NON-COOPERATIVE GAMES

MIHAI MANEA

Department of Economics, MIT,

1. Normal-Form Games

A **normal** (or **strategic**) **form game** is a triplet (N, S, U) with the following properties

- $N = \{1, 2, \ldots, n\}$ is a finite set of **players**
- S_i is the set of **pure strategies** of player i; $S = S_1 \times \ldots \times S_n$
- $u_i : S \to \mathbb{R}$ is the **payoff function** of player i; $u = (u_1, \ldots, u_n)$.

Player $i \in N$ receives payoff $u_i(s)$ when $s \in S$ is played. The game is **finite** if S is finite.

The structure of the game is **common knoweldge**: all players know (N, S, U), and know that their opponents know it, and know that their opponents know that they know, and so on.

1.1. **Detour on common knowledge.**

Common knowledge looks like an innocuous assumption, but may have strong consequences in some situations. Consider the following story. Once upon a time, there was a village with 100 married couples. The women had to pass a logic exam before being allowed to marry. The high priestess was not required to take that exam, but it was common knowledge that she was truthful. The village was small, so everyone would be able to hear any shot fired in the village. The women would gossip about adulterous relationships and each knows which of the other husbands are unfaithful. However, no one would ever inform a wife about her cheating husband.

The high priestess knows that some husbands are unfaithful and decides that such immorality should not be tolerated any further. This is a successful religion and all women agree with the views of the priestess.

The priestess convenes all the women at the temple and *publicly announces* that the well-being of the village has been compromised—there is at least one cheating husband. She

---

Date: May 13, 2010.

also points out that even though none of them knows whether her husband is faithful, each woman knows about the other unfaithful husbands. She orders that each woman shoot her husband on the midnight of the day she finds out. 39 silent nights went by and on the 40^{th} shots were heard. How many husbands were shot? Were all the unfaithful husbands caught? How did some wives learn of their husband's infidelity after 39 nights in which *nothing* happened?

Since the priestess was truthful, there must have been at least one unfaithful husband in the village. How would events have evolved if there was exactly one unfaithful husband? His wife, upon hearing the priestess' statement and realizing that she does not know of any unfaithful husband, would have concluded that her own marriage must be the only adulterous one and would have shot her husband on the midnight of the first day. Clearly, there must have been more than one unfaithful husband. If there had been exactly two unfaithful husbands, then each of the two cheated wives would have initially known of exactly one unfaithful husband, and after the first silent night would infer that there were exactly two cheaters and her husband is one of them. (Recall that the wives are all perfect logicians.) The unfaithful husbands would thus both be shot on the second night. As no shots were heard on the first two nights, all women concluded that there were at least three cheating husbands... Since shootings were heard on the 40^{th} night, it must be that exactly 40 husbands were unfaithful and they were all exposed and killed simultaneously.

For any measurable space X we denote by $\Delta(X)$ the set of probability measures (or distributions) on X.[1] A **mixed strategy** for player i is an element σ_i of $\Delta(S_i)$. A **correlated strategy profile** σ is an element of $\Delta(S)$. A strategy profile σ is **independent** (or **mixed**) if $\sigma \in \Delta(S_1) \times \ldots \times \Delta(S_n)$, in which case we write $\sigma = (\sigma_1, \ldots, \sigma_n)$ where $\sigma_i \in \Delta(S_i)$ denotes the marginal of σ on S_i. A **correlated belief** for player i is an element σ_{-i} of $\Delta(S_{-i})$. The set of **independent beliefs** for i is $\prod_{j \neq i} \Delta(S_j)$. It is assumed that player i has von Neumann-Morgenstern preferences over $\Delta(S)$ and u_i extends to $\Delta(S)$ as follows

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

[1]In most of our applications X is either finite or a subset of an Euclidean space.

2. Dominated Strategies

Are there obvious predictions about how a game should be played?

Example 1 (Prisoners' Dilemma). *Two persons are arrested for a crime, but there is not enough evidence to convict either of them. Police would like the accused to testify against each other. The prisoners are put in different cells, with no communication possibility. Each suspect is told that if he testifies against the other ("Defect"), he is released and given a reward provided the other does not testify ("Cooperate"). If neither testifies, both are released (with no reward). If both testify, then both go to prison, but still collect rewards for testifying. Each*

	C	D
C	$1, 1$	$-1, 2$
D	$2, -1$	$0, 0^*$

prisoner is better off defecting regardless of what the other does. Cooperation is a strictly dominated action for each prisoner. The only feasible outcome is (D, D), which is Pareto dominated by (C, C).

Example 2. *Consider the game obtained from the prisoners' dilemma by changing player 1's payoff for (C, D) from -1 to 1. No matter what player 1 does, player 2 still prefers*

	C	D
C	$1, 1$	$1, 2^*$
D	$2, -1$	$0, 0$

D to C. If player 1 knows that 2 never plays C, then he prefers C to D. Unlike in the prisoners' dilemma example, we use an additional assumption to reach our prediction in this case: player 1 needs to deduce that player 2 never plays a dominated strategy.

Formally, a strategy $s_i \in S_i$ is **strictly dominated** by $\sigma_i \in \Delta(S_i)$ if

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

We can iteratively eliminate dominated strategies, under the assumption that "I know that you know that I know... that I know the payoffs and that you would never use a dominated strategy."[2]

Definition 1. *For all $i \in N$, set $S_i^0 = S_i$ and define S_i^k recursively by*

$$S_i^k = \{s_i \in S_i^{k-1} | \nexists \sigma_i \in \Delta(S_i^{k-1}), u_i(\sigma_i, s_{-i}) > u_i(s_i, s_{-i}), \forall s_{-i} \in S_{-i}^{k-1}\}.$$

The set of pure strategies of player i that survive ***iterated deletion of strictly dominated strategies*** *is $S_i^\infty = \cap_{k \geq 0} S_i^k$. The set of surviving mixed strategies is*

$$\{\sigma_i \in \Delta(S_i^\infty) | \nexists \sigma_i' \in \Delta(S_i^\infty), u_i(\sigma_i', s_{-i}) > u_i(\sigma_i, s_{-i}), \forall s_{-i} \in S_{-i}^\infty\}.$$

Note that in a finite game the elimination procedure ends in a finite number of steps, so S^∞ is simply the set of surviving strategies at the last stage.

The definition above assumes that at each iteration all dominated strategies of each player are deleted simultaneously. Clearly, there are many other iterative procedures that can be used to eliminate strictly dominated strategies. However, the limit set S^∞ does not depend on the particular way deletion proceeds.[3] The intuition is that a strategy which is dominated at some stage is dominated at any later stage. Furthermore, the outcome does not change if we eliminate strictly dominated mixed strategies at every step. The reason is that a strategy is dominated against all pure strategies of the opponents if and only if it is dominated against all their mixed strategies. Eliminating mixed strategies for player i at any stage does not affect the set of strictly dominated pure strategies for any player $j \neq i$ at the next stage.

3. Rationalizability

Rationalizability is a solution concept introduced independently by Bernheim (1984) and Pearce (1984). Like iterated strict dominance, rationalizability derives restrictions on play from the assumptions that the payoffs and rationality of the players are common knowledge. Dominance: what actions should a player never use? Rationalizability: what strategies can

[2]Play the "guess 2/3 of the average" game with strategy space $\{1, 2, \ldots, 100\}$. Talk about the Keynesian beauty contest. Note that playing 1 is a winning strategy only if there is common knowledge of rationality, which is not usually true in class.

[3]This property does not hold for *weakly* dominated strategies.

a rational player choose? It is not rational for a player to choose a strategy that is not a best response to some beliefs about his opponents' strategies.

What is a "belief"? In Bernheim (1984) and Pearce (1984) each player i's beliefs σ_{-i} about the play of $j \neq i$ must be independent, i.e., $\sigma_{-i} \in \prod_{j \neq i} \Delta(S_j)$. Independent beliefs are consistent with the definition of mixed strategies, but in the context of an iterative procedure entail common knowledge of the fact that each player holds such beliefs. Alternatively, we may allow player i to believe that the actions of his opponents are correlated, i.e., any $\sigma_{-i} \in \Delta(S_{-i})$ is a possibility. The two definitions have different implications for $n \geq 3$. We focus on the case with correlated beliefs.

We can again iteratively develop restrictions imposed by common knowledge of the payoffs and rationality to obtain the definition of rationalizability.

Definition 2. *Set $S^0 = S$ and let S^k be given recursively by*

$$S_i^k = \{s_i \in S_i^{k-1} | \exists \sigma_{-i} \in \Delta(S_{-i}^{k-1}), u_i(s_i, \sigma_{-i}) \geq u_i(s_i', \sigma_{-i}), \forall s_i' \in S_i^{k-1}\}.$$

The set of ***correlated rationalizable strategies*** *for player i is $S_i^\infty = \cap_{k \geq 0} S_i^k$. A mixed strategy $\sigma_i \in \Delta(S_i)$ is rationalizable if there is a belief $\sigma_{-i} \in \Delta(S_{-i}^\infty)$ s.t. $u_i(\sigma_i, \sigma_{-i}) \geq u_i(s_i, \sigma_{-i})$ for all $s_i \in S_i^\infty$.*

The definition of **independent rationalizability** replaces $\Delta(S_{-i}^{k-1})$ and $\Delta(S_{-i}^\infty)$ above with $\prod_{j \neq i} \Delta(S_j^{k-1})$ and $\prod_{j \neq i} \Delta(S_j^\infty)$, respectively.

Definition 3. *A strategy $s_i^* \in S_i$ is a* ***best response*** *to a belief $\sigma_{-i} \in \Delta(S_{-i})$ if*

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

We say that a strategy s_i is **never a best response** for player i if it is not a best response to any $\sigma_{-i} \in \Delta(S_{-i})$. Recall that a strategy s_i of player i is **strictly dominated** if there exists $\sigma_i \in \Delta(S_i)$ s.t. $u_i(\sigma_i, s_{-i}) > u_i(s_i, s_{-i}), \forall s_{-i} \in S_{-i}$.

Theorem 1. *In a finite game, a strategy is never a best response if and only if it is strictly dominated.*

Proof. Clearly, a strategy s_i strictly dominated for player i by some σ_i cannot be a best response for any belief $\sigma_{-i} \in \Delta(S_{-i})$ as σ_i yields a strictly higher payoff than s_i against any such σ_{-i}.

We are left to show that a strategy which is never a best response must be strictly dominated. We prove that any strategy s_i of player i which is not strictly dominated must be a best response for some beliefs. Define the set of "dominated payoffs" for i by

$$D = \{x \in \mathbb{R}^{|S_{-i}|} | \exists \sigma_i \in \Delta(S_i), x \leq u_i(\sigma_i, \cdot)\}.$$

Clearly D is non-empty, closed and convex. Also, $u_i(s_i, \cdot)$ does not belong to the interior of D because it is not strictly dominated by any $\sigma_i \in \Delta(S_i)$. By the supporting hyperplane theorem, there exists $\alpha \in \mathbb{R}^{|S_{-i}|}$ different from the zero vector s.t. $\alpha \cdot u_i(s_i, \cdot) \geq \alpha \cdot x, \forall x \in D$. In particular, $\alpha \cdot u_i(s_i, \cdot) \geq \alpha \cdot u_i(\sigma_i, \cdot), \forall \sigma_i \in \Delta(S_i)$. Since D is not bounded from below, each component of α needs to be non-negative. We can normalize α so that its components sum to 1, in which case it can be interpreted as a belief in $\Delta(S_{-i})$ with the property that $u_i(s_i, \alpha) \geq u_i(\sigma_i, \alpha), \forall \sigma_i \in \Delta(S_i)$. Thus s_i is a best response to α. □

Corollary 1. *Correlated rationalizability and iterated strict dominance coincide.*

Theorem 2. *For every $k \geq 0$, each $s_i \in S_i^k$ is a best response (within S_i) to a belief in $\Delta(S_{-i}^k)$.*

Proof. Fix $s_i \in S_i^k$. We know that s_i is a best response within S_i^{k-1} to some $\sigma_{-i} \in \Delta(S_{-i}^{k-1})$. If s_i was not a best response within S_i to σ_{-i}, let s_i' be such a best response. Since s_i is a best response within S_i^{k-1} to σ_{-i}, and s_i' is a strictly better response than s_i to σ_{-i}, we need $s_i' \notin S_i^{k-1}$. This contradicts the fact that s_i' is a best response against σ_{-i}, which belongs to $\Delta(S_{-i}^{k-1})$. □

Corollary 2. *Each $s_i \in S_i^\infty$ is a best response (within S_i) to a belief in $\Delta(S_{-i}^\infty)$.*

Theorem 3. *S^∞ is the largest set $Z_1 \times \ldots \times Z_n$ with $Z_i \subset S_i, \forall i \in N$ s.t. each element in Z_i is a best response to a belief in $\Delta(Z_{-i})$ for all i.*

Proof. Clearly S^∞ has the stated property by Theorem 2. Suppose that there exists $Z_1 \times \ldots \times Z_n \not\subset S^\infty$ satisfying the property. Consider the smallest k for which there is an i such that $Z_i \not\subset S_i^k$. It must be that $k \geq 1$ and $Z_{-i} \subset S_{-i}^{k-1}$. By assumption, every element in Z_i is a best response to an element of $\Delta(Z_{-i}) \subset \Delta(S_{-i}^{k-1})$, contradicting $Z_i \not\subset S_i^k$. □

Example 3 (Rationalizability in Cournot duopoly). *Two firms compete on the market for a divisible homogeneous good. Each firm $i = 1, 2$ has zero marginal cost and simultaneously decides to produce an amount of output $q_i \geq 0$. The resulting price is $p = \max(0, 1 - q_1 - q_2)$. Hence, if $q_1 + q_2 \leq 1$, the profit of firm i is given by $q_i(1 - q_1 - q_2)$. The best response correspondence of firm i is $B_i(q_j) = (1 - q_j)/2$ ($j = 3 - i$). If i knows that $q_j \lesseqqgtr q$ then $B_i(q_j) \gtreqqless (1 - q)/2$.*

We know that $q_i \geq q^0 = 0$ for $i = 1, 2$. Hence $q_i \leq q^1 = B_i(q^0) = (1 - q^0)/2$ for all i. But then $q_i \geq q^2 = B_i(q^1) = (1 - q^1)/2$ for all i... We obtain

$$\forall i,\ q^0 \leq q^2 \leq \ldots \leq q^{2k} \leq \ldots \leq q_i \leq \ldots \leq q^{2k+1} \leq \ldots \leq q^1,$$

where $q^{2k} = \sum_{l=1}^{k} 1/4^l = (1 - 1/4^k)/3$ and $q^{2k+1} = (1 - q^{2k})/2$ for all $k \geq 0$. Clearly, $\lim_{k\to\infty} q^k = 1/3$, hence the only rationalizable strategy for firm i is $q_i = 1/3$. This is also the unique Nash equilibrium, which we define next.

