is non-negative for δ large enough, and the deviation is not profitable. ☐

[11] Since w is a payoff profile of some action profile then it is immediately feasible.

7.5 Perfect folk theorems

Consider the following base game (taken from Osborne and Rubinstein [29]):

	D	C
D	0,1	0,1
C	1,5	2,3

Here, an equilibrium built in Theorems 7.3 and 7.4 that achieves payoff profile $(2,3)$ has the players playing (C,C) on the equilibrium path, and punishing by playing D forever after a deviation. Note, however, that for the row player, action D is strictly dominated by C. Hence this equilibrium is not a subgame perfect equilibrium: regardless of what the column player does, the row player can increase her subgame utility by at least 1 by always playing C rather than D. It is therefore interesting to ask if there are subgame perfect equilibria that can achieve the same set of payoff profiles.

7.5.1 Perfect folk theorem for limiting means

The following theorem is due to Aumann and Shapley [6], as well as Rubinstein [30].

Theorem 7.6. *For every feasible, strictly enforceable payoff profile w there exists a subgame perfect Nash equilibrium of G with limiting means utilities whose associated payoff profile is w.*

The idea behind these equilibria is still of punishing, but just for some time rather than for all infinity. Since lower stage utilities are only attained in finitely many periods there is no loss in the limit.

Proof of Theorem 7.6. As in the proof of Theorem 7.3, let $w_i = \sum_{a\in A} \alpha_a \cdot u_i(a)$, and let $(a^1, a^2, \ldots)$ be a sequence in A such that $m(\mathbb{1}_{\{a^1=a\}}, \mathbb{1}_{\{a^2=a\}}, \ldots) = \alpha_a$. Likewise, for each player i let s_i^* be the strategy in which she chooses a_i^t, unless in some previous period τ some player j deviated and did not choose a_j^τ. In the latter case, we find for each such τ and j a τ' large enough so that, if all players but j play

$$b_{-j} \in \underset{a_{-i}}{\operatorname{argmin}} \max_{a_i} u_i(a_{-1}, a_i),$$

in time periods $(\tau+1, \ldots, \tau')$ then the average of player j's payoffs in periods $(\tau, \tau+1, \ldots, \tau')$ is lower than w_j. Such a τ' exists since the payoffs in periods $(\tau+1, \ldots, \tau')$ will all be at most u_j^{mm}, and since $w_j > u_j^{mm}$. We let all players but j play b_{-j} in time periods $\tau+1, \ldots, \tau'$. We do not consider these punishments as punishable themselves, and after period τ' all players return to playing a_i^t (until the next deviation).

To see that s^* is a subgame perfect equilibrium, we consider two cases. First, consider a subgame in which no one is currently being punished. In such a subgame anyone who deviates will be punished and lose more than they gain for each deviation. Hence a deviant j's long run average utility will tend to at most w_j, and there is no incentive to deviate.

Second, consider a subgame in which someone is currently being punished. In such a subgame the punishers have no incentive to deviate, since the punishment lasts only finitely many periods, and thus does not affect their utilities; deviating from punishing will not have any consequences (i.e., will not be punished) but will also not increase utilities, and therefore there is again no incentive to deviate.

□

7.5.2 Perfect folk theorems for discounting

We next turn to proving perfect folk theorems for discounted utilities. An early, simple result is due to Friedman [12].

Theorem 7.7. *Let G_0 have a pure Nash equilibrium s^* with payoff profile z. Let w be a payoff profile of some strategy profile $a \in A$ of G_0 such that $w_i > z_i$ for all $i \in \mathbb{N}$. Then there is a $\delta_0 > 0$ such that for all $\delta > \delta_0$ there exists a subgame perfect equilibrium of G under discounting, with payoff profile w.*

The idea behind this result is simple: the players all play a unless someone deviates. Once anyone has deviated, they all switch to playing s^*. Given that δ is close enough to 1, the deviant's utility from playing s^* henceforth will trump any gains from the deviation. Since s^* is an equilibrium, there is no reason for the punishers to deviate from the punishment.

A harder result is due to Fudenberg and Maskin [13] who, for the two player case, extend the Nash folk theorem 7.4 to a perfect Nash folk theorem.

Theorem 7.8. *Assume that $|N| = 2$. For every feasible,* strictly *enforceable payoff profile w and $\varepsilon > 0$ there is a $\delta_0 > 0$ such that for all $\delta > \delta_0$ there exists a perfect Nash equilibrium of G with δ-discounting utilities whose associated payoff profile w' satisfies $|w'_i - w_i| < \varepsilon$ for all $i \in N$.*

Before proving this theorem will we state the following useful lemma. The proof is straightforward.

Lemma 7.9. *Let G be a repeated game, with δ-discounted utilities. The utilities for playing $(b, a^1, a^2, \ldots)$ are given by*

$$v_i(b, a^1, a^2, \ldots) = (1-\delta)u_i(b) + \delta v_i(a^1, a^2, \ldots).$$

More generally,

$$v_i(b^1, \ldots, b^k, a^1, a^2, \ldots) = (1-\delta)\sum_{t=1}^{k} \delta^{t-1} u_i(b^t) + \delta^k v_i(a^1, a^2, \ldots). \tag{7.2}$$

A useful interpretation of this lemma is the following: the utility for playing $(b, a^1, a^2, \ldots)$ is $1-\delta$ times the stage utility for b, plus δ times the utility of $(a^1, a^2, \ldots)$ in the subgame that starts in the second period.

We will also use the following lemma, which is a one deviation principle for repeated game with discounting.

Lemma 7.10. *Let G be a repeated game with δ-discounting. Let s^* be a strategy profile that is* not *a subgame perfect equilibrium. Then there is a subgame G' of G and a player i who has a profitable deviation in G' that differs from s_i^* only in the first period of G'.*

Proof. Let s_i be a profitable deviation from s_i^*. Assume without loss of generality that stage utilities take values in $[0,1]$, and let $v_i(s_{-i}^*, s_i) = v_i(s^*) + \varepsilon$. Let $\bar{s}_i$ be the strategy for player i which is equal to s_i up to some time period $\ell > \log(\varepsilon/2)/\log(\delta)$, and thereafter is equal to s^*. Then $\delta^\ell < \varepsilon/2$, and so, by (7.2),

$$|v_i(s_{-i}^*, s_i) - v_i(s_{-i}^*, \bar{s}_i)| \le \delta^\ell.$$

Hence $v_i(s_{-i}^*, \bar{s}_i) > v_i(s^*) + \varepsilon/2$, and thus $\bar{s}_i$ is a profitable deviation. Note that $\bar{s}_i$ differs from s_i^* in only finitely many histories.

Assume now that s_i, among i's profitable deviations, has a minimal number of histories in which it can differ from s_i^*. By applying an argument identical to the one used to prove the one deviation principle for finite extensive form games, there also exists a profitable deviation that differs from s_i^* at only one history, in which it matches s_i. □

We prove Theorem 7.8 for the particular case that $(w_1, w_2) = (u_1(a), u_2(a))$ for some feasible and strictly enforceable $a \in A$; in this case we can take $\varepsilon = 0$. The proof of the general case uses the same idea, but requires the usual technique of choosing a sequence of changing action profiles.