4. Nash Equilibrium

Many games are not solvable by iterated strict dominance or rationalizability. The concept of Nash (1950) equilibrium has more bite in some situations. The idea of Nash equilibrium was implicit in the particular examples of Cournot (1838) and Bertrand (1883) at an informal level.

Definition 4. *A mixed-strategy profile σ^* is a Nash equilibrium if for each $i \in N$*

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

Note that if a player uses a nondegenerate mixed strategy in a Nash equilibrium (one that places positive probability weight on more than one pure strategy) then he must be indifferent between all pure strategies in the support. Of course, the fact that there is no profitable deviation in pure strategies implies that there is no profitable deviation in mixed strategies either.

Example 4 (Matching Pennies). *Pure strategy equilibria do not always exist. We will establish that equilibria in mixed strategies always exist.*

	H	T
H	$1,-1$	$-1,1$
T	$-1,1$	$1,-1$

Nash equilibria are "consistent" predictions of how the game will be played—if all players predict that a specific Nash equilibrium will arise then no player has incentives to play differently. Each player must have correct "conjectures" about the strategies of their opponents and play a best response to his conjecture. Formally, Aumann and Brandenburger (1995) provide a framework that can be used to examine the epistemic foundations of Nash equilibrium. The primitive of their model is an **interactive belief system** in which each player has a possible set of types, which correspond to beliefs about the types of the other players, a payoff for each action, and an action selection. Aumann and Brandenburger show that in a 2-player game, if the game being played (i.e., both payoff functions), the rationality of the players, and their conjectures are all *mutually known*, then the conjectures constitute a (mixed strategy) Nash equilibrium. Thus *common knowledge* plays no role in the 2-player case. However, for games with more than 2 players, we need to assume additionally that players have a common prior and and that conjectures are *commonly known*.

So far, we have motivated our solution concepts by presuming that players make predictions about their opponents' play by introspection and deduction, using knowledge of their opponents payoffs, knowledge that the opponents are rational, knowledge about this knowledge... Alternatively, we may assume that players extrapolate from past observations of play in "similar" games, with either current opponents or "similar" ones. They form expectations about future play based on past observations and adjust their actions to maximize their current payoffs with respect to these expectations. The idea of using adjustment processes to model **learning** originates with Cournot (1838). In that setting (Example 3), players take turns setting their outputs, each player choosing a best response to the opponent's last period action. Alternatively, we can assume simultaneous belief updating, best responding to sample average play, populations of players being anonymously matched, etc. If the process *converges* to a particular steady state, then the steady state is a Nash equilibrium. While convergence occurs in Example 3, this is not always the case. How sensitive is the convergence to the initial state? If convergence obtains for all initial strategy profiles sufficiently close to the steady state, we say that the steady state is asymptotically stable. See figures 1.13-1.15 (pp. 24-26) in FT. The Shapley (1964) cycling example is interesting. Also, adjustment processes are myopic and do not offer a compelling description of behavior.

	L	M	R
U	$0,0$	$0,1$	$1,0$
M	$1,0$	$0,0,$	$0,1$
D	$0,1$	$1,0$	$0,0$

Definitely such processes do not provide good predictions for behavior in the actual repeated game.

5. Existence and Continuity of Nash Equilibria

Follow Muhamet's slides. We need the following result for future reference.

Theorem 4. *Suppose that each S_i is a convex and compact subset of an Euclidean space and that each u_i is continuous in s and quasi-concave in s_i. Then there exists a* ***pure strategy*** *Nash equilibrium.*

6. Bayesian Games

When some players are uncertain about the characteristics or **types** of others, the game is said to have **incomplete information**. Most often a player's type is simply defined by his payoff function. More generally, types may embody any private information that is relevant to players' decision making. This may include, in addition to the player's payoff function, his beliefs about other players' payoff functions, his beliefs about what other players believe his beliefs are, and so on. The idea that a situation in which players are unsure about each other's payoffs and beliefs can be modeled as a Bayesian game, in which a player's type encapsulates all his uncertainty, is due to Harsanyi (1967, 1968) and has been formalized by Mertens and Zamir (1985). For simplicity, we assume that a player's type is his own payoff and the type captures all the private information.

A **Bayesian game** is a list $\mathcal{B} = (N, S, \Theta, u, p)$ with

- $N = \{1, 2, \ldots, n\}$ is a finite set of **players**
- S_i is the set of **pure strategies** of player i; $S = S_1 \times \ldots \times S_n$

- Θ_i is the set of **types** of player i; $\Theta = \Theta_1 \times \ldots \times \Theta_n$
- $u_i : \Theta \times S \to \mathbb{R}$ is the **payoff function** of player i; $u = (u_1, \ldots, u_n)$
- $p \in \Delta(\Theta)$ is a common prior (we can relax this assumption).

We often assume that Θ is finite and the marginal $p_i(\theta_i)$ is positive for each type θ_i.

Example 5 (First Price Auction with I.I.D. Private Values). *One object is up for sale. Suppose that the value θ_i of player $i \in N$ for the object is uniformly distributed in $\Theta_i = [0,1]$ and that the values are independent across players. This means that if $\tilde{\theta}_i \in [0,1], \forall i$ then $p(\theta_i \leq \tilde{\theta}_i, \forall i) = \prod_i \tilde{\theta}_i$. Each player i submits a bid $s_i \in S_i = [0, \infty)$. The player with the highest bid wins the object and pays his bid. Ties are broken randomly. Hence the payoffs are given by*

$$u_i(\theta, s) = \begin{cases} \frac{\theta_i - s_i}{|\{j \in N | s_i = s_j\}|} & \text{if } s_i \geq s_j, \forall j \in N \\ 0 & \text{otherwise.} \end{cases}$$

Example 6 (An exchange game). *Each player $i = 1, 2$ receives a ticket on which there is a number in some finite set $\Theta_i \subset [0,1]$. The number on a player's ticket represents the size of a prize he may receive. The two prizes are independently distributed, with the value on i's ticket distributed according to F_i. Each player is asked independently and simultaneously whether he wants to exchange his prize for the other player's prize, hence $S_i = \{agree,\ disagree\}$. If both players agree then the prizes are exchanged; otherwise each player receives his own prize. Thus the payoff of player i is*

$$u_i(\theta, s) = \begin{cases} \theta_{3-i} & \text{if } s_1 = s_2 = agree \\ \theta_i & \text{otherwise.} \end{cases}$$

In the **normal form representation** $G(\mathcal{B})$ **of the Bayesian game** $\mathcal{B}$ player i has strategies $(s_i(\theta_i))_{\theta_i \in \Theta_i} \in S_i^{\Theta_i}$ and utility function U_i given by

$$U_i((s_i(\theta_i))_{\theta_i \in \Theta_i, i \in N}) = \mathbb{E}_p(u_i(\theta, s_1(\theta_1), \ldots, s_n(\theta_n))).$$

The **agent-normal form representation** $AG(\mathcal{B})$ **of the Bayesian game** $\mathcal{B}$ has player set $\cup_i \Theta_i$. The strategy space of each player θ_i is S_i. A strategy profile $(s_{\theta_i})_{\theta_i \in \Theta_i, i \in N}$ yields utility

$$U_{\theta_i}((s_{\theta_i})_{\theta_i \in \Theta_i, i \in N}) = \mathbb{E}_p(u_i(\theta, s_{\theta_1}, \ldots, s_{\theta_n})|\theta_i)$$

for player θ_i. For the conditional expectation to be well-defined we need $p_i(\theta_i) > 0$.

Definition 5. *A Bayesian Nash equilibrium of $\mathcal{B}$ is a Nash equilibrium of $G(\mathcal{B})$.*

Proposition 1. *If $p_i(\theta_i) > 0$ for all $\theta_i \in \Theta_i, i \in N$, a strategy profile is a Nash equilibrium of $G(\mathcal{B})$ iff it is a Nash equilibrium of $AG(\mathcal{B})$ (strategies are mapped across the two games by $s_i(\theta_i) \to s_{\theta_i}$).*

Theorem 5. *Suppose that*

- *N and Θ are finite*
- *each S_i is a compact and convex subset of an Euclidean space*
- *each u_i is continuous in s and concave in s_i.*

Then $\mathcal{B}$ has a pure strategy Bayesian Nash equilibrium.

Proof. By Proposition 1, it is sufficient to show that $AG(\mathcal{B})$ has a pure Nash equilibrium. The latter follows from Theorem 4. We use the concavity of u_i in s_i to show that the corresponding U_{θ_i} is quasi-concave in s_{θ_i}. Quasi-concavity of u_i in s_i does not typically imply quasi-concavity of U_{θ_i} in s_{θ_i} because U_{θ_i} is an integral of u_i over variables other than s_{θ_i}.[4] □

We can show that the set of Bayesian Nash equilibria of $\mathcal{B}^x$ is upper-hemicontinuous with respect to x when payoffs are given by u^x, assumed continuous in x in a compact set X, if S, Θ are finite. Indeed, $BNE(\mathcal{B}^x) = NE(AG(\mathcal{B}^x))$. Furthermore, we have upper-hemicontinuity with respect to beliefs.

Theorem 6. *Suppose that N, S, Θ are finite. Let $P \subset \Delta(\Theta)$ be such that for every $p \in P$ $p_i(\theta_i) > 0, \forall \theta_i \in \Theta_i, i \in N$. Then $BNE(\mathcal{B}^p)$ is upper-hemicontinuous in p over P.*

Proof. Since $BNE(\mathcal{B}^p) = NE(G(\mathcal{B}^p))$, it is sufficient to note that

$$U_i((s_i(\theta_i))_{\theta_i \in \Theta_i, i \in N}) = \mathbb{E}_p(u_i(\theta, s_1(\theta_1), \ldots, s_n(\theta_n)))$$

(as defined for $G(\mathcal{B}^p)$) is continuous in p. □

[4]Sums of quasi-concave functions are not necessarily quasi-concave.

7. Extensive Form Games

An **extensive form game** consists of

- a finite set of **players** $N = \{1, 2, \ldots, n\}$; nature is denoted as "player 0"
- the order of moves specified by a **tree**
- each player's payoffs at the terminal nodes in the tree
- information partition
- the set of actions available at every information set and a description of how actions lead to progress in the tree
- moves by nature.

A **tree** is a **directed graph** $(X, >)$—there is a link from node x to node y if $x > y$, which we interpret as "x precedes y." We assume that X is **finite**, there is an initial node $\phi \in X$, $>$ is transitive ($x > y, y > z \Rightarrow x > z$) and asymmetric ($x > y \Rightarrow y \not> x$). Hence the tree has no cycles. We also require that each node $x \neq \phi$ has exactly one immediate predecessor, i.e., $\exists x' > x$ such that $x'' > x, x'' \neq x'$ implies $x'' > x'$. A node is **terminal** if it does not precede any other node; this means that the set of terminal nodes is $Z = \{z| \not\exists x, z > x\}$. Each $z \in Z$ completely determines a path of moves though the tree (recall the finiteness assumption), with associated payoff $u_i(z)$ for player i.

An **information partition** is a partition of $X \setminus Z$. Node x belongs to the information set $h(x)$. The same player, denoted $i(h) \in N \cup \{0\}$, moves at each node $x \in h$ (otherwise players might disagree on whose turn to move is). The interpretation is that $i(h)$ is uncertain whether he is at x or some other $x' \in h(x)$. We abuse notation writing $i(x) = i(h(x))$.

The set of available **actions** at $x \in X \setminus Z$ for player $i(x)$ is $A(x)$. We assume that $A(x) = A(x') =: A(h), \forall x' \in h(x)$ (otherwise $i(h)$ might play an infeasible action). A function l labels each node $x \neq \phi$ with the last action taken to reach it. We require that the restriction of l to the immediate successors of x be bijective on $A(x)$. Finally, a move by nature at some node x corresponds to a probability distribution over $A(x)$.

Let $H_i = \{h|i(h) = i\}$. The set of **pure strategies** for player i is $S_i = \prod_{h \in H_i} A(h)$. As usual, $S = \prod_{i \in N} S_i$. A strategy is a complete contingent plan specifying an action to be taken at each information set (if reached). We can define **mixed strategies** as probability distributions over pure strategies, $\sigma_i \in \Delta(S_i)$. Any mixed strategy profile $\sigma \in \prod_{i \in N} \Delta(S_i)$, along with the distribution of the moves by nature and the labeling of nodes with actions, leads to a probability distribution $O(\sigma) \in \Delta(Z)$. We denote by $u_i(\sigma) = \mathbb{E}_{O(\sigma)}(u_i(z))$. The associated **normal form** game is (N, S, u).

Two strategies $s_i, s_i' \in S_i$ are equivalent if $O(s_i, s_{-i}) = O(s_i', s_{-i}), \forall s_{-i}$, that is, they lead to the same distribution over outcomes regardless of how the opponents play. See figure 3.9 in FT p. 86. S_i^R is a subset of S_i that contains exactly one strategy from each equivalence class. The **reduced normal form** game is given by (N, S^R, u).

A **behavior strategy** specifies a distribution over actions for each information set. Formally, a behavior strategy $b_i(h)$ for player $i(h)$ at information set h is an element of $\Delta(A(h))$. We use the notation $b_i(a|h)$ for the probability of action a at information set h. A behavior strategy b_i for i is an element of $\prod_{h \in H_i} \Delta(A(h))$. A profile b of behavior strategies determines a distribution over Z in the obvious way. Clearly, b_i is equivalent to σ_i with

$$\sigma_i(s_i) = \prod_{h \in H_i} b_i(s_i(h)|h),$$

where $s_i(h)$ denotes the projection of s_i on $A(h)$.

To guarantee that every mixed strategy is equivalent to a behavior strategy we need to impose the additional requirement of **perfect recall**. Basically, prefect recall means that no player ever forgets any information he once had and all players know the actions they have chosen previously. See figure 3.5 in FT, p. 81. Formally, perfect recall stipulates that if $x'' \in h(x')$, x is a predecessor of x' and the same player i moves at both x and x' (and thus at x'') then there is a node $\hat{x}$ in the same information set as x (possibly x itself) such that $\hat{x}$ is a predecessor of x'' and the action taken at x along the path to x' is the same as the action taken at $\hat{x}$ along the path to x''. Intuitively, the nodes x' and x'' are distinguished by information i does not have, so he cannot have had it at $h(x)$; x' and x'' must be consistent with the same action at $h(x)$ since i must remember his action there.