We assume that $u_i^{mm} = 0$. This is without loss of generality, since otherwise we can define a game G_0' in which the utilities are $u_i' = u_i - u_i^{mm}$. The analysis of the G_0'-repeated game will be identical, up to an additive constant for each player's utility.

Fix

$$b_1 \in \operatorname*{argmin}_{a_1} \max_{a_2} u_2(a_1, a_2)$$

and

$$b_2 \in \operatorname*{argmin}_{a_2} \max_{a_1} u_1(a_1, a_2)$$

Note that $u_1(b_1, b_2) \le 0$ and likewise $u_2(b_1, b_2) \le 0$, since we assume that $u_1^{mm} = u_2^{mm} = 0$.

As an example, consider the following base game:

	D	C	F
D	$0,0$	$1,0$	$0,1$
C	$0,1$	$2,2$	$-2,3$
F	$1,0$	$2,-3$	$-2,-2$

With $a = (C, C)$. It is easy to see that $u_1^{mm} = u_2^{mm} = 0$ and that necessarily $b = (F, F)$.

Consider the following strategy profile s^* for the repeated game. Recall that $(w_1, w_2) = (u_1(a), u_2(a))$ for some feasible and strictly enforceable $a \in A$. In s^*, the game has two "modes": *on-path mode* and *punishment mode*.

- In on-path mode each player i plays a_i.
- In punishment mode each player i plays b_i.

Both players start in the on-path mode. If any player deviates, the game enters punishment mode for some fixed number of ℓ rounds. This also applies when the game is already in punishment mode: if a player deviates when in punishment mode (i.e., stops punishing and does not play b_i), the game re-enters punishment mode for ℓ rounds.

Player i's utility on path is given by

$$v_i(s^*) = (1-\delta)\sum_{t=1}^{\infty}\delta^{t-1}u_i(a) = w_i.$$

Denote by $M = \max_{a\in A, i\in N} u_i(a)$ the maximal utility achievable by any player in G_0. Denote the punishment utility penalty by $p_i = -u_i(b_1,b_2) \geq 0$.

In the example above, $w_1 = w_2 = 2$ and $u_1(b_1,b_2) = u_2(b_1,b_2) = -2$.

To show that s^* is a subgame perfect equilibrium, we consider the $\ell+1$ types of possible subgames: the ℓ possible subgames that start in punishment mode (which differ by the number of punishment periods left), and those that start in on-path mode.

Let the subgame G^k start in punishment mode, with k punishment periods left. Then on the (subgame) equilibrium path, player i's utility is given by

$$\begin{aligned}v_i^k &= (1-\delta)\sum_{t=1}^{k}\delta^{t-1}\cdot u_i(b_1,b_2) + (1-\delta)\sum_{t=k+1}^{\infty}\delta^{t-1}\cdot w_i \\ &= \delta^k w_i - (1-\delta^k)p_i.\end{aligned}$$

Note that this is strictly decreasing in k, since $-p_i < w_i$. In particular

$$v_i^\ell = \delta^\ell w_i - (1-\delta^\ell)p_i \tag{7.3}$$

is less than v_i^k for all $k < \ell$.

Recall that by Lemma 7.10 we need only consider deviations at the first period. Since the other player is punishing player i, the utility for i for deviating at G^k is at most

$$(1-\delta)u_i^{mm} + \delta v_i^\ell = \delta v_i^\ell$$

since, applying This is independent of k, and so, since $v_i^\ell < v_i^k$ for all $k < \ell$, if it is profitable to deviate at any G^k then it is profitable to deviate at G^ℓ. In order for it to not be profitable to deviate at G^ℓ it must be that

$$\delta v_i^\ell \leq v_i^\ell,$$

or that $v_i^\ell \geq 0$. Examining (7.3), we can achieve this if we choose δ and ℓ so that δ^ℓ is close enough to 1, since $w_i > 0$.

In on-path mode a player's utility for deviating is at most

$$(1-\delta)M + \delta v_i^\ell.$$

Therefore, in order to make this deviation not profitable, we need to choose δ and ℓ in such a way that

$$w_i - (1-\delta)M - \delta v_i^\ell \geq 0. \tag{7.4}$$

Substituting the expression for v_i^ℓ and rearranging yields the condition

$$(1-\delta)M + \delta^{\ell+1} w_i - \delta(1-\delta^\ell)p_i \leq w_i.$$

Note that there is some balancing that needs to be done between δ and ℓ: The l.h.s. is a weighted average, and in order for it to be lower than w_i the weight of M must be sufficiently lower than the weight of p_i. The ratio between these weights is

$$\frac{\delta(1-\delta^\ell)}{1-\delta} = \delta + \delta^2 + \cdots + \delta^\ell.$$

This can be made arbitrarily large by choosing a large enough ℓ, and then taking δ to 1. This completes the proof of Theorem 7.8.

The picture is a little more complicated once the number of players is increased beyond 2. Consider the following 3 player base game: the actions available to each player are $A_i = \{0,1\}$, and the utility to each player is 1 if all players choose the same action, and 0 otherwise. Clearly, the minmax utilities here are $u_i^{mm} = 0$.

If we try to implement the idea of the two person proof to this game we immediately run into trouble, since there is no strategy profile $(b_1, b_2, \ldots, b_n)$ such that, for *every player* i, b_{-i} satisfies

$$b_{-i} \in \operatorname*{argmin}_{a_{-i}} \max_{a_i} u_i(a_{-i}, a_i).$$

To see this, assume that the above is satisfied for $i = 3$; that is, that b_{-3} is a minmax strategy for player 3. Then $b_1 \neq b_2$. Hence either $b_3 = b_1$ or $b_3 = b_2$. In the first case b_{-2} is not a minmax strategy for player 2, while in the second case b_{-1} is not a minmax strategy for player 1. In other words, for any strategy profile b there is a player who can guarantee a payoff of 1, either by playing b or by deviating from it.

In fact, it can be shown [13] that in this repeated game there are no perfect equilibria in which all players have utility less than 1! Fix a discount factor $\delta \in (0,1)$, and let

$$\alpha = \inf\{w : \exists \text{ a subgame perfect equilibrium with utility } w \text{ for all players}\}.$$

By the above observation, in any subgame there will be a player who can, by perhaps deviating, guarantee a payoff of at least $(1-\delta) + \delta\alpha$.

Now, for every $\varepsilon > 0$ there is a subgame perfect equilibrium in which the utility for each player is at most $\alpha + \varepsilon$. Hence

$$(1-\delta) + \delta\alpha \leq \alpha + \varepsilon,$$

or

$$1 - \frac{\varepsilon}{1-\delta} \leq \alpha.$$

Since this holds for every ε we have that $\alpha \geq 1$.