Let $R_i(h)$ be the set of pure strategies for player i that do not preclude reaching the information set $h \in H_i$, i.e., $R_i(h) = \{s_i | h \text{ is on the path of some } (s_i, s_{-i})\}$. If the game has prefect recall, a mixed strategy σ_i is equivalent to a behavior strategy b_i defined by

$$b_i(a|h) = \frac{\sum_{\{s_i \in R_i(h) | s_i(h) = a\}} \sigma_i(s_i)}{\sum_{s_i \in R_i(h)} \sigma_i(s_i)},$$

when the denominator is positive and any distribution when it is zero.

Many different mixed strategies can generate the same behavior strategy. Consider the example from FT p. 88, figure 3.12. Player 2 has four pure strategies, $s_2 = (A, C), s_2' = (A, D), s_2'' = (B, C), s_2''' = (B, D)$. Now consider two mixed strategies, $\sigma_2 = (1/4, 1/4, 1/4, 1/4)$, which assigns probability 1/4 to each pure strategy, and $\sigma_2 = (1/2, 0, 0, 1/2)$, which assigns probability 1/2 to each of s_2 and s_2'''. Both of these mixed strategies generate the behavior strategy b_2 with $b_2(A|h) = b_2(B|h) = 1/2$ and $b_2(C|h') = b_2(D|h') = 1/2$. Moreover, for any strategy σ_1 of player 1, all of σ_2, σ_2', b_2 lead to the same probability distribution over terminal nodes. For example, the probability of reaching node z_1 equals the probability of player 1 playing U times 1/2.

The relationship between mixed and behavior strategies is different in the game illustrated in FT p. 89, figure 3.13, which is not a game of perfect recall (player 1 forgets what he did at the initial node; formally, there are two nodes in his second information set which obviously succeed the initial node, but are not reached by the same action at the initial node). Player 1 has four strategies in the strategic form, $s_1 = (A, C), s_1' = (A, D), s_1'' = (B, C), s_1''' = (B, D)$. Now consider the mixed strategy $\sigma_1 = (1/2, 0, 0, 1/2)$. As in the last example, this generates the behavior strategy $b_1 = \{(1/2, 1/2), (1/2, 1/2)\}$, where player 1 mixes $50 - 50$ at each information set. But b_1 is *not* equivalent to the σ_1 that generated it. Indeed (σ_1, L) generates a probability 1/2 for the terminal node corresponding to (A, L, C) and a 1/2 probability for (B, L, D). However, since behavior strategies describe independent randomizations at each information set, (b_1, L) assigns probability 1/4 to each of the four paths $(A, L, C), (A, L, D), (B, L, C), (B, L, D)$. Since both A vs. B and C vs. D are choices made by player 1, the strategy σ_1 under which player 1 makes all his decisions at once allows choices at different information sets to be correlated. Behavior strategies cannot produce this correlation in the example, because when it comes time to choose between C and D, player 1 has forgotten whether he chose A or B.

Theorem 7 (Kuhn 1953)**.** *Under perfect recall, mixed and behavioral strategies are equivalent.*

Hereafter we restrict attention to games with perfect recall, and use mixed and behavior strategies interchangeably. Behavior strategies prove more convenient in many arguments and constructions. We drop the notation b for behavior strategies and instead use $\sigma_i(a_i|h)$ to denote player i's probability of playing action a_i at information set h...

8. Backward Induction and Subgame Perfection

An extensive form game has **perfect information** if all information sets are singletons. Backward induction can be applied to any extensive form game of perfect information with finite X (which means that the number of "stages" and the number of actions feasible at any stage are finite). The idea of backward induction is formalized by Zermelo's algorithm. Since the game is finite, it has a set of penultimate nodes, i.e., nodes whose (all) immediate successors are terminal nodes. Specify that the player who moves at each such node chooses the strategy leading to the terminal node with the highest payoff for him. In case of a tie, make an arbitrary selection. Next each player at nodes whose immediate successors are penultimate (or terminal) nodes chooses the action maximizing his payoff over the feasible successors, given that players at the penultimate nodes play as assumed. We can now roll back through the tree, specifying actions at each node. At the end of the process we have a pure strategy for each player. It is easy to check that the resulting strategies form a Nash equilibrium.

Theorem 8 (Zermelo 1913; Kuhn 1953)**.** *A finite game of perfect information has a pure-strategy Nash equilibrium.*

Moreover, the backward induction solution has the nice property that each player's actions are optimal at every possible history if the play of the opponents is held fixed, which we call subgame perfection. More generally, subgame perfection extends the logic of backward induction to games with imperfect information. The idea is to replace the "smallest" proper subgame with one of its Nash equilibria and iterate this procedure on the reduced tree. In stages where the "smallest" subgame has multiple Nash equilibria, the procedure implicitly assumes that all players believe the same equilibrium will be played. To define subgame perfection formally we first need the definition of a proper subgame. Informally, a proper subgame is a portion of a game that can be analyzed as a game in its own right.

Definition 6. *A* ***proper subgame*** *G of an extensive form game T consists of a single node x and all its successors in T, with the property that if $x' \in G$ and $x'' \in h(x')$ then*

$x'' \in G$. The information sets and payoffs of the subgame are inherited from the original game. That is, two nodes are in the same information set in G if and only if they are in the same information set in T, and the payoff function on the subgame is just the restriction of the original payoff function to the terminal nodes of G.

The requirements that all the successors of x be in the subgame and that the subgame does not "chop up" any information set ensure that the subgame corresponds to a situation that could arise in the original game. In figure 3.16, p. 95 of FT, the game on the right is not a subgame of the game on the left, because on the right player 2 knows that player 1 has not played L, which he did not know in the original game.

Together, the requirements that the subgame begin with a single node x and respect information sets imply that in the original game x must be a singleton information set, i.e. $h(x) = \{x\}$. This ensures that the payoffs in the subgame, conditional on the subgame being reached, are well defined. In figure 3.17, p. 95 of FT, the "game" on the right has the problem that player 2's optimal choice depends on the relative probabilities of nodes x and x', but the specification of the game does not provide these probabilities. In other words, the diagram on the right cannot be analyzed as a separate game; it makes sense only as a component of the game on the left, which provides the missing probabilities.

Since payoffs conditional on reaching a proper subgame are well-defined, we can test whether strategies yield a Nash equilibrium when restricted to the subgame.

Definition 7. *A behavior strategy profile σ of an extensive form game is a* ***subgame perfect equilibrium*** *if the restriction of σ to G is a Nash equilibrium of G for every proper subgame G.*

Because any game is a proper subgame of itself, a subgame perfect equilibrium profile is necessarily a Nash equilibrium. If the only proper subgame is the whole game, the sets of Nash and subgame perfect equilibria coincide. If there are other proper subgames, some Nash equilibria may fail to be subgame perfect.

It is easy to see that subgame perfection coincides with backward induction in finite games of perfect information. Consider the penultimate nodes of the tree, where the last choices are made. Each of these nodes begins a trivial one-player proper subgame, and Nash equilibrium in these subgames requires that the player make a choice that maximizes his payoff. Thus any subgame perfect equilibrium must coincide with a backward induction solution at every penultimate node, and we can continue up the tree by induction.

9. Important Examples of Extensive Form Games

9.1. Repeated games with observable actions.

- **time** $t = 0, 1, 2, \ldots$ (usually infinite)
- **stage game** is a normal-form game G
- G is played every period t
- players **observe** the realized actions at the end of each period
- **payoffs** are the sum of discounted payoffs in the stage game.

Repeated games are a widely studied class of dynamic games. There is a lot of research dealing with various restrictions on the information about past play.

9.2. Multi-stage games with observable actions.

- **stages** $k = 0, 1, 2, \ldots$
- at stage k, after having **observed** a "non-terminal" history of play $h = (a^0, \ldots, a^{k-1})$, each player i simultaneously chooses an **action** $a_i^k \in A_i(h)$
- **payoffs** given by $u(h)$ for each terminal history h.

Often it is natural to identify the "stages" of the game with time periods, but this is not always the case. A game of perfect information can be viewed as a multi-stage game in which every stage corresponds to some node. At every stage all but one player (the one moving at the node corresponding to that stage) have singleton action sets ("do nothing"; can refer to these players as "inactive"). Repeated versions of perfect information extensive form games also lead to multi-stage games, e.g., the Rubinstein (1982) alternating bargaining game, which we discuss later.

10. Single (or One-Shot) Deviation Principle

Consider a multi-stage game with observed actions. We show that in order to verify that a strategy profile σ is subgame perfect, it suffices to check whether there are any histories h_t where some player i can gain by deviating from the actions prescribed by σ_i at h_t and conforming to σ_i elsewhere.

If σ is a strategy profile and h_t a strategy, write $u_i(\sigma|h_t)$ for the (expected) payoff to player i that results if play starts at h_t and continues according to σ in each stage.

Definition 8. *A strategy σ_i is **unimprovable** given σ_{-i} if $u_i(\sigma_i, \sigma_{-i}|\, h_t) \geq u_i(\sigma_i', \sigma_{-i}|\, h_t)$ for every $t \geq 0, h_t \in H_i$ and $\sigma_i' \in \Delta(S_i)$ with $\sigma_i'(h_{t'}') = \sigma_i(h_{t'}')$ for all $h_{t'}' \in H_i \setminus \{h_t\}$.*

Hence a strategy σ_i is unimprovable if after every history, no strategy that differs from it at only one information set can increase utility. Obviously, if σ is a subgame perfect equilibrium then σ_i is unimprovable given σ_{-i}. To establish the converse, we need an additional condition.

Definition 9. *A game is **continuous at infinity** if for each player i the utility function u_i satisfies*

$$\lim_{t \to \infty} \sup_{\{(h, \tilde{h}) | h_t = \tilde{h}_t\}} |u_i(h) - u_i(\tilde{h})| = 0.$$

Continuity at infinity requires that events in the distant future are relatively unimportant. It is satisfied if the overall payoffs are a discounted sum of per-period payoffs and the stage payoffs are uniformly bounded.

Theorem 9. *Consider a (finite or infinite horizon) multi-stage game with observed actions[5] that is continuous at infinity. If σ_i is unimprovable given σ_{-i} then σ_i is a best response to σ_{-i} conditional on any history h_t.*

Proof. Suppose that σ_i is unimprovable given σ_{-i}, but σ_i is not a best response to σ_{-i} following some history h_t. Let σ_i^1 be a strictly better response and define

$$\varepsilon = u_i(\sigma_i^1, \sigma_{-i}|h_t) - u_i(\sigma_i, \sigma_{-i}|h_t) > 0. \tag{10.1}$$

Since the game is *continuous at infinity*, there exists $t' > t$ and σ_i^2 such that σ_i^2 is identical to σ_i^1 at all information sets up to (and including) stage t', σ_i^2 coincides with σ_i across all longer histories and

$$|u_i(\sigma_i^2, \sigma_{-i}|h_t) - u_i(\sigma_i^1, \sigma_{-i}|h_t)| < \varepsilon/2. \tag{10.2}$$

In particular, 10.1 and 10.2 imply that

$$u_i(\sigma_i^2, \sigma_{-i}|h_t) > u_i(\sigma_i, \sigma_{-i}|h_t).$$

[5]We allow for the possibility that the action set be infinite at some stages.

Denote by σ_i^3 the strategy obtained from σ_i^2 by replacing the stage t' actions following any history $h_{t'}$ with the corresponding actions under σ_i. Conditional on any history $h_{t'}$, the strategies σ_i and σ_i^3 coincide, hence

$$u_i(\sigma_i^3, \sigma_{-i}|h_{t'}) = u_i(\sigma_i, \sigma_{-i}|h_{t'}). \tag{10.3}$$

As σ_i is *unimprovable* given σ_{-i}, and σ_i and σ_i^2 only differ at stage t' conditional on $h_{t'}$, we need

$$u_i(\sigma_i, \sigma_{-i}|h_{t'}) \geq u_i(\sigma_i^2, \sigma_{-i}|h_{t'}). \tag{10.4}$$

Then 10.3 and 10.4 lead to

$$u_i(\sigma_i^3, \sigma_{-i}|h_{t'}) \geq u_i(\sigma_i^2, \sigma_{-i}|h_{t'})$$

for all histories $h_{t'}$ (consistent with h_t). Since σ_i^2 and σ_i^3 coincide before reaching stage t', we obtain

$$u_i(\sigma_i^3, \sigma_{-i}|h_t) \geq u_i(\sigma_i^2, \sigma_{-i}|h_t).$$

Similarly, we can construct $\sigma_i^4, \ldots, \sigma_i^{t'-t+3}$. The strategy $\sigma_i^{t'-t+3}$ is identical to σ_i conditional on h_t and

$$u_i(\sigma_i, \sigma_{-i}|h_t) = u_i(\sigma_i^{t'-t+3}, \sigma_{-i}|h_t) \geq \ldots \geq u_i(\sigma_i^3, \sigma_{-i}|h_t) \geq u_i(\sigma_i^2, \sigma_{-i}|h_t) > u_i(\sigma_i, \sigma_{-i}|h_t),$$

a contradiction. □

11. Iterated Conditional Dominance

Definition 10. *In a multi-stage game with observable actions, an action a_i is conditionally dominated at stage t given history h_t if in the subgame starting at h_t every strategy for player i that assigns positive probability to a_i is strictly dominated.*

Proposition 2. *In any perfect information game, every subgame perfect equilibrium survives iterated elimination of conditionally dominated strategies.*

12. Bargaining with Alternating Offers

The set of players is $N = \{1, 2\}$. For $i = 1, 2$ we write $j = 3 - i$. The set of feasible utility pairs is $U \subset \mathbb{R}^2$, assumed to be compact and convex with $(0, 0) \in U$.[6] Time is discrete and infinite, $t = 0, 1, \ldots$ Each player i discounts payoffs by δ_i, so receiving u_i at time t is worth $\delta_i^t u_i$.

Rubinstein (1982) analyzes the following perfect information game. At every time $t = 0, 1, \ldots$, player $i(t)$ proposes an alternative $u = (u_1, u_2) \in U$ to player $j = 3 - i(t)$; the bargaining protocol specifies that $i(t) = 1$ for t even and $i(t) = 2$ for t odd. If j accepts the offer, then the game ends yielding a payoff vector $(\delta_1^t u_1, \delta_2^t u_2)$. Otherwise, the game proceeds to period $t + 1$. If agreement is never reached, each player receives a 0 payoff.

For each player i, it is useful to define the function g_i by setting

$$g_i(u_j) = \max\{u_i |\, (u_1, u_2) \in U\}.$$

Notice that the graphs of g_1 and g_2 coincide with the Pareto-frontier of U.