Note that in this game the set of feasible, enforceable payoff profiles is $\{(w,w,w) : w \in [0,1]\}$, which is one dimensional. It turns out that in base games in which this set has full dimension — i.e., dimension that is equal to the number of players — a folk theorem does apply. This result is also due to Fudenberg and Maskin [13].

Theorem 7.11. *Assume that the set of feasible, enforceable payoff profiles has dimension n. For every feasible,* strictly *enforceable payoff profile w and $\varepsilon > 0$ there is a $\delta_0 > 0$ such that for all $\delta > \delta_0$ there exists a perfect Nash equilibrium of G with δ-discounting utilities whose associated payoff profile w' satisfies $|w'_i - w_i| < \varepsilon$ for all $i \in N$.*

To prove this theorem, let w be a feasible, strictly enforceable payoff profile. Then there is some payoff profile z so that $z_j < w_j$ for all j. Furthermore, because of the full dimensionality assumption, for each $i = 1, \ldots, n$ there is a payoff profile z^i such that

- $z^i_i = z_i$.
- For $j \neq i$, $z_j < z^i_j < w_j$.

As in the two-player case, we will prove Theorem 7.11 for the case that there are action profiles $a^0, a^1, \ldots, a^n$ for G_0 that, respectively, realize the payoff profiles $w, z^1, \ldots, z^n$.

For each player i let b^i be the profile given by

$$b^i_{-i} \in \operatorname*{argmin}_{a_{-i}} \max_{a_i} u_i(a_{-i}, a_i),$$

with b^i_i a best response to b^i_{-i}. We consider the strategy profile s^* with the following modes.

- In on-path mode the players play a^0.
- In i-punishment mode, the players play b^i.
- In i-reconciliation mode the players play a^i.

The game starts in on-path mode. Assuming it stays there, the payoff profile is indeed w. If any player i deviates, the game enters i-punishment mode for some number of ℓ rounds. After these ℓ rounds the game enters i-reconciliation mode, in which it stays forever. A deviation by player j in i-reconciliation model or i-punishment mode are likewise met with entering j-punishment mode for ℓ periods, followed by j-reconciliation mode.

As in the two player case, denote by $\bar{u}_i = \max_{a \in A} u_i(a)$ the maximal utility achievable by player i in G_0.

For s^* to be an equilibrium we have to verify that there are no profitable deviations in any of the possible subgames. By Lemma 7.10 it suffices to check that no one-shot profitable deviation exists in them. Note that the possible subgames correspond to one-path mode, i-punishment mode with k periods left, and i-reconciliation mode.

The equilibrium path utility for player j is w_j in on-path mode. In i-punishment mode with k periods left it is

$$(1-\delta^k)u_j(b^i) + \delta^k z_j^i,$$

which we will denote by $v_j^{i,k}$. Note that $u_i(b^i) = 0$ by the definition of b^i, and so

$$v_i^{i,k} = \delta^k z_i^i.$$

In i-reconciliation mode the utility on equilibrium path for player j is z_j^i.

For a deviation of player j in on-path mode to *not* be profitable it suffices to ensure that

$$(1-\delta)\bar{u}_j + \delta v_j^{j,\ell} \leq w_j.$$

Substituting $v_j^{j,\ell}$ yields

$$(1-\delta)\bar{u}_j + \delta^{\ell+1} z_j^j \leq w_j.$$

Since $z_j^j < w_j$ this holds for all δ close enough to 1. Similarly, in i-reconciliation mode it suffices that

$$(1-\delta)\bar{u}_j + \delta^{\ell+1} z_j^j \leq z_j^i,$$

which holds for all δ close enough to 1 and ℓ large enough, since $z_j^j \leq z_j^i$.

In i-punishment mode with k periods left there is clearly no profitable deviation for i, who is already best-responding to her punishment b_{-i}. For there to not be a profitable deviation for $j \neq i$, it must hold that

$$(1-\delta)\bar{u}_j + \delta v_j^{j,\ell} \leq v_j^{i,k}.$$

Substituting yields

$$(1-\delta)\bar{u}_j + \delta^{\ell+1} z_j^j \leq (1-\delta^k)u_j(b^i) + \delta^k z_j^i.$$

By again choosing δ close enough to 1 we can make the left hand side smaller than $\delta^k z_j^i$, since $z_j^j < z_j^i$ (recall that $j \neq i$), and thus smaller than the right hand side. This completes the proof of Theorem 7.11, for the case that the relevant payoff profiles can be realized using pure strategy profiles.

7.6 Finitely repeated games

In this section we consider a finitely repeated game with T periods. The utility will always be the sum of the stage utilities:

$$v_i(a^1, a^2, \ldots, a^T) = \frac{1}{T} \sum_{t=1}^{T} u_i(a^t).$$

7.6.1 Nash equilibria and folk theorems

A simple but important first observation about finitely repeated games is the following.

Claim 7.12. *In every Nash equilibrium of a finitely repeated game, the last action profile played on path is a Nash equilibrium of the base game.*

In general finitely repeated games one cannot hope to prove a folk theorem that is as strong as those available in infinitely repeated games, as the following example illustrates.

Let G_0 be the following prisoners' dilemma:

	D	C
D	6,6	14,2
C	2,14	10,10

Claim 7.13. *In every Nash equilibrium of G both players play D in every period on the equilibrium path.*

Proof. Let s^* be a Nash equilibrium of the repeated game. Assume by contradiction that, when playing s^*, some player i plays C on some period, and let t_ℓ be the last such period.

Let s_i be the strategy for player i that is identical to s_i^* in all periods $t < t_\ell$, and under which, for periods $t \geq t_\ell$, player i always plays D. We claim that s_i is a profitable deviation: the stage utilities for i in periods $t < t_\ell$ are the same under s_i and s_i^*. In period t_ℓ the utility is strictly larger, since D is a strictly dominant strategy. In periods $t > t_\ell$ both players played D under s^* (by the definition of t_ℓ), and so the utility for player i under s_i is either the same (if the other player still plays D) or greater than the utility under s^* (if the other player now plays C in some of the periods). $\square$

Hence the payoff profile in every Nash equilibrium (subgame perfect or not) is $(6,6)$. This is in stark contrast to the infinitely repeated case.

This result can be extended to any game in which every equilibrium achieves minmax payoff profiles. In contrast we consider games in which there is a Nash equilibrium in which every player's payoff is larger than her minmax payoff. In such games we again have a strong folk theorem [20].

An example of such a game is the following:

	D	C	F
D	6,6	14,2	6,6
C	2,14	10,10	6,6
F	6,6	6,6	7,7

In this game there is an equilibrium in which both players play (C,C) for all periods except the last four, in which both players play (F,F). If anyone deviates the other plays D henceforth, thus erasing any possible gains achieved by deviating.