12.1. Stationary subgame perfect equilibrium.

Let (m_1, m_2) be the unique solution to the following system of equations

$$\begin{aligned} m_1 &= \delta_1 g_1(m_2) \\ m_2 &= \delta_2 g_2(m_1). \end{aligned}$$

Note that (m_1, m_2) is the intersection of the graphs of the functions $\delta_2 g_2$ and $(\delta_1 g_1)^{-1}$.

We are going to argue that the following "stationary" strategies constitute the unique subgame perfect equilibrium. In any period where player i has to make an offer to j, he offers u with $u_j = m_j$ and j accepts only offers u with $u_j \geq m_j$. We can use the *single-deviation principle* to check that this is a subgame perfect equilibrium.

12.2. Equilibrium uniqueness.

We prove that the subgame perfect equilibrium is unique by arguing that it is essentially the only strategy profile that survives *iterated conditional dominance*.

[6] The set of feasible utility outcomes U can be generated from a set of contracts or decisions X in a natural way. Define $U = \{(v_1(x), v_2(x)) \,|\, x \in X\}$ for a pair of utility functions v_1 and v_2 over X. With additional assumptions on X, v_1, v_2 we can ensure that the resulting U is compact and convex.

Theorem 10. *If a strategy profile survives iterative elimination of conditionally dominated strategies, then it is identical to the stationary subgame perfect equilibrium except for the nodes at which a player is indifferent between accepting and rejecting an offer in the subgame perfect equilibrium.*

Proof. Since player i can get 0 by never reaching an agreement, offering an alternative that gives him less than

$$m_i^0 = 0$$

or accepting such an offer at any history is conditionally dominated. All such offers are eliminated at the first stage of the iteration. Then i should never expect to receive more than

$$M_i^0 = \delta_i g_i (0)$$

in any future period following a disagreement. Hence rejecting an offer u with $u_i > M_i^0$ is conditionally dominated by accepting such an offer for i. Once we eliminate the latter strategies, i always accepts offers u with $u_i > M_i^0$ from j. Then making offers u with $u_i > M_i^0$ is dominated for j by offers $\bar{u} = \lambda u + (1 - \lambda) (M_i^0, g_j (M_i^0))$ for $\lambda \in (0, 1)$. We remove all the strategies involving such offers.

Under the surviving strategies, j can reject an offer from i and make an offer next period that leaves him with slightly less than $g_j (M_i^0)$, which i accepts. Hence accepting any offer that gives him less than

$$m_j^1 = \delta_j g_j \left(M_i^0\right)$$

is dominated for j. Moreover, making such offers is dominated for j because we argued above that offers with $u_i > M_i^0$ are dominated. After we eliminate such moves, i cannot expect more than

$$M_i^1 = \delta_i g_i \left(m_j^1\right) = \delta_i g_i \left(\delta_j g_j \left(M_i^0\right)\right)$$

in any future period following a disagreement.

We can recursively define the sequences

$$\begin{aligned} m_j^{k+1} &= \delta_j g_j \left(M_i^k\right) \\ M_i^{k+1} &= \delta_i g_i \left(m_j^{k+1}\right) \end{aligned}$$

for $i = 1, 2$ and $k \geq 1$. Since both g_1 and g_2 are decreasing functions, we can easily show that the sequence (m_i^k) is increasing and (M_i^k) is decreasing. By arguments similar to those above, we can prove by induction on k that, at some stage in the iteration, player $i = 1, 2$

- never accepts or makes an offer with $u_i < m_i^k$
- always accepts offers with $u_i > M_i^k$, but making such offers is dominated for j.

The sequences (m_i^k) and (M_i^k) are monotonic and bounded, so they need to converge. The limits satisfy

$$\begin{aligned} m_j^\infty &= \delta_j g_j \left(\delta_i g_i \left(m_j^\infty\right)\right) \\ M_i^\infty &= \delta_i g_i \left(m_j^\infty\right). \end{aligned}$$

It follows that (m_1^∞, m_2^∞) is the (unique) intersection point of the graphs of the functions $\delta_2 g_2$ and $(\delta_1 g_1)^{-1}$. Moreover, $M_i^\infty = \delta_i g_i \left(m_j^\infty\right) = m_i^\infty$. Therefore, no strategy for i that rejects u with $u_i > M_i^\infty = m_i^\infty$ or accepts u with $u_i < m_i^\infty = M_i^\infty$ survives iterated elimination of conditionally dominated strategies. Also, no strategy for i to offer u with $u_i \neq M_i^\infty = m_i^\infty$ survives. □

12.3. **Properties of the subgame perfect equilibrium.** The subgame perfect equilibrium is efficient—agreement is obtained in the first period, without delay. The subgame perfect equilibrium payoffs are given by by $(g_1(m_2), m_2)$, where (m_1, m_2) solve

$$\begin{aligned} m_1 &= \delta g_1 \left(m_2\right) \\ m_2 &= \delta g_2 \left(m_1\right). \end{aligned}$$

It can be easily shown that the payoff of player i is increasing in δ_i and decreasing in δ_j. For a fixed $\delta_j \in (0, 1)$, the limit payoff of player i converges to 0 as $\delta_i \to 0$ and to $\max_{u \in U} u_i$ as $\delta_i \to 1$. If U is symmetric and $\delta_1 = \delta_2$, player 1 enjoys a first mover advantage because $m_1 = m_2$ and $g_1(m_2) > m_2$.

13. Nash Bargaining

Assume that U is such that g_2 is decreasing, strictly concave and continuously differentiable (derivative exists and is continuous). The **Nash (1950) bargaining solution** u^* is defined by $\{u^*\} = \arg\max_{u \in U} u_1 u_2 = \arg\max_{u \in U} u_1 g_2(u_1)$. It is the outcome $(u_1^*, g_2(u_1^*))$ uniquely pinned down by the first order condition $g_2(u_1^*)+u_1^*g_2'(u_1^*)=0$. Indeed, since g_2 is decreasing and strictly concave, the function f, given by $f(x)=g_2(x)+xg_2'(x)$, is strictly decreasing and continuous and changes sign on the relevant range.

Theorem 11 (Binmore, Rubinstein and Wolinsky 1985)**.** *Suppose that $\delta_1=\delta_2=:\delta$ in the alternating bargaining model. Then the unique subgame perfect equilibrium payoffs converge to the Nash bargaining solution as $\delta\to 1$.*

Proof. Recall that the subgame perfect equilibrium payoffs are given by $(g_1(m_2),m_2)$ where (m_1,m_2) satisfies

$$\begin{aligned} m_1 &= \delta g_1(m_2) \\ m_2 &= \delta g_2(m_1). \end{aligned}$$

It follows that $g_1(m_2)=m_1/\delta$, hence $m_2=g_2(g_1(m_2))=g_2(m_1/\delta)$. We rewrite the equations as follows

$$\begin{aligned} g_2(m_1/\delta) &= m_2 \\ g_2(m_1) &= m_2/\delta. \end{aligned}$$

By the mean value theorem, there exists $\xi\in(m_1,m_1/\delta)$ such that $g_2(m_1/\delta)-g_2(m_1)=(m_1/\delta-m_1)g_2'(\xi)$, hence $(m_2-m_2/\delta)=(m_1/\delta-m_1)g_2'(\xi)$ or, equivalently, $m_2+m_1g_2'(\xi)=0$. Substituting $m_2=\delta g_2(m_1)$ we obtain $\delta g_2(m_1)+m_1g_2'(\xi)=0$.

Note that $(g_1(m_2),m_2)$ converges to u^* as $\delta\to 1$ if and only if (m_1,m_2) does. In order to show that (m_1,m_2) converges to u^* as $\delta\to 1$, it is sufficient to show that any limit point of (m_1,m_2) as $\delta\to 1$ is u^*. Let (m_1^*,m_2^*) be such a limit point corresponding to a sequence $(\delta_k)_{k\geq 0}\to 1$. Recognizing that m_1,m_2,ξ are functions of δ, we have

$$\delta_k g_2(m_1(\delta_k))+m_1(\delta_k)g_2'(\xi(\delta_k))=0. \tag{13.1}$$

Since $\xi(\delta_k)\in(m_1(\delta_k),m_1(\delta_k)/\delta_k)$ with $m_1(\delta_k),m_1(\delta_k)/\delta_k\to m_1^*$ as $k\to\infty$ and g_2' is continuous by assumption, in the limit 13.1 becomes $g_2(m_1^*)+m_1^*g_2'(m_1^*)=0$. Therefore, $m_1^*=u_1^*$. □

14. Sequential Equilibrium

In multi-stage games with incomplete information, say where payoffs depend on initial moves by nature, the only proper subgame is the original game, even if players observe one another's actions at the end of each period. Thus the refinement of Nash equilibrium to subgame perfect equilibrium has no bite. Since players do not know the others' types, the start of a period can only be analyzed as a separate subgame when the players' posterior beliefs are specified. The concept of sequential equilibrium proposes a way to derive plausible beliefs at every information set. Based on the beliefs, one can test whether the continuation strategies form a Nash equilibrium.

The complications that incomplete information causes are easiest to see in "signaling games"—leader-follower games in which only the leader has private information. The leader moves first; the follower observes the leader's action, but not the leader's type, before choosing his own action. One example is Spence's (1974) model of the job market. In that model, the leader is a worker who knows her productivity and must choose a level of education; the follower, a firm (or number of firms), observes the worker's education level, but not her productivity, and then decides what wage to offer her. In the spirit of subgame perfection, the optimal wage should depend on the firm's beliefs about the worker's productivity given the observed education. An equilibrium needs to specify not only contingent actions, but also beliefs. At information sets that are reached with positive probability in equilibrium, beliefs should be derived using Bayes' rule. However, there are some theoretical issues about belief update following zero-probability events.

Refer for more motivation to the example in FT, figure 8.1 (p. 322). The strategy profile (L, A) is a Nash equilibrium, which is subgame perfect as player 2's information set does not initiate a proper subgame. However, it is not a very plausible equilibrium, since player 2 prefers playing B rather than A at his information set, regardless of whether player 1 has chosen M or R. So, a good equilibrium concept should rule out the solution (L, A) in this example and ensure that 2 always plays B. The problem with the considered equilibrium is that player 2 does not play a best response to any possible belief at his information set.

For most definitions, we focus on extensive form games of prefect recall with finite sets of decision nodes. We use some of the notation introduced earlier.

A **sequential equilibrium** (Kreps and Wilson 1982) is an assessment (σ, μ), where σ is a (behavior) strategy profile and μ is a **system of beliefs**. The latter component consists of a belief specification $\mu(h)$ over the nodes at each information set h. The definition of sequential equilibrium is based on the concepts of sequential rationality and consistency. **Sequential rationality** requires that conditional on every information set h, the strategy $\sigma_{i(h)}$ be a best response to $\sigma_{-i(h)}$ given the beliefs $\mu(h)$. Formally,

$$u_{i(h)}(\sigma_{i(h)}, \sigma_{-i(h)}|h, \mu(h)) \geq u_{i(h)}(\sigma'_{i(h)}, \sigma_{-i(h)}|h, \mu(h))$$

for all information sets h and alternative strategies σ'. Here the conditional payoff $u_i(\sigma|h, \mu(h))$ now denotes the payoff that results when play begins at a randomly selected node in the information set h, where the probability distribution on these nodes is given by $\mu(h)$, and subsequent play at each information set is as specified by the profile σ.

Beliefs need to be **consistent** with strategies in the following sense. For any fully mixed strategy profile $\tilde{\sigma}$—that is, one where each action is played with positive probability at every information set—all information sets are reached with positive probability and Bayes' rule leads to a unique system of beliefs $\mu^{\tilde{\sigma}}$. The assessment (σ, μ) is consistent if there exist a sequence of fully mixed strategy profiles $(\sigma^m)_{m\geq 0}$ converging to σ such that the associated beliefs μ^{σ^m} converge to μ as $m \to \infty$.

Definition 11. *A sequential equilibrium is an assessment which is sequentially rational and consistent.*

The definition of sequential equilibrium rules out the strange equilibrium in the earlier example (FT figure 8.1). Since player 1 chooses L under the proposed equilibrium strategies, consistency does not pin down player 2's beliefs at his information set. However, sequential rationality requires that player 2 have some beliefs and best-respond to them, which ensures that A is not played.

Consistency imposes more restrictions than Bayes' rule alone. Consider figure 8.3 in FL (p. 339). The information set h_1 of player 1 consists of two nodes x, x'. Player 1 can take an action D leading to y, y' respectively. Player 2 cannot distinguish between y and y' at the information set h_2. If 1 never plays D in equilibrium, then Bayes' rule does not pin down beliefs at h_2. However, consistency implies that $\mu_2(y|h_2) = \mu_1(x|h_1)$. The idea is that since

1 cannot distinguish between x and x', he is equally likely to tremble at either node. Hence trembles ensure that players' beliefs respect the information structure.

More generally, consistency imposes common beliefs following deviations from equilibrium behavior. There are criticisms of this requirement—why should different players have the same theory about something that was not supposed to happen? A contra-argument is that consistency matches the spirit of equilibrium analysis, which normally assumes that players agree in their beliefs about other players' strategies (namely, players share correct conjectures about each other's strategies).

15. Properties of Sequential Equilibrium

Theorem 12. *A sequential equilibrium exists for every finite extensive-form game.*

This is a consequence of the existence of perfect equilibria, which we prove later.

Proposition 3. *The sequential equilibrium correspondence is upper hemi-continuous with respect to payoffs.*

Proof. Let $u^k \to u$ be a convergent sequence of payoff functions and $(\sigma^k, \mu^k) \to (\sigma, \mu)$ be a convergent sequence of sequential equilibria of the games with corresponding payoffs u^k. We need to show that (σ, μ) is a sequential equilibrium for the game with payoffs given by u. Sequential rationality of (σ, μ) is straightforward because the expected payoffs conditional on reaching any information set are continuous in the payoff functions and beliefs.

We also have to check consistency of (σ, μ). As (σ^k, μ^k) is a sequential equilibrium of the game with payoff function u^k, there exists a sequence of completely mixed strategies $(\sigma^{m,k})_m \to \sigma^k$, with corresponding induced beliefs given by $(\mu^{m,k})_m \to \mu^k$. For every k, we can find a sufficiently large m_k so that each component of σ^{k,m_k} and μ^{m_k} are within $1/k$ from the corresponding one under σ^k and μ^k. Since $\sigma^k \to \sigma, \mu^k \to \mu$, it must be that $\sigma^{m_k,k} \to \sigma, \mu^{m_k,k} \to \mu$. Thus we have obtained a sequence of fully mixed strategies converging to σ, which induces beliefs converging to μ. □

Kreps and Wilson show that in generic games (i.e., a space of payoff functions such that the closure of its complement has measure zero), the set of sequential equilibrium outcome distributions is finite. Nevertheless, it is not generally true that the set of sequential equilibria is finite, as there may be infinitely many belief specifications for off-path information sets that support some equilibrium strategies. We provide an illustration in the context of the beer-or-quiche signaling game of Cho and Kreps (1987).