Theorem 7.14. *Assume that G_0 has a Nash equilibrium a^* whose associated payoff profile w^* satisfies $w^*_i > u^{mm}_i$. Let a be an action profile in G_0 whose payoff profile w is strictly enforceable. Then for any $\varepsilon > 0$ and T large enough there is a Nash equilibrium of G with payoff profile w' such that $|w_i - w'_i| < \varepsilon$ for all i.*

Proof. Consider a strategy profile s^* with the following modes:

- In on-path mode, players play a in all periods except the last ℓ periods, in which they play a^*.
- In i-punishment mode the players play b^i where
$$b^i_{-i} \in \operatorname*{argmin}_{a_{-i}} \max_{a_i} u_i(a_{-i}, a_i)$$
and b^i_i a best response to b^i_{-i}

The game starts in on-path mode, and switches to i-punishment mode for the rest of the game if i deviates.

Fix T. We will show that s^* is an equilibrium for an appropriate choice of ℓ, and assuming T is large enough. Note that player i's utility when no one deviates is
$$T \cdot v_i(s^*) = (T - \ell)w_i + \ell w^*_i.$$

Assume by contradiction that player i has a profitable deviation s_i, and let t be the last time period in which she does not deviate. Clearly, t cannot occur in the last ℓ periods, since in these the players play an equilibrium and thus there is no profitable deviation. Consider then the case that t is not in the last ℓ periods. Then
$$T \cdot v_i(s^*_{-i}, s_i) \le t w_i + M + (T - \ell - t - 1)u^{mm}_i + \ell u^{mm}_i,$$
where M is the maximum she can earn at period t. In particular, since $u^{mm}_i < w_i$ we have that
$$T \cdot v_i(s^*_{-i}, s_i) < (T - \ell - 1)w_i + M + \ell u^{mm}_i.$$
Hence
$$T \cdot (u_i(s^*) - u_i(s^*_{-i}, s_i)) = w_i - M + \ell(w^*_i - u^{mm}_i).$$
Therefore, if we choose
$$\ell > \frac{M - w_i}{w^*_i - u^{mm}_i},$$
which is only possible if T is at least this large, then this will not be a profitable deviation.

Finally, the utility for playing s^* can be written as
$$v_i(s^*) = w_i + \frac{\ell}{T}(w^*_i - w_i + \ell w^*_i),$$
and so $|v_i(s^*) - w_i| < \varepsilon$ for T large enough. □

7.6.2 Perfect Nash equilibria and folk theorems

Claim 7.15. *In every subgame perfect Nash equilibrium of a finitely repeated game, the last action profile played after any history is a Nash equilibrium of the base game.*

Exercise 7.16. *Show that if the base game has a unique Nash equilibrium a^* then the payoff profile of any subgame perfect equilibrium of the repeated game is the same as that of a^*.*

When there are sufficiently many diverse equilibria of the base game it is possible to prove a perfect folk theorem for the finite repeated game.

8 Social learning

8.1 Bayesian hypothesis testing

Consider the following single agent Bayesian game. There is a *state of nature* $S = \{0,1\}$, which is initially picked with some probability $p = \mathbb{P}[S = 1]$. There is an experiment that the agent can carry out repeatedly, and where each execution $i = 1,\dots,n$ results in an observation X_i, which is picked from some finite set Ω. If $S = 0$ then the X_i's are distributed i.i.d. with some distribution ν_0. If $S = 1$ then they are distributed i.i.d. with distribution ν_1. We say the the *private signals* $\{X_i\}_{i=1}^n$ are *conditionally independent*.

A common example is $\Omega = \{H,T\}$, $\nu_0(H) = \nu_1(T) = 0.6$ and $\nu_1(H) = \nu_0(T) = 0.4$. We will assume (as in this example) that ν_0 and ν_1 both given positive probability to every $\omega \in \Omega$.

Now, after observing her n signals the agent has to guess whether the state of nature is 0 or 1. That is, she has to take an action $a \in \{0,1\}$, and her utility is 1 if $a = S$ and zero otherwise. Hence the agent's expected utility is equal to $\mathbb{P}[a = S]$.

It is easy to see that the only rational choices are to choose $a = 1$ if $\mathbb{P}[S = 1|X_1,\dots,X_n] \geq \frac{1}{2}$, and to choose $a = 0$ if $\mathbb{P}[S = 1|X_1,\dots,X_n] \leq \frac{1}{2}$. We will assume for now that p,ν_0,ν_1 are chosen in such a way that we never have equality (i.e., indifference), and so

$$a = \operatorname*{argmax}_{s\in\{0,1\}} \mathbb{P}[s = S|X_1,\dots,X_n].$$

Now,

$$\mathbb{P}[a = S] = \mathbb{P}[a = 1|S = 1]\cdot p + \mathbb{P}[a = 0|S = 0]\cdot(1-p).$$

Let us calculate $\mathbb{P}[a = 1|S = 1]$. By the remarks above the event $a = 1$ is equal to the event $\mathbb{P}[S = 1|X_1,\dots,X_n] > \frac{1}{2}$. The latter is equal to the event that

$$\frac{\mathbb{P}[S = 1|X_1,\dots,X_n]}{\mathbb{P}[S = 0|X_1,\dots,X_n]} > 1.$$

We can rewrite this as

$$\frac{\mathbb{P}[X_1,\dots,X_n|S = 1]}{\mathbb{P}[X_1,\dots,X_n|S = 0]}\cdot\frac{p}{1-p} > 1.$$

By conditional independence, this is the same as

$$\sum_{i=1}^n \log\frac{\mathbb{P}[X_i|S = 1]}{\mathbb{P}[X_i|S = 0]}(X_i) + \ell_0 > 0.$$

Let

$$L_i = \log\frac{\nu_1(X_i)}{\nu_0(X_i)}$$

be the *private log likelihood ratio* of signal X_i. Denote $M_i = \sum_{k=1}^{i} L_k$. Then we have shown that $a = 1$ iff $\ell_0 + M_n > 0$. To simplify calculations henceforth we assume that $\ell_0 = 0$.

Note that, conditioned on $S = 1$, the L_i's are i.i.d. random variables. Their expectation is

$$\mathbb{E}[L_k|S=1] = \sum_{\omega\in\Omega} \nu_1(\omega)\cdot\log\frac{\nu_1(\omega)}{\nu_0(\omega)}.$$

This is also called the *Kullback-Leibler divergence* between ν_1 and ν_0, and is a measure of how different the distributions are. It is easy to show that this number is always non-negative, and is zero iff $\nu_0 = \nu_1$.

To estimate $\mathbb{P}[a=1|S=1]$ we need to estimate the probability that $M_n > 0$. To this end we will need a few definitions. First, denote $L = L_1$, and let $\varphi\colon \mathbb{R}\to\mathbb{R}$ be given by

$$\varphi(t) = -\log\mathbb{E}\left[e^{-t\cdot L}\middle|S=1\right].$$

Note that $\mathbb{E}\left[e^{-t\cdot L}\right]$ is called the *Laplace transform* of L.

Proposition 8.1. *1.* $\varphi(0) = \varphi(1) = 0$.

2. φ is smooth and concave.

We will not prove this claim here. Note that it follows from it that φ is positive in the interval $(0,1)$.