See figure 11.6 in FT (p. 450). Player 1 is wimpy or surly, with respective probabilities 0.1 or 0.9. Player 2 is a bully who would like to fight the wimpy type but not the surly one. Player 1 orders breakfast and 2 decides whether to fight him after observing his breakfast choice. Player 1 gets a utility of 1 from having his favorite breakfast—beer if surly, quiche if weak—but a disutility of 2 from fighting. When player 1 is weak, player 2's utility is 1 if he fights and 0 otherwise; when 1 is surly, the payoffs to the two actions are reversed. One can show that there are two classes of sequential equilibria, corresponding to two distinct outcomes. In one set of sequential equilibria, both types of player 1 drink beer, while in the other both types of player 1 eat quiche. In both cases, player 2 must fight with probability at least 1/2 when observing the out-of-equilbrium breakfast in order to make the mismatched type of player 1 endure gastronomic horror. Note that either type of equilibrium can be supported with any belief for player 2 placing a probability weight of at least 1/2 on player 1 being wimpy following the out-of-equilbrium breakfast. Hence there is an infinity of sequential equilibrium assessments.

Kohlberg and Mertens (1986) criticized sequential equilibrium for allowing "strategically neutral" changes in the game tree to affect the equilibrium. Compare, for instance, the two games in FT figure 8.6 (p. 343). The game on the right is identical to the one on the left, except that player 1's first move is split into two moves in a seemingly irrelevant way. Whereas (A, L_2) can be supported as a sequential equilibrium for the game on the left, the strategy A is not part of a sequential equilibrium for the one on the right. For the latter game, in the simultaneous-move subgame following NA, the only Nash equilibrium is (R_1, R_2), as L_1 is strictly dominated by R_1 for player 1. Hence the unique sequential equilibrium strategies for the right-hand game are (NA, R_1, R_2).

Note that the sensitivity of sequential equilibrium to the addition of "irrelevant moves" is not a direct consequence of consistency, but is rather implied by sequential rationality. In the example above, the problem arises even for subgame perfect equilibria. Kohlberg and Mertens (1986) further develop these ideas in their concept of a stable equilibrium. However, their proposition that mistakes be "conditionally optimal" is not necessarily compelling. If we take seriously the idea that players make mistakes at each information set, then it is not clear that the two extensive forms above are equivalent. In the game on the right, if player 1 makes the mistake of not playing A, he is still able to ensure that R_1 is more likely than L_1; in the game on the left, he might take either action by mistake when intending to play A.

16. Perfect Bayesian Equilibrium

Perfect Bayesian equilibrium was the original solution concept for extensive-form games with incomplete information, when subgame-perfection does not have enough force. It incorporated the ideas of sequential rationality and Bayesian updating of beliefs. Nowadays sequential equilibrium (which was invented later) is the preferred way of expressing these ideas, but it's worthwhile to know about PBE since older papers refer to it.

The idea is similar to sequential equilibrium but with more basic requirements about how beliefs are updated. Fudenberg & Tirole (1991) have a paper that describes various formulations of PBE. The basic requirements are that strategies should be sequentially rational and that beliefs should be derived from Bayes's rule wherever applicable, with no constraints on beliefs at information sets reached with probability zero in equilibrium.

Other properties that can be imposed:

- In a multi-stage game with independent types — i.e. exactly one move by Nature, at the beginning of the game, assigning types to players and such that types are independently distributed, with all subsequent actions of observed — beliefs about different players should remain independent at each history. (PBE is usually applied to games in which Nature moves only at the beginning and actions are observed.)
- Updating should be "consistent": given a probability-zero history h^t at time t, from which strategies do call for a positive-probability transition to history h^{t+1}, the belief at h^{t+1} should be given by updating beliefs at h^t via Bayes's rule.
- "Not signaling what you don't know": beliefs about player i at the beginning of period $t+1$ depend only on h^t and action by player i at time t, not also on other players' actions at time t.
- Two different players i, j should have the same belief about a third player k even at probability-zero histories.

All of these conditions are implied by consistency.

Anyhow, there does not seem to be a single clear definition of PBE in the literature. Different sets of conditions are imposed by different authors. For this reason, using sequential equilibrium is preferable.

17. Perfect Equilibrium

Now consider the following game:

	L	R
U	$1,1$	$0,0$
D	$0,0$	$0,0$

Both (U, L) and (D, R) are sequential equilibria (sequential equilibrium coincides with Nash equilibrium in a normal-form game). But (D, R) seems non-robust: if player 1 thinks that player 2 might make a mistake and play L with some small probability, he would rather deviate to U. This motivates the definition of **(trembling-hand) perfect equilibrium** (Selten, 1975) for normal-form games. A profile σ is a PE if there is a sequence of "trembles" $\sigma^m \to \sigma$, where each σ^m is a completely mixed strategy, such that σ_i is a best reply to σ_{-i}^m for each m.

An equivalent approach is to define a strategy profile σ^ε to be an **ε-perfect equilibrium** if there exist $\varepsilon(s_i) \in (0, \varepsilon)$ for all i, all s_i, such that σ^ε is a Nash equilibrium of the game where players are restricted to play mixed strategies where every strategy s_i has probability at least $\varepsilon(s_i)$. A PE is a profile that is a limit of some sequence of ε-perfect equilibria σ^ε as $\varepsilon \to 0$. (We will not show the equivalence here but it's not too hard.)

Theorem 13. *Every finite normal-form game has a perfect equilibrium.*

Proof. For any $\varepsilon > 0$, we can certainly find a Nash equilibrium of the modified game, where each player is restricted to play mixed strategies that place probability at least ε on every pure action. (Just apply the usual Nash existence theorem for compact strategy sets and quasiconcave payoffs.) By compactness, there is some subsequence of these strategy profiles as $\varepsilon \to 0$ that converges, and the limit point is a perfect equilibrium by definition. □

We would like to extend this definition to extensive-form games. Consider the game in Fig 8.11 (p. 353) of FT. They show an extensive-form game and its reduced normal form. There is a unique SPE $(L_1 L_1', L_2)$. But (R_1, R_2) is a PE of the reduced normal form. Thus perfection in the normal form does *not* imply subgame-perfection. The perfect equilibrium is sustained only by trembles such that, conditional on trembling to L_1 at the first node, player 1 is also much more likely to play R_1' than L_1' at his second node. This seems unreasonable — R_1' is only explainable as a tremble. Perfect equilibrium as defined so far thus has the disadvantage of allowing correlation in trembles at different information sets.

The solution to this is to impose perfection in the **agent-normal form**. We treat the two different nodes of player 1 as being different players, thus requiring them to tremble independently. More formally, in the agent-normal form game, we have a player corresponding to every information set. Given a strategy profile for all the players, each "player" corresponding to an information set h gets payoff given by the payoff of player $i(h)$ from the corresponding strategies in the extensive-form game. Thus, the game in figure 8.11 turns into a three-player game. The only perfect equilibrium of this game is (L_1, L_1', L_2).

More generally, a **perfect equilibrium** in an extensive-form game is defined to be a perfect equilibrium of the corresponding agent-normal form.

Theorem 14. *Every PE of a finite extensive-form game is a sequential equilibrium (for some appropriately chosen beliefs).*

Proof. Let σ be the given PE. So there exist fully mixed strategy profiles $\sigma^m \to \sigma$ which are ε-perfect equilibria of the agent-normal form game with $\varepsilon \to 0$. For each σ^m we have a well-defined belief system induced by Bayes's rule. Pick a subsequence for which these belief systems converge, to some μ. Then by definition (σ, μ) is consistent. Sequential rationality follows exactly from the fact that σ is a perfect equilibrium of the agent-normal form, using the first definition of perfect equilibrium. (More properly, this implies that there are no one-shot deviations that benefit any player; by an appropriate adaptation of the one-shot deviation principle this shows that σ is in fact fully sequentially rational at every information set.) $\square$

The converse is not true — not every sequential equilibrium is perfect, as we already saw with the simple normal-form example above. But for generic payoffs it is true (Kreps & Wilson, 1982).

The set of perfect equilibrium outcomes is not upper-hemicontinuous (unlike sequential equilibrium or subgame-perfect equilibrium). Consider the following game:

	L	R
U	$1, 1$	$0, 0$
D	$0, 0$	$1/n, 1/n$

It has (D, R) as a perfect equilibrium for each $n > 0$, but in the limit where (D, R) has payoffs $(0, 0)$ it is no longer a perfect equilibrium. We can think of this as an order-of-limits problem: as $n \to \infty$ the trembles against which D and R remain best responses become smaller and smaller.

18. Proper Equilibrium

Myerson (1978) considered the notion that when a player trembles, he is still more likely to play better actions than worse ones. Myerson's notion is that a player's probability of playing the second-best action is at most ε times the probability of the best action, the probability of the third-best action is at most ε times the probability of the second-best action, and so forth. Consider the game in Fig. 8.15 of FT (p. 357). (M, M) is a perfect equilibrium, but Myerson argues that it can be supported only using unreasonable trembles, where each player has to be likely to tremble to a very bad reply rather than an almost-best reply.

Definition 12. *A ε-**proper equilibrium** is a totally mixed strategy profile σ^ε such that, if $u_i(s_i, \sigma^\varepsilon_{-i}) < u_i(s'_i, \sigma^\varepsilon_{-i})$, then $\sigma^\varepsilon_i \leq \varepsilon\sigma^\varepsilon_i(s'_i)$. A **proper equilibrium** is any limit of some ε-proper equilibria as $\varepsilon \to 0$.*

Theorem 15. *Every finite normal-form game has a proper equilibrium.*

Proof. First prove existence of ε-proper equilibria, using the usual Kakutani argument applied to the "almost-best-reply" correspondences BR^ε_i rather than the usual best-reply correspondences. ($BR^\varepsilon_i(\sigma_{-i})$ is the set of mixed strategies for player i in a suitable compact space of fully mixed strategies that satisfy the inequality in the definiton of ε-proper equilibrium.) Then use compactness to see that there exists a sequence that converges as $\varepsilon \to 0$; its limit is a proper equilibrium. □

Given an extensive-form game, a proper equilibrium of the corresponding normal form is automatically subgame-perfect; we don't need to go to the agent-normal form. We can show this by a backward-induction-type argument.

Kohlberg and Mertens (1986) showed that a proper equilibrium in a normal-form game is sequential in every extensive-form game having the given normal form. However, it will not necessarily be a trembling-hand perfect equilibrium in (the agent-normal form of) every such game. See Figure 8.16 of FT (p. 358): (Lr) is proper (and so sequential) but not perfect in the agent-normal form.

19. Forward Induction in Signaling Games

Consider now a **signaling game**. There are two players, a sender S and a receiver R. There is a set T of types for the sender; the realized type will be denoted by t. $p(t)$ denotes the probability of type t. The sender privately observes his type t, then sends a message $m \in M(t)$. The receiver observes the message and chooses an action $a \in A(m)$. Finally both players receive payoffs $u_S(t, m, a), u_R(t, m, a)$; thus the payoffs potentially depend on the true type, the message sent, and the action taken by the receiver.

In such a game we will use $T(m)$ to denote the set $\{t \mid m \in M(t)\}$.

The beer-quiche game from before is an example of such a game. T is the set $\{weak, surly\}$; the messages are $\{beer, quiche\}$; the actions are $\{fight, not\ fight\}$. As we saw before, there are two sequential equilibria: one in which both types of sender choose beer, and another in which both types choose quiche. In each case, the equilibrium is supported by some beliefs such that the sender is likely to have been weak if he chose the unused message, and the receiver responds by fighting in this case.

Cho and Kreps (1987) argued that the equilibrium in which both types choose quiche is unreasonable for the following reason. It does not make any type for the weak type to deviate to ordering beer, no matter how he thinks that the receiver will react, because he is already getting payoff 3 from quiche, whereas he cannot get more than 2 from switching to beer. On the other hand, the surly type can benefit if he thinks that the receiver will react by not fighting. Thus, conditional on seeing beer ordered, the receiver should conclude that the sender is surly and so will not want to fight.

On the other hand, this argument does not rule out the equilibrium in which both types drink beer. In this case, in equilibrium the surly type is getting 3, whereas he gets at most 2 from deviating no matter how the receiver reacts; hence he cannot want to deviate. The weak type, on the other hand, is getting 2, and he can get 3 by switching to quiche if he thinks this will induce the receiver not to fight him. Thus only the weak type would deviate, so the sender's belief (that the receiver is weak if he orders quiche) is reasonable.

Now consider modifying the game by adding an extra option for the receiver: paying a million dollars to the sender. Now the preceding argument doesn't rule out the quiche equilibrium — either type of sender might deviate to beer if he thinks this will induce the receiver to pay him a million dollars. Hence, in order for the argument to go through, we need the additional assumption that the sender cannot expect the receiver to play a bad strategy.

Cho and Kreps formalized this line of reasoning in the **intuitive criterion**, as follows. For any set of types $T' \subseteq T$, write

$$BR(T', m) = \cup_{\mu \ | \ \mu(T')=1} BR(\mu, m)$$

— the set of strategies that R could reasonably play if he observes m and is sure that the sender's type is in T'. Now with this notation established, consider any sequential equilibrium, and let $u_S^*(t)$ be the equilibrium payoff to a sender of type t. Define

$$\tilde{T}(m) = \{t \mid u_S^*(t) > \max_{a \in BR(T(m),m)} u_S(t, m, a)\}.$$

This is the set of types that do better in equilibrium than they could possibly do by sending m, no matter how R reacts, as long as R is playing a best reply to some belief. We then say that the proposed equilibrium fails the intuitive criterion if there exist a type t' and a message m such that

$$u_S^*(t') < \min_{a \in BR(T(m) \setminus \tilde{T}(m),m)} u_S(t', m, a).$$

In words, the equilibrium fails the intuitive criterion if some type t' of the sender is getting a lower payoff than any payoff he could possibly get by playing m if he could thereby convince the sender that he could not possibly be in $\tilde{T}(m)$.

In the beer-quiche example, the all-quiche equilibirum fails this criterion: let $t' = surly$ and $m = beer$; check that $\tilde{T}(m) = \{weak\}$.

Now we can apply this procedure repeatedly, giving the **iterated intuitive criterion**. We can use the intuitive criterion as above to rule out some pairs (t, m) — type t cannot conceivably send message m. Now we can rule out some actions of the receiver, by requiring that the receiver should be playing a best reply to some belief about the types that have not yet been eliminated (given the message). Given this elimination, we can go back and possibly rule out more pairs (t, m), and so forth.

This idea has been further developed by Banks and Sobel (1987). They say that type t' is **infinitely more likely** to choose the out-of-equilibrium message m than type t under the following condition: the set of possible best-replies by the receiver (possibly mixed) that make t' strictly prefer to deviate to m is a strict superset of the possible best-replies that make t weakly prefer to deviate. If this holds, then conditional on observing m, the receiver should put belief 0 on type t. The analogue of the Intuitive Criterion under this elimination procedure is known as D1. If we allow t' to vary across different best replies by the sender, requiring only that every mixed best reply that weakly induced t to deviate would also strictly induce *some* t' to deviate, then this gives criterion D2. We can also apply either of these restrictions on beliefs to eliminate possible actions by the receiver, and proceed iteratively. Iterating D2 leads to the equilibrium refinement criterion known as **universal divinity**.

The motivating application is Spence's job-market signaling model. With just two types of job applicant, the intuitive criterion selects the equilibrium where the low type gets the lowest level of education and the high type gets just enough education to deter the low type. With more types, the intuitive criterion no longer accomplishes this. D1 does manage to uniquely select the separating equilibrium that minimizes social waste by having each type get just enough education to deter the next-lower type from imitating him.

20. Forward Induction in General

The preceding ideas are all attempts to capture some kind of **forward induction**: players should believe in the rationality of their opponents, even after observing a deviation; thus if you observe an out-of-equilibrium action being played, you should believe that your opponent expected you to play in a way that made his action reasonable, and this in turn is informative about his type (or, in more general extensive forms, about how he plans to play in the future). Forward induction is not itself an equilibrium concept, since equilibrium means that the specified strategies are to be followed even after a deviation; rather, it is an attempt to describe reasoning by players who are not quite certain about what will be played.

Consider now the extensive-form game as follows: 1 can play O, leading to $(2, 2)$, or I, leading to the following battle-of-the-sexes game:

	T	W
T	$0, 0$	$3, 1$
W	$1, 3$	$0, 0$

There is an SPE in which player 1 first plays O; conditional on playing I, they play the equilibrium (W, T). But the following forward-induction argument suggests this equilibrium is unreasonable: if player 1 plays I, this suggests he is expecting to coordinate on (T, W) in the battle-of-the-sexes game, so player 2, anticipating this, will play W. Thus if 1 can convince 2 to play W by playing I in the first stage, he can get the higher payoff $(3, 1)$.

This can also be represented in (reduced) normal form.

	T	W
O	$2, 2$	$2, 2$
IT	$0, 0$	$3, 1$
IW	$1, 3$	$0, 0$

This representation of the game shows a connection between forward induction and strict dominance. We can rule out IW because it is dominated by O; then the only perfect equilibrium of the remaining game is (IT, W) giving payoffs $(3, 1)$. However, (O, T) can be enforced as a perfect (in fact a proper) equilibrium in the normal-form game.

Kohlberg and Mertens (1986) argue that an equilibrium concept that is not robust to deletion of strictly dominated strategies is troubling. The above example, together with other cases of such non-robustness, leads them to define the notion of **stable equilibria**. It is a set-valued concept — not a property of individual equilibrium but of sets of strategies, one for each player. They first argue that a solution concept should meet the following requirements:

- Iterated dominance: every strategically stable set must contain a strategically stable set of any game obtained by deleting a strictly dominated strategy.
- Admissibility: no mixed strategy in a strategically stable set assigns positive probability to a strictly dominated strategy.
- Invariance to extensive-form representation: they define an equivalence relation between extensive forms and require that any stable set in one game should be stable in any equivalent game.

They then define strategic stability in a way such that these criteria are satisfied. Their definition is as follows: A closed set S of NE is **strategically stable** if it is minimal among sets with the following property: for every $\eta > 0$, there exists $\varepsilon' > 0$ such that, for all $\varepsilon < \varepsilon'$, all choices of $0 < \varepsilon(s_i) \leq \varepsilon$ for each player i and strategies s_i, the game where each player i is constrained to play every s_i with probability at least $\varepsilon(s_i)$ has a Nash equilibrium which is within distance η of some equilibrium in S.

Thus, any sequence of ε-perturbed games as $\varepsilon \to 0$ should have equilibria corresponding to an equilibrium in S. Notice that we need the minimality property of S to give bite to this definition — otherwise, by upper hemi-continuity, we know that the set of all Nash equilibria would be strategically stable, and we get no refinement.

The difference with trembling-hand perfection is that there should be convergence to one of the selected equilibria for *any* sequence of perturbations, not just some sequence of perturbations.

They have a theorem that there exists some stable set that is contained in a connected component of the set of Nash equilibria. Generically, each component of the set of Nash equilibria leads to a single distribution over outcomes in equilibrium; thus, generically, there exists a stable set that determines a unique outcome distribution. Moreover, any stable set contains a stable set of the game obtained by elimination of a weakly dominated strategy.

Moreover, stable sets have an even stronger property, "never a weak best reply": given a stable set S of equilibria of a game, if we remove a strategy for some player i that is not a best reply to the strategies of players $-i$ at any equilibrium in the set, then the remaining game has a stable set of equilibria contained in S. This means that the concept of stable equilibrium robust to forward induction: knowing that player i will not use a particular strategy does not create new equilibria in the stable set.

There are actually a lot of stability concepts in the literature. Mertens has more papers with alternative definitions.

Every equilibrium in a stable set has to be a perfect equilibrium. This follows from the minimality condition — if an equilbrium is not a limiting equilibrium along some sequence of trembles, then there's no need to include it in the stable set. But notice, these equilibria are only guaranteed to be perfect in the normal form, not in the agent-normal form (if the game represented is an extensive-form one).

Some recent papers further develop these ideas. Battigalli and Siniscalchi (2002) are interested in the epistemic conditions that lead to forward induction. They have an epistemic model, with state of nature of the form $\omega = (s_i, t_i)_{i \in N}$, where s_i represents player i's disposition to act and t_i represents his disposition to believe. t_i specifies a belief $g_{i,h} \in \Delta(\Omega_{-i})$ over states of the other players for each information set h of player i. We saw i is rational at state ω if s_i is a best reply to his beliefs t_i at each information set. Let R be the set of states at which every player is rational. For any event $E \subseteq \Omega$, we can define the set $B_{i,h}(E) = \{(s,t) \in \Omega \mid g_{i,h}(E) = 1\}$, i.e. the set of states where i is sure that E has occurred (at information set h). We can define $B_h(E) = \cap_i B_{i,h}$. Finally $SB_i(E) = \cap_h B_{i,h}(E)$, the set of states at which i **strongly believes** in event E, meaning the set of states at which i would be sure of E as long as he's reached an information set where E is possible. Finally, they show that $SB(R)$ identifies forward induction — that is, in the states of the world where everyone strongly believes that everyone is sequentially rational, strategies must form a profile that is not ruled out by forward induction.

Battigalli and Siniscalchi take this a level further by iterating the strong-beliefs operator — everyone strongly believes that everyone strongly believes that everyone is rational, and so forth — and this operator leads to backward induction in games of perfect information; without perfect information, it leads to iterated deletion of strategies that are never a best reply. This gives a formalization of the idea of rationalizability in extensive-form games.

21. Repeated games

We now consider the standard model of repeated games. Let $G = (N, A, u)$ be a normal-form stage game. At time $t = 0, 1, \ldots$, the players simultaneously play game G. At each period, the players can all observe play in each previous period; the history is denoted $h^t = (a^0, \ldots, a^{t-1})$. Payoffs in the repeated game $RG(\delta)$ are given by $U_i = (1 - \delta_i)\sum_{t=0}^{\infty} \delta_i^t u_i(a^t)$. The $(1 - \delta_i)$ factor normalizes the sum so that payoffs in the repeated game are on the same scale as in the stage game. We assume players play behavior strategies (by Kuhn's theorem), so a strategy for player i is given by a choice of $\sigma_i(h^t) \in \Delta(A_i)$ for each history h^t.

Given such strategies, we can define continuation payoffs after any history h^t: $U_i(\sigma|h^t)$.

If α^* is a Nash equilibrium of the static game, then playing α^* at every history is a subgame-perfect equilibrium of the repeated game. Conversely: for any finite game G and any $\varepsilon > 0$, there exists $\bar{\delta}$ with the property that, for any $\delta < \bar{\delta}$, any SPE of the repeated game $RG(\delta)$ has the property that, at every history, play is within ε of a static NE. However, we usually care about players with high discount factors, not low discount factors.

The main results for repeated games are "Folk Theorems": for high enough δ, every feasible and individually rational payoff in the stage game can be enforced as an equilibrium of the repeated game. There are several versions of such a theorem, which is why we use the plural. For now, we look at repeated games with perfect monitoring (as just defined), where the appropriate equilibrium concept is SPE. The way to check an SPE is via the one-shot deviation principle. Payoffs from playing a at history h^t are given by the value function

$$V_i(a) = (1 - \delta)u_i(a) + \delta U_i(\sigma|h^t, a). \tag{21.1}$$

This gives us an easy way to check whether or not a player wants to deviate from a proposed strategy, given other player's strategies. σ is an SPE if and only if, for every history h^t, $\sigma|h^t$ is a NE of the induced game $G(h^t, \sigma)$ whose payoffs are given by (21.1).

To state a folk theorem, we need to explain the terms "individually rational" and "feasible." The **minmax payoff** of player i is the worst payoff his opponents can hold him down to if he knows their strategies:

$$\underline{v}_i = \min_{\alpha_{-i} \in \prod_{j \neq i} \Delta(A_j)} \left[\max_{a_i \in A_i} u_i(a_i, \alpha_{-i}) \right].$$

We will let m^i, a **minmax profile for** i, denote a profile of strategies (a_i, α_{-i}) that solves this minimization and maximization problem. Note that we require independent mixing by i's opponents.

In any SPE — in fact, any Nash equilibrium — i's payoff is at least his minmax payoff, since he can always get at least this much by just best-responding to his opponents' (possibly independently mixed) actions in each period separately. This motivates us to say that a payoff vector v is **individually rational** if $v_i \geq \underline{v}_i$ for each i, and it is **strictly individually rational** if the inequality is strict for each i.

The set of **feasible payoffs** is the convex hull of the set $\{u(a) \mid a \in A\}$. Again note that this can include payoffs that are not obtainable in the stage game using mixed strategies, because correlation between players may be required.

Also, in studying repeated games we usually assume the availability of a **public randomization device** that produces a publicly observed signal $\omega^t \in [0,1]$, uniformly distributed and independent across periods, so that players can condition their actions on the signal. Properly, we should include the signals (or at least the current period's signal) in the specification of the history, but it is conventional not to write it out explicitly. The public randomization device is a convenient way to convexity the set of possible equilibrium payoff vectors. (Fudenberg and Maskin (1991) showed that one can actually do this without the public randomization device for sufficiently high δ, by appropriate choice of which periods to play each action profile involved in any given convex combination.)

An easy folk theorem is that of Friedman (1971):

Theorem 16. *If e the payoff vector of some Nash equilibrium of G, and v is a feasible payoff vector with $v_i > e_i$ for each i, then for all sufficiently high δ, there exists an SPE with payoffs v.*

Proof. Just specify that the players play whichever action profile gives payoffs v (using the public randomization device to correlate their actions if necessary), and revert to the static Nash permanently if anyone has ever deviated. □

So, for example, if there is a Nash equilibrium that gives everyone their minmax payoff (for example, in the prisoner's dilemma), then every individually rational and feasible payoff vector is obtainable in SPE.

However, it would be nice to have a full, or nearly full, characterization of the set of possible equilibrium payoff vectors (for large δ). In many repeated games, the Friedman folk theorem is not strong enough for this. A more general folk theorem would say that every individually rational, feasible payoff is achievable in SPE under general conditions. This is harder to show, because in order for one player to be punished by minmax if he deviates, others need to be willing to punish him. Thus, for example, if all players have equal payoffs, then it may not be possible to punish a player for deviating, because the punisher hurts himself as well as the deviator.

For this reason, the standard folk theorem (due to Fudenberg and Maskin, 1986) requires a full-dimensionality condition.

Theorem 17. *Suppose the set of feasible payoffs V has full dimension n. For any feasible and strictly individually rational payoff vector v, there exists $\underline{\delta}$ such that whenever $\delta > \underline{\delta}$, there exists an SPE of $RG(\delta)$ with payoffs v.*

Actually we don't quite need the full-dimensionality condition — all we need, conceptually, is that there are no two players who have the same payoff functions; more precisely, no player's payoff function can be a positive affine transformation of any other's (Abreu, Dutta, & Smith, 1994). But the proof is easier under the stronger version.

Proof. We will assume that i's minmax action profile m^i is pure. Consider the action profile a for which $u(a) = v$. Choose v' in the interior of the feasible, individually rational set with $v'_i < v_i$ for each i. We can do this by full-dimensionality. Let w^i denote v'_i with ε added to each player's payoff except for player i; with ε low enough, this will again be a feasible payoff vector.

Strategies are now specified as follows.

- Phase I: play a, as long as there are no deviations. If i deviates, switch to II_i.
- Phase II_i: play m^i. If player j deviates, switch to II_j. Note that if m^i is a pure strategy profile it is clear what we mean by j deviating. If it requires mixing it is not so clear; this will be dealt with in the second part of the proof. Phase II_i lasts for N periods, where N is a number to be determined, and if there are no deviations during this time, play switches to III_i.

- Phase III_i: play the action profile leading to payoffs w^i forever. If j deviates, go to II_j. (This is the "reward" phase that gives players $-i$ incentives to punish in phase II_i.)

We check that there are no incentives to deviate, using the one-shot deviation principle for each of the three phases: calculate the payoff to i from complying and from deviating in each phase. Phases II_i and II_j ($j \neq i$) need to be considered separately, as do III_i and III_j.