Now, the probability of mistake (that is, the probability that $a = 0$ conditioned on $S = 1$) is, for any $t > 0$,

$$\mathbb{P}[M_n < 0|S=1] = \mathbb{P}\left[e^{-t\cdot M_n} > 1\middle|S=1\right].$$

Since $e^{-t\cdot M_n}$ is a positive random variable, we can apply Markov's inequality and write

$$\mathbb{P}[M_n < 0|S=1] \le \mathbb{E}\left[e^{-t\cdot M_n}\middle|S=1\right].$$

Now, recall that $M_n = \sum_i L_i$, and that the L_i's are conditionally independent. Hence

$$\begin{aligned}\mathbb{P}[M_n < 0|S=1] &\le \mathbb{E}\left[e^{-t\sum_{i=1}^n L_i}\middle|S=1\right]\\ &= \prod_{i=1}^n \mathbb{E}\left[e^{-t\cdot L_i}\middle|S=1\right]\\ &= e^{-\varphi(t)\cdot n}.\end{aligned}$$

This holds for any $t > 0$, and so we have show that

Theorem 8.2. $\mathbb{P}[a \neq 1|S=1] \le \inf_{t>0} e^{-\varphi(t)\cdot n}$.

Since φ is positive in the interval $(0,1)$, it follows that $\mathbb{P}[a\neq 1|S=1]$ decreases (at least) exponentially with n. It turns out (but will not be proven here) that this estimate is asymptotically tight: if we denote $r = \sup_{t>0}\varphi(t)$, then $\mathbb{P}[a\neq 1|S=1] = e^{-r\cdot n + o(n)}$.

8.2 Herd behavior

The text of this section is largely taken from [25].

Banerjee [7] and concurrently Bikhchandani, Hirshleifer and Welch [9], consider the following model: There is a binary state of the world $S \in \{0,1\}$, conditionally i.i.d. private signals with conditional distributions ν_1 and ν_0. There is an infinite sequence of agents $N = \{1,2,\ldots\}$, each one with a single private signal X_i. For simplicity, we consider the prior $\mathbb{P}[S=1] = \mathbb{P}[S=0] = 1/2$. We assume that ν_0 and ν_1 satisfy the *bounded likelihood ratio* assumption: that is, that there is some $M > 0$ such that, with probability one,

$$L_i = \log \frac{\mathrm{d}\nu_1}{\mathrm{d}\nu_0}(X_i)$$

is in $(-M, M)$. This is satisfied, for example, when ν_0 and ν_1 have the same finite support.

The agents act sequentially, with agent i having to choose $a_i \in \{0,1\}$ and receiving utility $u_i(a_i, S) = \mathbb{1}_{\{a_i = S\}}$. Each agent i, in addition to her private signal, observes the actions of her predecessors $\{1,2,\ldots,i-1\}$. Hence

$$a_i = \operatorname*{argmax}_{s \in \{0,1\}} \mathbb{P}[s = S | a_1, \ldots, a_{i-1}, X_i].$$

We assume that $a_i = 1$ whenever the agent is indifferent.

By Theorem 8.2, if agent i had access to the private signals of her predecessors then she should choose the correct action, except with exponentially small probability. It turns out that this is not exactly what happens here.

Theorem 8.3. *The limit* $\lim_i \mathbb{P}[a_i \neq S]$ *exists and is strictly greater than zero.*

To analyze this model it is useful to consider an outside observer x who sees the agents' actions, but not their private signals. Thus the information available to x at round i is $\{a_1, a_2, \ldots, a_{i-1}\}$. We denote by

$$B_x^i = \mathbb{P}[S = 1 | a_1, \ldots, a_{i-1}]$$

x's belief at period i. It follows from Lévy's zero-one law (Theorem 4.22) that

$$B_x^\infty := \mathbb{P}[S = 1 | a_1, a_2, \ldots] = \lim_i B_x^i.$$

We also define an action a_x^i for x at round i; this is given by

$$a_x^i = \operatorname*{argmax}_{s \in \{0,1\}} \mathbb{P}[s = S | a_1, \ldots, a_{i-1}],$$

where, as with the actual players of the game, indifference results in choosing $a_x^i = 1$.

Claim 8.4. $a_x^i = a_{i-1}$.

That is, the outside observer simply copies the action of the last agent she observes. The proof is simple, and follows from the fact that at time i, agent $i-1$ has more information than x does. It follows that $\mathbb{P}[a_{i-1}=S]=\mathbb{P}\left[a_x^i=S\right]$, which will help us prove Theorem 8.3.

We define agent i's log-likelihood ratio by

$$L_i^a = \log\frac{\mathbb{P}[S=1|a_1,\ldots,a_{i-1},X_i]}{\mathbb{P}[S=0|a_1,\ldots,a_{i-1},X_i]}.$$

As before, by Bayes' law we can write this as

$$L_i^a = \log\frac{\mathbb{P}[a_1,\ldots,a_{i-1},X_i|S=1]}{\mathbb{P}[a_1,\ldots,a_{i-1},X_i|S=0]}.$$

Now, the actions $a_1,\ldots,a_{i-1}$ are not conditionally independent. However, they are conditionally independent of X_i, and so

$$L_i^a = \log\frac{\mathbb{P}[a_1,\ldots,a_{i-1}|S=1]}{\mathbb{P}[a_1,\ldots,a_{i-1}|S=0]}+\log\frac{\mathbb{P}[X_i|S=1]}{\mathbb{P}[X_i|S=0]}.$$

If we denote by

$$L_x^i = \log\frac{B_x^i}{1-B_x^i} = \log\frac{\mathbb{P}[a_1,\ldots,a_{i-1}|S=1]}{\mathbb{P}[a_1,\ldots,a_{i-1}|S=0]}$$

the outside observer's log-likelihood ratio then we have that

$$L_i^a = L_x^i + L_i. \tag{8.1}$$

We observe here without proof (although it is straightforward) that there is a *continuous* function $f:[0,1]\times\{0,1\}\to[0,1]$ such that for all rounds i

$$B_x^{i+1} = f(B_x^i, a_i). \tag{8.2}$$

That is, to update her belief the outside observer need only know the new action she observes, and this calculation is the same in all rounds. The key observation behind this is that

$$B_x^{i+1} = \mathbb{P}[S=1|a_1,\ldots,a_{i-1},a_i] = \mathbb{P}\left[S=1\middle|B_x^i,a_i\right] = f(B_x^i,a_i).$$

That is, B_x^i already includes all the S-relevant information contained in $a_1,\ldots,a_{i-1}$, and so, given B_x^i, the history before a_i is irrelevant for making inferences regarding S.

Theorem 8.5. *The limit* $\lim_i a_i$ *exists almost surely.*

Proof. As we noted above, $a_i = a_x^{i+1}$, and so it suffices to show that $\lim_i a_x^i$ exists almost surely.

Assume by contradiction that a_x^i takes both values infinitely often. Hence B_x^i is infinitely often above 1/2, and infinitely often below 1/2. It therefore converges to 1/2.

Taking the limit of (8.2), it follows from the continuity of f that $f(1/2,0)=f(1/2,1)=1$. But $B_x^0 = 1/2$, and so $B_x^i = 1/2$ for all i. Hence $a_x^i = 1$ for all i, and we have reached a contradiction. □

Since $\lim_i a_i$ exists almost surely we can define a random variable $a = \lim_i a_i = \lim_i a_x^i$; this is the action that almost all agents choose. To prove Theorem 8.3 it remains to show that $\mathbb{P}[a = S] \neq 1$. An important observation is that

$$\mathbb{P}\left[a_i^x = S \middle| a_1, \ldots, a_{i-1}\right] = \max\{B_x^i, 1 - B_x^i\}.$$

It therefore suffices to show that $\lim_i B_x^i$ is almost surely in $(0,1)$.

Since the private signals have bounded log-likelihood ratios (say with bound M), it follows from (8.1) that when $L_x^i > M$ then, with probability one, $L_i^a > 0$, and hence (again with probability one) $a_i = 1$. Thus when $L_x^i > M$ it is not informative to observe a_i; the outside observer already knew that agent i would choose $a_i = 1$. Hence, in this case, $B_x^{i+1} = B_x^i$. It follows that, with probability one, $L_x^i < 2M$, and thus B_x^i is bounded away from one. An identical argument shows that it is bounded away from zero. This completes a sketch of the proof of Theorem 8.3.

9 Better response dynamics and potential games

9.1 Better response dynamics

Better response dynamics describe a mechanism by which a group of players playing a game might find an equilibrium.

Formally, let $G = (N, \{S_i\}, \{u_i\})$ be a strategic form game. A *better response path* is a sequence of strategy profiles s^k such that

1. Each $s^k = (s^k_1, \dots, s^k_n)$ differs from its successor s^{k+1} by exactly one coordinate.

2. If $s^k_i \neq s^{k+1}_i$ then $u_i(s^{k+1}) > u_i(s^k)$. That is, s'_i is a profitable deviation for i.

Exercise 9.1. *Explain why in a game with no Nash equilibria every better response path* $(s^1, s^2, \dots, s^k)$ *can be extended by some* s^{k+1} *to a better response path* $(s^1, s^2, \dots, s^k, s^{k+1})$*. Conclude that such games have infinite better response paths.*

Exercise 9.2. *Find a finite game in which there exists an infinite better response path, as well as a pure Nash equilibrium.*

Exercise 9.3. *Explain why there are no infinite better response paths in prisoners' dilemma.*

Proposition 9.4. *If a game has no infinite better response paths then it has a pure Nash equilibrium.*

Proof. Since there are no infinite response paths then there must be a path $(s^1, s^2, \dots, s^k)$ that cannot be extended by any s^{k+1} to a longer path. Thus in s^k no player has a profitable deviation, and so s^k is a pure Nash equilibrium. □

9.2 A congestion game

Let $G = (V, E)$ be a finite graph, where we think of each directed edge $e = (v, w)$ as representing a road between cities v and w. For each edge $e = (v, w)$ let $c_e : \mathbb{N} \to \mathbb{R}^+$ be a monotone increasing *congestion function*, where $c_e(k)$ is the travel time when k cars are on the road from v to w.

Consider a finite set of players N. Each player i has to travel from some city v_i to some other city w_i, and so has to choose a path s_i that connects v_i to w_i. We assume that the chosen paths are always *simple* (i.e., do not repeat edges) and so we can think of s_i simply as a subset of E.

Consider the game G in which each player i's set of strategies is the set of all simple paths from v_i to w_i. Given a strategy profile $s = (s_1, \dots, s_n)$ and an edge e, we denote by $n_e(s)$ the number of players who travel on e:

$$n_e(s) = |\{i \in N : e \in s_i\}|.$$

Hence the travel time on e when the players choose s is $c_e(n_e(s))$. Player i's utility is minus her travel time:

$$u_i(s) = -\sum_{e \in s_i} c_e(n_e(s)).$$

Proposition 9.5. *G has a pure Nash equilibrium.*

To prove this, let $S = \prod_i S_i$ denote the set of strategy profiles, and define the function $\Phi\colon S \to \mathbb{R}$ by

$$\Phi(S) = -\sum_{e \in E} \sum_{k=1}^{n_e(s)} c_e(k).$$

Note that Φ is not the social welfare: that is equal to $-\sum_e n_e(s) c_e(n_e(s))$.

Claim 9.6. *For every $s = (s_{-i}, s_i) \in S$ and $s' = (s_{-i}, s'_i) \in S_i$ it holds that*

$$u_i(s') - u_i(s) = \Phi(s') - \Phi(s).$$

That is, the change in utility for player i when switching from s_i to s'_i is equal to the change in Φ that caused by this switch. The proof of this claim is left as an exercise.

Note that the existence of this Φ implies that there is a pure NE for this game, since any s such that $\Phi(s)$ is maximal has to be a NE, and Φ attains its maximum since it has finite domain. Thus we have proved 9.5.

In fact, we can prove an even stronger statement:

Claim 9.7. *G has no infinite better response paths.*

Proof. Assume by contradiction that $(s^1, s^2, \dots)$ is an infinite better response path. Since the game is finite, there must be some $\varepsilon > 0$ such that the improvement in utility for each deviating player in each s^{k+1} is at least ε. It then follows from Claim 9.6 that $\Phi(s^{k+1})$ is at least $\Phi(s^1) + k \cdot \varepsilon$, for every k. But Φ is bounded since it has finite domain, and so we have arrived at a contradiction. □

9.3 Potential games

Let $G = (N, \{S_i\}, \{u_i\})$ be a strategic form game. We say that G is a potential game if there exists a $\Phi\colon S \to \mathbb{R}$ with the same property as in the example above: For every $s = (s_{-i}, s_i) \in S$ and $s' = (s_{-i}, s'_i) \in S_i$ it holds that

$$u_i(s') - u_i(s) = \Phi(s') - \Phi(s).$$

The proof of 9.7 applies to any finite potential game, showing that they have no infinite better response paths. Thus better response dynamics always converges to a pure Nash equilibrium for finite potential games.

9.4 The conformism game

Let $G = (V, E)$ be a *social network graph*: the nodes are players and $e = (i, j)$ is an edge if i and j are friends. We assume that $(i, j) \in E$ iff $(j, i) \in E$, so that all friendships are bidirectional.

Consider the following strategic form game. For all players the set of strategies is $\{0, 1\}$. A player's payoff is the number of her neighbors who choose the same strategy:

$$u_i(s_1, \ldots, s_n) = \sum_{j:(i,j)\in E} \mathbb{1}_{\{s_i = s_j\}}.$$

Exercise 9.8. *Show that this is a potential game whenever the number of players is finite.*

The same holds for the *hipsters game*, where the utility is

$$u_i(s_1, \ldots, s_n) = \sum_{j:(i,j)\in E} \mathbb{1}_{\{s_i \neq s_j\}}.$$

Exercise 9.9. *Find an infinite graph (but where each player has finitely many neighbors) in which this game has an infinite better response path.*

Exercise 9.10. *Find an infinite connected graph (where each player has finitely many neighbors) for which, in this game, every infinite better response path includes each player only finitely many times. Find a graph for which this is not the case.*

Given a graph $G = (V, E)$ and a vertex $v \in V$, let $f_v(r)$ equal the number of vertices that are of distance exactly r from v in G.