- Phase I: deviating gives at most $(1-\delta)M + \delta(1-\delta^N)\underline{v}_i + \delta^{N+1}v_i'$, where M is some upper bound on all of i's feasible payoffs, and complying gives v_i. Whatever N we have chosen, it is clear that as long as δ is sufficiently close to 1, complying produces a higher payoff than deviating, since $v_i' < v_i$.
- Phase II_i: Suppose there are $N' \leq N$ remaining periods in this phase. Then complying gives i a payoff of $(1-\delta^{N'})\underline{v}_i + \delta^{N'}v_i'$, whereas since i is being minmaxed, deviating can't help in the current period and leads to N more periods of punishment, for a total payoff of at most $(1-\delta^{N+1})\underline{v}_i + \delta^{N+1}v_i'$. Thus deviating is always worse than complying.
- Phase II_j: With N' remaining periods, i gets $(1-\delta^{N'})u_i(m^j) + \delta^{N'}(v_j' + \varepsilon)$ from complying and at most $(1-\delta)M + (\delta - \delta^N)\underline{v}_i + \delta^N v_i'$ from deviating. When δ is large enough, complying is preferred.
- Phase III_i: This is the one case that affects the choice of N. Complying gives v_i' in every period, while deviating gives at most $(1-\delta)M + \delta(1-\delta^N)\underline{v}_i + \delta^{N+1}v_i'$. Canceling out common terms, the comparison is between $((1-\delta^{N+1})/(1-\delta))v_i'$ and $M + ((1-\delta^N)/(1-\delta))\underline{v}_i$. The fractions approach $N+1$ and N as $\delta \to 1$. So for sufficiently large N and δ close enough to 1, the desired inequality will hold.
- Phase III_j: Complying gives $v_i' + \varepsilon$ forever, whereas deviating leads to a switch to phase II_i and so gives at most $(1-\delta)M + \delta(1-\delta^N)\underline{v}_i + \delta^{N+1}v_i'$. Again, for sufficiently large δ, complying is preferred.

Now we need to deal with the part where minmax strategies are mixed. For this we need to change the strategies so that, during phase II_j, player i is indifferent among all the possible sequences of N realizations of his prescribed mixed action. We accomplish this by choosing a different reward ε for each such sequence, so as to balance out their different short-term payoffs. We're not going to talk about this in detail; see the Fudenberg and Maskin paper for this.

□

22. Repeated games with fixed $\delta < 1$

The folk theorem shows that many payoffs are possible in SPE. But the construction of strategies in the proof is fairly complicated, since we have to have punishments and then rewards for punishers to induce them not to deviate. Also, the folk theorem is concerned with limits as $\delta \to 1$, whereas we may be interested in the set of equilibria for a particular value of $\delta < 1$.

We will now approach the question of identifying equilibrium payoffs for a given $\delta < 1$. In repeated games with perfect information, it turns out that an insight of Abreu (1988) will simplify the analysis greatly: equilibrium strategies can be enforced by using a worst possible punishment for any deviator. First we need to show that there is a well-defined worst possible punishment.

Theorem 18. *Suppose each player's action set in the stage game is compact and payoffs are continuous in actions, and some pure-strategy SPE of the repeated game exists. Then, among all pure-strategy SPEs, there is one that is worst for player i.*

That is, the infimum of player i's payoffs, across all pure-strategy SPEs, is attained.

Proof. We prove this for every player i simultaneously. Fix a sequence of equilibrium play paths (*not* strategy profiles) $s^{i,k}$, $k = 0, 1, 2, \ldots$ such that $U_i(s^{i,k})$ converges to the specified infimum $y(i)$. We want to define a limit of the strategy profiles, in such a way that the limiting profile is again an SPE with payoff $y(i)$ to player i.

Each strategy profile is an element of the strategy space $\prod_t A$, where A is the action space of the stage game and t ranges over all periods. By Tychonoff's theorem, this strategy space, with the product topology, is compact. Convergence in the product topology is defined componentwise — that is, $s^{i,k} \to s^{i,\infty}$ if and only if $s^{i,k}_t \to s^{i,\infty}_t$ for each t. Because the strategy space is compact, by passing to a subsequence if necessary, we can ensure that the $s^{i,k}$ have a limiting play path. It is easy to check that the resulting payoff to player i is $y(i)$.

Now we just have to check that this limiting play path $s^{i,\infty}$ is supportable as an SPE by some strategy profile. We construct the following profile: All players play according to $s^{i,\infty}$. If j deviates, switch to the strategy profile supporting $s^{j,\infty}$ constructed above.

Now, we need to check that the N strategy profiles constructed this way are really SPEs. But suppose we are in the strategy profile $s^{i,\infty}$ and j (who may or may not equal i) deviates at some period τ. His payoff from deviating is

$$(1-\delta)u_j(\widehat{a}_j, a^{i,\infty}_{-j}(\tau)) + \delta y(j).$$

We want to show that this is at most the continuation payoff,

$$(1-\delta)\sum_{t=0}^{\infty}\delta^t u_j(a^{i,\infty}(\tau+t)).$$

But we know that for each k, j does not have incentive to deviate in the SPE whose equilibrium play path is $s^{i,k}$; and by deviating his value in future periods is at least $y(j)$ (by definition of $y(j)$). So for each k we have

$$(1-\delta)\sum_{t=0}^{\infty}\delta^t u_j(a^{i,k}(\tau+t)) \geq u_j(\widehat{a}_j, a^{i,k}_{-j}(\tau)) + \delta y(j).$$

By taking limits at $k \to \infty$, we see that there is no incentive to deviate in the strategy profile supporting $s^{i,\infty}$, either.

This shows there are never incentives for a one-shot deviation. So by the one-shot deviation principle, we do have an SPE giving i his infimum of SPE payoffs, for any player i. □

Abreu refers to an SPE that gives i his worst possible payoff as an **optimal penal code.**

The above theorem applies when there exists a pure-strategy SPE. If the stage game is finite, there frequently will not be any pure-strategy SPE. In this case, there will be mixed-strategy SPE, and we would like to again prove that an optimal (mixed-strategy) penal code exists. A different method is required:

Theorem 19. *Consider an infinite-horizon repeated game where the stage game has finite action spaces. Each player i's strategy space is simply the countable product* $\prod_{h_t}\Delta(A_i)$, *taken over all possible finite histories h_t. Put the product topology on this space. Then, the set of SPE profiles is nonempty and compact.*

They show this by giving an equivalence between SPE of the infinitely repeated game and limits of near-equilibria of the T-period repeated game as $T \to \infty$. One can verify that the payoff functions are continuous on the space of strategy profiles of the repeated game; hence it follows that for each player i, there is some SPE that is worst for player i, that is, an optimal penal code.

Now we give a result applying the existence theorem. This result holds in either of the settings where an optimal penal code exists — either for pure strategies when the stage game has continuous action spaces (and some SPE exists), or for mixed strategies when the stage game is finite.

Theorem 20. *(Abreu, 1988) Suppose the stage game has compact action spaces and continuous payoffs. Then any distribution over outcomes achievable by an SPE can be generated by strategies that have optimal penal codes used off the equilibrium path, i.e. when i is the first to deviate, continuation play follows the optimal penal code for i.*

Proof. This is straightforward. Let $\widehat{s}$ be the given SPE. Form a new strategy profile s by leaving play on the equilibrium path as proposed by $\widehat{s}$, and replacing play off the equilibrium path by the optimal penal code for i when i is the first deviator (or one of the first deviators, if there is more than one). By the one-shot deviation principle, we need only check that i does not want to deviate when play so far is on the equilibrium path — but this is immediate, because i is punished with $y(i)$ in the continuation if he deviates, whereas in the original profile $\widehat{s}$ he would get at least $y(i)$ in the continuation (by definition of $y(i)$) and we know this was already low enough to deter deviation (because $\widehat{s}$ was an SPE). □

Now we look at Abreu (1986). Among other things, this paper looks at symmetric games and considers **strongly symmetric** equilibria — equilibria which have every player playing identically at every history, including asymmetric histories. This is a simple setting because everyone gets the same payoff, so there is one such equilibrium that is worst for everyone. One can similarly show that there is an equilibrium that is best for everyone. (Abreu does this in a setting where punishments can be made arbitrarily bad, so the good equilibrium can be sustained by punishments that last only one period. This is a very big simplifying assumption.) Then, describing the set of strongly symmetric equilibrium payoffs is simple — there are just two numbers, a worst and a best payoff, and we just have to write the incentive constraints relating the two, which makes computing these extremal equilibria fairly easy. Typically the best outcome is better than the best static Nash equilibrium.

23. Imperfect Public Monitoring

Next is Green and Porter (1984). This is the classic example of a repeated game with imperfect public monitoring: players only see a signal of other players' past actions, rather than observing the actions fully.

More specifically, each period there is a publicly observed signal y which follows some probability distribution conditional on the action profile a. Each player i's payoff is $r_i(a_i, y)$, something that depends only on his own action and the signal. His expected payoff from a strategy profile is then $u_i(a) = \sum_{y \in Y} r_i(a_i, y)p(y|a)$. In the Green-Porter model, each player is a firm in a cartel that sets a production quantity. Quantities are only privately observed. There is also a market price, which is publicly observed and depends stochastically on the players' quantity choices (thus there is a demand shock each period). Each firm's payoff is the product of the market price and its quantity as usual. So the firms are trying to collude by keeping quantities low and prices high, but in any given period prices may be low, and each firm doesn't know if prices are low because of a low shock or because some other firm deviated and produced a high quantity. In particular, Green and Porter assume that the support of the price signal y does not depend on the action profile played, which ensures that a low price may occur even when no firm has deviated.

Green and Porter did not try to solve for all equilibria of their model. Instead they simply discussed the idea of threshold equilibria: everyone plays the collusive action profile $\hat{a}$ for a while; if the price y is ever observed to be below some threshold $\hat{y}$, revert to static Nash for some number of periods T, and then return to the collusion phase. (Note: this is not pushing the limits of what is feasible, since, for example, Abreu's work implies that there are worse punishments possible than just reverting to static Nash.)

They define $\lambda(\hat{a}) = P(y > \hat{y}|\hat{a})$, the probability of seeing a high price when there was no deviation. Equilibrium values are then given by

$$\hat{v} = (1-\delta)u_i(\hat{a}) + \delta\lambda(\hat{a})\hat{v} + \delta(1-\lambda(\hat{a}))\delta^T\hat{v}$$

(writing 0 for the static Nash payoffs). This lets us calculate $\hat{v}$ for any proposed $\hat{a}$ and T. These strategies form an equilibrium only if no player wants to deviate in the collusive phase:

$$u_i(a_i', \hat{a}_{-i}) - u_i(\hat{a}) \leq \delta(1-\delta^T)u_i(\hat{a})(\lambda(\hat{a}) - \lambda(a_i', \hat{a}_{-i}))$$

for all possible deviations a_i'. This compares the short-term incentives to deviate, the relative probability that deviation will trigger a reversion to static Nash, and the severity of the punishment.

Green and Porter showed that it is possible to sustain payoffs at least slightly above static Nash with strategies of this sort. As already remarked, they did not find the best possible equilibria.

Now we return to the setting of general repeated games with imperfect public monitoring; the notation is as laid out at the beginning of this section. We will present the theory of these games as developed by Abreu, Pearce, and Stacchetti (1990) (hereafter referred to as APS).

For convenience we will assume that the action spaces A_i and the space Y of possible signals are finite. We will also write $\pi_y(a)$ for the probability distribution over y given action profile a (previously notated $p(y|a)$). It is clear how to generalize this to the distribution $\pi_y(\alpha)$ where α is a mixed action profile.

If there were just one period, players would just be playing the normal-form game with action sets A_i and payoffs $u_i(a) = \sum_{y \in Y} \pi_y(a) r_i(a_i, y)$. With repetition, this is no longer the case since play can be conditioned on the history — but may not be able to be conditioned exactly on past actions of opponents, as in the earlier, perfect-monitoring setting, because players do not see their opponents' actions.

Notice that the perfect monitoring setting can be embedded into this framework, by simply letting $Y = A$ be the space of action profiles, and y be the action profile actually played with probability 1. We can also embed "noisy" repeated games with perfect monitoring, where each agent tries to play a particular action a_i in each period but ends up playing any other action a_i' with some small probability ε; each player can only observe the action profile actually played, rather than the actions that the opponents "tried" to play.

In such a game, at any time t, player i's information is given by his private history

$$h_i^t = (y^0, \ldots, y^{t-1}; a_i^0, \ldots, a_i^{t-1}).$$

That is, he knows the history of public signals and his own actions (but not others' actions). He can condition his action in the present period on this information. The **public history** $h^t = (y^0, \ldots, y^{t-1})$ is commonly known.

APS restrict attention to pure strategies, which is a nontrivial restriction.

A strategy σ_i for player i is a **public strategy** if $\sigma_i(h_i^t)$ depends only on the history of public signals $y^0, \ldots, y^{t-1}$.

Lemma 1. *Every pure strategy is equivalent to a public strategy.*

Proof. Let σ_i be a pure strategy. Define a public strategy σ_i' on length-t histories by induction: $\sigma_i'(y^0, \ldots, y^{t-1}) = \sigma_i(y^0, \ldots, y^{t-1}; a_i^0, \ldots, a_i^{t-1})$ where $a_i^s = \sigma_i'(y^0, \ldots, y^{s-1})$ for each $s < t$. That is, at each period, i plays the actions specified by σ_i for the given public signals and the history of private actions that i was supposed to play. It is straightforward to check that σ_i' is equivalent to σ_i, since they differ only at "off-path" histories reachable only by deviations of player i. □

This shows that if attention is restricted to pure strategies, it is no further loss to restrict in turn to public strategies. However, instead of doing this, we will follow the exposition of Fudenberg and Tirole and restrict to public (but potentially mixed) strategies.

Lemma 2. *If i's opponents use public strategies, then i has a best reply in public strategies.*

Proof. This is straightforward — i always knows what the other players will play, since their actions depend only on the public history; hence i can just play a best response to their anticipated future play, which does not depend on i's private history of past actions. □

This allows us to define a **public perfect equilibrium** (PPE): a profile $\sigma = (\sigma_i)$ of public strategies such that, at every public history $h^t = (y^0, \ldots, y^{t-1})$, the strategies $\sigma_i|_{h^t}$ form a Nash equilibrium of the continuation game.

(This is the straightforward adaptation of the concept of subgame-perfect equilibrium to our setting. Notice that we cannot simply use subgame-perfect equilibrium because it has no bite in general — there are no subgames.)

The set of PPE's is stationary — they are the same at every history. This is why we look at PPE. Sequential equilibrium does not share this stationarity property, because a player may want to play differently in one period depending on the realization of his mixing in a previous period. Such correlation across periods can be self-sustaining in equilibrium: if i and j both mixed at a previous period s, then the signal in that period gives i information about the realization of j's mixing, which means it is informative about what j will do in the current period, and therefore affects i's current best reply. Consequently, different players can have different information at time t about what will be played at time t, and stationarity is destroyed. We stick to public equilibria in order to avoid this difficulty.

Importantly, the one-shot deviation principle applies to our setting. That is, a set of public strategies constitutes a PPE if and only if there is no beneficial one-shot deviation for any player.