Theorem 9.11. *If in $G = (V, E)$ there is a bound on the degrees, and if $f_v(r)$ is sub-exponential, then there exists a $C \in \mathbb{N}$ such that, in any better response path, agent v participates at most C times.*

10 Social choice

10.1 Preferences and constitutions

Consider a set of n voters $N = \{1,\dots,n\}$ who each have a preference regarding k alternatives A. A *preference* [12] or a *ranking* here is a bijection from A to $\{1,\dots,k\}$, so that if some $a \in A$ is mapped to 1 then it is the least preferred alternative, and if it is mapped to k then it is the more preferred. We denote the set of all preferences P_A.

A *profile (of preferences)* $\pi = (\pi_1,\dots,\pi_n) \in P_A^n$ includes a preference for each voter. A *constitution* is a map from P_A^n to P_A, assigning to each profile a preference called the *social preference*. Given a constitution $\varphi \colon P_A^n \to P_A$ and a profile π we will sometimes write φ_π instead of the usual $\varphi(\pi)$.

A simple example of a constitution is a *dictatorship*: if we define $d : P_A^n \to P_A$ by $d(\pi_1,\dots,\pi_n) = \pi_1$ then d is a constitution in which the preferences of voter 1 are always adopted as the social preference.

When $k = 2$ (in which case we will denote $A = \{a,b\}$) and n is odd, a natural example of a constitution is *majority rule* $m \colon P_A^n \to P_A$ given by

$$m_\pi(a) = \begin{cases} 1 & \text{if } |\{i : \pi_i(a) < \pi_i(b)\}| > n/2 \\ 2 & \text{otherwise} \end{cases}.$$

Majority rule has a few desirable properties that we now define for a general constitution.

- A constitution φ satisfies *non-dictatorship* if it is not a dictatorship. It is a dictatorship if there exists an $i \in N$ such that $\varphi(pi) = \pi_i$ for all profiles π.
- A constitution φ is said to satisfy *unanimity* if, for any profile π and pair $a,b \in A$ it holds that if $\pi_i(a) < \pi_i(b)$ for all $i \in N$ then $\varphi_\pi(a) < \varphi_\pi(b)$.
- A constitution φ is said to satisfy the *weak Pareto principle* (WPP) if for any profile π such that $\pi_1 = \pi_2 = \cdots = \pi_n$ it holds that $\varphi(\pi) = \pi_1$. Clearly unanimity implies WPP, but not vice versa.
- A constitution φ is said to be *anonymous* if for any permutation $\eta \colon N \to N$ it holds that $\varphi(\pi_{\eta(1)},\dots,\pi_{\eta(n)}) = \varphi(\pi_1,\dots,\pi_n)$ for every profile π. More generally, we say that a permutation η is a *symmetry* of φ if the above holds, and so φ is anonymous if it has every permutation as a symmetry. We say that φ is *equitable* if for any two voters $i,j \in V$ there exists a symmetry η of φ such that $\eta(i) = j$. That is, φ is equitable if its group of symmetries acts transitively on the set of voters.
- A constitution φ is said to be *neutral* if it is indifferent to a renaming of the alternatives: for any permutation $\zeta \colon A \to A$ it holds that $\varphi(\pi_1 \circ \zeta,\dots,\pi_n \circ \zeta) = \varphi(\pi) \circ \zeta$.

It is easy to verify that majority rule has these properties.

[12] This is usually known in the literature as a strict preference.

10.2 The Condorcet Paradox

When there are three alternatives or more one could imagine generalizing majority rule to a constitution φ that satisfies that $\varphi_\pi(a) < \varphi_\pi(b)$ for any $a, b \in A$ such that $\pi_i(a) < \pi_i(b)$ for the majority of voters. The *Condorcet paradox* is the observation that this is impossible [11]. To see this, consider three voters with the following preferences:

$$\begin{aligned} &\pi_1(c) < \pi_1(b) < \pi_1(a) \\ &\pi_2(a) < \pi_2(c) < \pi_2(b) \\ &\pi_3(b) < \pi_3(a) < \pi_3(c). \end{aligned} \tag{10.1}$$

Here, two voters prefer a to b, two prefer b to c, and two prefer c to a. Thus it is impossible to rank the three alternatives in a way that is consistent with the majority rule applied to each pair.

10.3 Arrow's Theorem

Condorcet's paradox shows us that we cannot choose the relative ranking of each pair using majority. Can we still choose the relative ranking of each pair independently, using some other rule?

A constitution φ is said to satisfy *independence of irrelevant alternatives* (IIA) if the relative ranking of a and b in $\varphi(\pi)$ is determined by their relative rankings in π. Formally, given two alternatives $a, b \in A$ and profiles π, π' such that, for all $i \in N$,

$$\pi_i(a) < \pi_i(b) \text{ iff } \pi'_i(a) < \pi'_i(b),$$

it holds that

$$\varphi_{\pi'}(a) < \varphi_{\pi'}(b) \text{ iff } \varphi_{\pi'}(a) < \varphi_{\pi'}(b).$$

Arrow's Theorem [1, 2] shows that any constitution for three or more alternatives that satisfies IIA is in some sense trivial.

Theorem 10.1 (Arrow's Theorem)**.** *Let φ be a constitution for $|A| \geq 3$ alternatives and $|N| \geq 2$ voters. If φ satisfies unanimity and IIA then φ is a dictatorship.*

In fact, the theorem still holds if we replace unanimity with WPP.

10.4 The Gibbard-Satterthwaite Theorem

As above, let A be a set of alternatives and N a set of voters. A *social choice function* $f : P_A^n \to A$ chooses a single alternative, given a preference profile.

We can associate with every social choice function f and preference profile π a natural game $G_{f,\pi}$ played between the voters. In this game the strategy space is the space of preferences P_A, and voter i's utility u_i when the voters play $(s_1, \ldots, s_n) \in P_A^n$ is $\pi_i(f(s))$. That is, a voter's utility is the ranking of the chosen alternative.

We say that a social choice function is *strategy-proof* if for any preference profile π, in the game $G_{f,\pi}$ it holds that for every player i the strategy π_i is weakly dominant. That is, f is strategy proof if for every $\pi \in P_A^n$, $i \in N$ and $\tau_i \in P_A$ it holds that $\pi_i(f(\pi_{-i},\pi_i)) \geq \pi_i(f(\pi_{-i},\tau_i))$. Informally, if f is strategy proof then players always get the best result by reporting their true preferences.

We say that a social choice function f is a *dictatorship* if there is some voter i such that $f(\pi_1,\ldots,\pi_n) = \mathrm{argmax}_a\, \pi_i(a)$ for every $\pi \in P_A^n$.

We say that a social choice function f satisfies the *Pareto principle* if for all $a \in A$ it holds that $f(\pi) \neq a$ whenever there exists a $b \in A$ such that $\pi_i(a) < \pi_i(b)$ for all $i \in N$.

Claim 10.2. *Let $|N|$ be odd and let $|A| = 2$. Then majority rule is strategy-proof and satisfies the Pareto principle.*

The proof is left as an exercise to the reader.

In the next claim we define the *plurality* social choice function and prove that it is not strategy proof, but satisfies Pareto optimality.

Claim 10.3. *Let $|A| = 3$, and let the social choice function $p: P_A^n \to A$ be given by $p(\pi) = a$ if a was ranked the highest more times than any other alternative. If there is more than one alternative that is ranked highest the most times then whichever of the highest ranked alternatives that is lexicographically earlier is chosen. Then f is not strategy proof and satisfies the Pareto principle.*

Proof. We prove for the case that $N = \{1,2,3\}$; the proof for the general case is similar.

Let $A = \{a,b,c\}$. Let $\pi_1(a) < \pi_1(b) < \pi_1(c)$, $\pi_2(a) < \pi_2(c) < \pi_2(b)$ and $\pi_3(b) < \pi_3(c) < \pi_3(a)$. Then $p(\pi) = b$, but $p(\pi_1,\pi_2,\tau_3) = c$, where $\tau_3(a) < \tau_3(b) < \tau_3(c)$. Since $\pi_3(b) < \pi_3(c)$ it follows that p is not strategy proof.

If $\pi_i(a) < \pi_i(b)$ for all i then a is ranked highest by no voter, and so $f(\pi) \neq a$. Thus p satisfies the Pareto principle. ☐

The Gibbard-Satterthwaite Theorem [16, 31] states that any strategy-proof social choice function for three or more alternatives is in some sense trivial.

Theorem 10.4 (Gibbard-Satterthwaite)**.** *Let $|A| \geq 3$, and let $f: P_A^n \to P_A$ be a social choice function that satisfies the Pareto principle. Then if f is strategy-proof then it is a dictatorship.*

In fact, this theorem can be proved even if we replace the Pareto principle with the condition that for every $a \in A$ there exists a $\pi \in P_A^n$ such that $f(\pi) = a$.

References

[1] K. Arrow. A difficulty in the concept of social welfare. *J. of Political Economy*, 58:328–346, 1950.

[2] K. Arrow. *Social choice and individual values*. John Wiley and Sons, 1963.

[3] R. J. Aumann. Agreeing to disagree. *The annals of statistics*, pages 1236–1239, 1976.

[4] R. J. Aumann. Correlated equilibrium as an expression of bayesian rationality. *Econometrica*, 55(1):1–18, 1987.

[5] R. J. Aumann and S. Hart. Long cheap talk. *Econometrica*, 71(6):1619–1660, 2003.

[6] R. J. Aumann and L. S. Shapley. *Long-term competition—a game-theoretic analysis*. Springer, 1994.

[7] A. V. Banerjee. A simple model of herd behavior. *The Quarterly Journal of Economics*, 107(3):797–817, 1992.

[8] P. Battigalli. Strategic independence and perfect bayesian equilibria. *Journal of Economic Theory*, 70(1):201 – 234, 1996.

[9] S. Bikhchandani, D. Hirshleifer, and I. Welch. A theory of fads, fashion, custom, and cultural change as informational cascades. *Journal of political Economy*, pages 992–1026, 1992.

[10] P. Dasgupta, P. Hammond, and E. Maskin. The implementation of social choice rules: Some general results on incentive compatibility. *The Review of Economic Studies*, 46(2):185–216, 1979.

[11] J.-A.-N. de Condorcet. *Essai sur l'application de l'analyse à la probabilité des décisions rendues à la pluralité des voix*. De l'Imprimerie Royale, 1785.

[12] J. W. Friedman. A non-cooperative equilibrium for supergames. *The Review of Economic Studies*, 38(1):1–12, 1971.

[13] D. Fudenberg and E. Maskin. The folk theorem in repeated games with discounting or with incomplete information. *Econometrica*, 54(3):533–554, 1986.

[14] D. Fudenberg and J. Tirole. Perfect bayesian equilibrium and sequential equilibrium. *Journal of Economic Theory*, 53(2):236–260, 1991.

[15] J. D. Geanakoplos and H. M. Polemarchakis. We can't disagree forever. *Journal of Economic Theory*, 28(1):192–200, 1982.

[16] A. Gibbard. Manipulation of voting schemes: a general result. *Econometrica*, 41(4):587–601, 1973.

[17] I. Gilboa, E. Kalai, and E. Zemel. On the order of eliminating dominated strategies. *Operations Research Letters*, 9(2):85–89, 1990.

[18] P. W. Goldberg and C. H. Papadimitriou. Reducibility among equilibrium problems. In *Proceedings of the thirty-eighth annual ACM symposium on Theory of computing*, pages 61–70. ACM, 2006.

[19] B. R. Holmström. *On incentives and control in organizations*. Stanford University, 1977.

[20] V. K. Jean-Pierre Benoit. Finitely repeated games. *Econometrica*, 53(4):905–922, 1985.

[21] D. M. Kreps and R. Wilson. Sequential equilibria. *Econometrica*, 50(4):863–894, 1982.

[22] H. W. Kuhn. Extensive games and the problem of information. *Contributions to the Theory of Games*, 2(28):193–216, 1953.

[23] M. Maschler, E. Solan, and S. Zamir. Game theory, 2013.

[24] P. Milgrom and N. Stokey. Information, trade and common knowledge. *Journal of Economic Theory*, 26(1):17–27, 1982.

[25] E. Mossel and O. Tamuz. Opinion exchange dynamics. *arXiv preprint arXiv:1401.4770*, 2014.

[26] R. B. Myerson. Incentive compatibility and the bargaining problem. *Econometrica*, 47(1):61–73, 1979.

[27] J. F. Nash. Equilibrium points in n-person games. *Proc. Nat. Acad. Sci. USA*, 36(1):48–49, 1950.

[28] B. O'Neill. Nonmetric test of the minimax theory of two-person zerosum games. *Proceedings of the National Academy of Sciences*, 84(7):2106–2109, 1987.

[29] M. J. Osborne and A. Rubinstein. *A course in game theory*. MIT press, 1994.

[30] A. Rubinstein. Strong perfect equilibrium in supergames. *International Journal of Game Theory*, 9(1):1–12, 1980.

[31] M. A. Satterthwaite. Strategy-proofness and Arrow's conditions: Existence and correspondence theorems for voting procedures and social welfare functions. *Journal of economic theory*, 10(2):187–217, 1975.

[32] R. Selten. Reexamination of the perfectness concept for equilibrium points in extensive games. *International journal of game theory*, 4(1):25–55, 1975.

[33] J. Watson. Perfect bayesian equilibrium: General definitions and illustrations. 2015.

[34] E. Zermelo. Über eine Anwendung der Mengenlehre auf die Theorie des Schachspiels. In *Proceedings of the fifth international congress of mathematicians*, volume 2, pages 501–504. II, Cambridge UP, Cambridge, 1913.