Let $w : Y \to \mathbb{R}^n$ be a function; $w_i(y)$ denotes the continuation payoff player i expects to get when the signal y is realized, from some strategy profile. This gives rise to the following definition:

Definition 13. *A pair consisting of a (mixed) action profile α and payoff vector $v \in \mathbb{R}^n$ is **enforceable** with respect to $W \subseteq \mathbb{R}^n$ if there exists $w : Y \to W$ such that*

$$v_i = (1-\delta)u_i(\alpha) + \delta \sum_{y \in Y} \pi_y(\alpha) w_i(y)$$

and

$$v_i \geq (1-\delta)u_i(a_i', \alpha_{-i}) + \delta \sum_{y \in Y} \pi_y(a_i', \alpha_{-i}) w_i(y)$$

for all $a_i' \in A_i$. (Here u_i denotes the expected stage payoff function defined above.)

The idea of enforceability is that it is incentive-compatible for each player to play according to α in the present period if continuation payoffs are given by w, and the resulting (expected) payoffs starting from the present period are given by v.

Let $B(W)$ be the set of all v that are enforceable with respect to W, for some action profile α. This is the set of payoffs **generated** by W.

Theorem 21. *Let E be the set of payoff vectors that are achieved by some PPE. Then $E = B(E)$.*

Proof. For any $v \in E$ generated by some strategy profile σ, let $\alpha_i = \sigma_i(\emptyset)$ and $w_i(y)$ be the expected continuation payoff of player i in subsequent periods given that y is the realized signal. Since play in subsequent periods again forms a PPE, $w(y) \in E$ for each signal realization y. Then (α, v) is enforced by w on E — this is exactly the statement that v represents the overall expected payoffs and the players are willing to play according to α in the first period. So $v \in B(E)$.

Conversely, if $v \in B(E)$, let (α, v) be enforced by w on E. Consider the strategies defined as follows: play α in the first period, and whatever signal y is observed, play in subsequent periods follows a PPE with payoffs $w(y)$. These strategies form a PPE, by the one-shot deviation principle: enforcement means that there is no incentive to deviate in the first period, and the fact that continuation play is given by a PPE ensures that there is no incentive to deviate in any subsequent period. Finally it is straightforward from the definition of enforcement that the payoffs are in fact given by v. Thus $v \in E$. □

Definition 14. *$W \subseteq \mathbb{R}^n$ is* ***self-generating*** *if $W \subseteq B(W)$.*

Thus, we have just shown that E, the set of equilibrium payoffs, is self-generating.

Theorem 22. *If W is a bounded, self-generating set, then $W \subseteq E$.*

This theorem shows that E is actually the largest bounded self-generating set.

Proof. Let $v \in W$. We want to construct a PPE with payoffs given by v. We construct the strategies iteratively. Suppose we have specified play for periods $0, \ldots, t-1$. We want to specify how players should play at period t. We do this by simultaneously specifying the continuation payoffs players should receive at each public history beginning in period $t+1$. (The base case, $t = 0$, has players receiving continuation payoffs given by v.)

Suppose the history of public signals so far is $y^0, \ldots, y^{t-1}$ and promised continuation payoffs are given by $v' \in W$. Because W is self-generating, there is some action profile α and some $w : Y \to W$ such that (α, v) is enforced by w. Specify that the players play α at this history, and whatever signal y is observed, their continuation payoffs starting from the next period should be $w(y)$.

These strategies form a PPE — this is easily checked using the one-shot deviation principle; enforcement means exactly that there are no incentives to deviate in any period. Finally, the expected payoffs at time 0 are just given by v; this follows from adding up (and using boundedness to make sure that the series of payoffs coverges appropriately). □

In addition to having this characterization of the set of PPE payoffs, Abreu, Pearce, and Stacchetti also show a monotonicity property with respect to the discount factor. Let $E(\delta)$ be the set of PPE payoffs when the discount factor is δ. Suppose that $E(\delta)$ is convex: this can be achieved either by incorporating public randomization into the model, or by having a sufficiently rich space of public signals (we can't do this in our model because Y is finite). Then if $\delta_1 < \delta_2$ they have $E(\delta_1) \subseteq B(E(\delta_1), \delta_2)$, and therefore, by the previous theorem, $E(\delta_1) \subseteq E(\delta_2)$. This is shown by the following approach: given $v \in E(\delta_1) = B(E(\delta_1), \delta_1)$, find α and w that enforce v when the discount factor is δ_1; by replacing w by a suitable convex combination of w and a constant function, we can enforce (α, v) when the discount factor is δ_2.

Some other facts about the B operator:

- If W is compact, so is $B(W)$. This is shown by a straightforward (boring) topological argument.
- B is monotone: if $W \subseteq W'$ then $B(W) \subseteq B(W')$. Also easy to see.
- If W is nonempty, so is $B(W)$. To see this, just let α be a Nash equilibrium of the stage game and v the resulting payoffs.

Now let V be the set of all feasible payoffs, which is certainly compact. Think about the sequence of iterates $B^0(V), B^1(V), \ldots$, where $B^0(v) = V$ and $B^k(V) = B(B^{k-1}(V))$. By induction, these sets are all compact and they form a decreasing sequence. Hence, their intersection is compact. If we let $B^\infty(V)$ denote this intersection, then $B^\infty(V) = B(B^\infty(V))$, and hence $B^\infty(V) \subseteq E$. On the other hand, $E \subseteq V$, and hence by induction E is contained in each term of the sequence. Therefore $E \subseteq B^\infty(V)$. In conclusion:

Theorem 23. *$E = B^\infty(V)$.*

This is the main theorem of APS. It gives a characterization of the set of PPE payoffs: start with the set of all feasible payoffs, and apply the operator B repeatedly; the resulting sequence of sets converges to the set of equilibrium payoffs.

Corollary 3. *The set of PPE payoff vectors is nonempty and compact.*

(Nonemptiness is immediate because, for example, the infinite repetition of any static NE is a PPE.)

APS also show a "bang-bang" property of public perfect equilibria. We say that $w : Y \to W$ has the bang-bang property if $w(y)$ is an extreme point of W for each y. Under appropriate assumptions on the signal structure, they show that if (α, v) is enforceable on a compact W, it is in fact enforceable on the set $ext(W)$ of extreme points of W. Consequently, every vector in E can be achieved as the vector of payoffs from a PPE such that the vector of continuation payoffs at every history lies in $ext(E)$.

Fudenberg, Levine, and Maskin (1994) (hereafter FLM) show a folk theorem for repeated games with imperfect public monitoring. They find conditions on the game under which they can find a convex set W with a smoothly curved boundary, approximating the set of feasible, individually rational payoffs arbitrarily closely; then they can show that W is self-generating for a sufficiently high discount factor. This implies that a folk theorem obtains.

We will briefly discuss what the technical difficulties are in the course of proving this. First, there has to be identifiability of each player's actions. If player i's deviation to a_i' generates exactly the same distribution over signals as some a_i he is supposed to play (given opponents' play α_{-i}), but gives him a higher payoff on average, then clearly there is no way to enforce the action profile a in equilibrium. The same problem may arise if a_i is not effectively imitated by playing another pure action a_i', but it is imitated by playing some mixture of other actions. To avoid this problem, FLM assume a full-rank condition: given α_{-i}, the different signal distributions generated by varying i's action a_i are linearly independent.

They need to further assume a "pairwise full rank" condition: deviations by player i are statistically distinguishable from deviations by player j. Intuitively this is necessary because, if the signal suggests that someone has deviated, the players need to know who to punish. (Radner, Myerson, and Maskin, 1986, give an example of a game that violates this condition and where the folk theorem does not hold. There are two workers who put in effort to increase the probability that a project succeeds; they both get 1 if it succeeds and 0 otherwise. The outcome of the project does not statistically distinguish between shirking by player 1 and shirking by player 2. So if the project fails, both players have to be punished by giving them lower continuation payoffs than if it succeeds. Because it fails some of the time even if both players are working, this means that equilibrium payoffs are bounded away from efficiency, even as $\delta \to 1$.)

The statement of the pairwise full rank condition is as follows: given the action profile α, if we form one matrix whose rows represent the signal distributions from (a_i, α_{-i}) as a_i varies over A_i, and another matrix whose rows represent the signal distributions from (a_j, α_{-j}) as a_j varies over A_j, and stack these two matrices, the combined matrix has rank $|A_i| + |A_j| - 1$. (This is effectively "full rank" — it is not possible to have literal full rank $|A_i| + |A_j|$, since the signal distribution generated by α is a linear combination of the rows of the first matrix and is also a linear combination of the rows of the second matrix.)

When this condition is satisfied, it is possible to use continuation payoffs to transfer utility between the two players i, j in any desired ratio, depending on the signal, so as to incentivize i and j to play according to the desired action profile.

FLM show that the W they construct is **locally self-generating**: for every $v \in W$, there is an open neighborhood U and a $\underline{\delta} < 1$ such that $U \cap W \subseteq B(W)$ when $\delta > \underline{\delta}$. This definition allows $\underline{\delta}$ to vary with v. For W compact and convex, they show that local self-generation implies self-generation for all sufficiently high δ.

The intuition behind their approach is best grasped with a picture. Suppose we want to achieve some payoff vector v on the boundary of W. The full-rank conditions ensure we can enforce it using some continuation payoffs that lie below the tangent hyperplane to W at v, by "transferring" continuation utility between players as described above. As $\delta \to 1$, the continuation payoffs sufficient to enforce v contract toward v, and the smoothness condition on the boundary of W ensures that they will eventually lie inside W. Thus (α, v) is enforced on W.

[PICTURE — See p. 1013 of Fudenberg, Levine, Maskin (1994)]

Some extra work is needed to take care of the points v where the tangent hyperplane is a coordinate hyperplane (i.e. one player's payoff is constant on this hyperplane).

An argument along these lines shows that every vector on the boundary of W is achievable using continuation payoffs in W, when δ is high enough. Using public randomization among boundary points, we can then achieve any payoff vector v in the interior of W as well. It follows that W is self-generating (for high δ).

24. Reputation

The earliest repeated-game models of reputation were by the Gang of Four (Kreps, Milgrom, Roberts, and Wilson); in various combinations they wrote three papers that were simultaneously published in JET 1982.

The motivating example was the "chain-store paradox." In the chain-store game, there are two players, an entering firm and an incumbent monopolist. The entrant (player 1) can enter or stay out; if it enters, the incumbent (player 2) can fight or not. If the entrant stays out, payoffs are $(0, a)$ where $a > 1$. If the entrant enters and the incumbent does not fight, the payoffs are $(b, 0)$ where $b \in (0, 1)$. If they do fight, payoffs are $(b - 1, -1)$. There is a unique SPE in which the entrant enters and the incumbent does not fight.

In reality, incumbent firms seem to fight when a rival enters, and thereby deter other potential rivals. Why would they do this? In a one-shot game, it is irrational for the incumbent to fight the entrant. As pointed out by Selten, even if the game is repeated finitely many times, the unique SPE still has the property that there is entry and accommodation in every period, by backward induction.

The Kreps-Wilson explanation for entry deterrence is as follows: with some small positive probability, the monopolist does not have the payoffs described above, but rather is obsessed with fighting and has payoffs such that it always chooses to fight. Then, when there are a large number of periods, they show that there is no entry for most of the game, with entry occurring only in the last few periods.

Their analysis is tedious, so we will instead begin with a simpler example: the centipede game. Initially both players have 1 dollar. Player 1 can end the game (giving payoffs $(1, 1)$), or he can give up 1 dollar for player 2 to get 2 dollars. Player 2 can then end the game (giving $(0, 3)$), or can give up 1 for player 1 to get 2. Player 1 can then end the game (with payoffs $(2, 2)$), or can give up 1 for player 2 to get 2. And so forth — until the payoffs reach $(100, 100)$, at which point the game is forced to end. (See the illustration on Muhamet's slides. We will refer to continuing the game as "playing across" and ending as "playing down," due to the shape of the centipede diagram.)

There is a unique SPE in this game, in which case both players play down at every opportunity. But believing in SPE requires us to hold very strong assumptions about the players' higher-order knowledge of each other's rationality.

Suppose instead that player 1 has two types. With probability 0.999, he is a "normal" type and his payoffs are above. With probability 0.001, he is a "crazy" type who always gets utility -1 if he ends the game and 0 if player 2 ends the game. (Player 2's payoffs are the same regardless of 1's type.) The crazy type of player 1 thus always wants to continue the game. Player 2 never observes player 1's type.

What happens in equilibrium? Initially player 1 has a low probability of being the crazy type. If the normal player 1 plays down at some information set, and the crazy player 1 does not, then after 1 plays across, player 2 must infer that he is crazy. But if player 1 is crazy then he will continue the game until the end; knowing this, player 2 also wants to play across. Anticipating this, the normal type of player 1 in turn also wants to play across in order to get a high payoff.

With this intuition laid out, let's analyze the game formally and describe all the sequential equilibria. Number the periods, starting from the end, with 1 being player 2's last information set, 2 being player 1's previous information set, ..., 198 being 1's first information set. It is easy to see that the crazy player 1 always plays across.

Player 2 always plays across with positive probability at every period $n > 1$. (Proof: if not, then the normal player 1 must play down at period $n+1$. Then, conditional on reaching n, player 2 knows that 1 is crazy with probability 1, hence he would rather go across and continue the game to the end.)

Hence there is positive probability of going across at every period, so the beliefs are uniquely determined from the equilibrium strategies by Bayes's rule.

Next we see that the normal player 1 plays across with positive probability at every $n > 2$. Proof: if not, then again, at $n-1$ player 2 is sure that he is facing a crazy type and therefore wants to go across. Given this strategy by player 2, then, the normal 1 also has incentives to go across at n so that he can go down at $n-2$, contradicting the assumption that 1 only goes down at n.

Next, if 2 goes across with probability 1 at n, then 1 goes across with probability 1 at $n+1$, and this in turn implies that 2 goes across with probability 1 at $n+2$. This is also seen by the same argument as in the previous paragraph. Therefore there is some cutoff $n^* \geq 3$ such that both players play across with probability 1 at $n > n^*$, and there is mixing for $2 < n \leq n^*$. (We know that both the normal 1 and 2 play down with probability 1 at $n = 1, 2$.)

Let q^n be the probability of player 2 going down at node n, if n is odd; let p^n be the probability of player 1 going down at n, if n is even. Let μ_n be the probability player 2 assigns to the crazy type at node n.

At each odd node n, $2 < n \leq n^*$, player 2 is to be indifferent between going across and down. The payoff to going down is some x. The payoff to going across is $(1-\mu_n)p_{n-1}(x-1) + [1-(1-\mu_n)p_{n-1}](x+1)$, using the fact that player 2 is again indifferent (or strictly prefers going down) two nodes later. Hence we get $(1-\mu_n)p_{n-1} = 1/2$: player 2 expects player 1 to play down with probability 1/2. But $\mu_{n-2} = \mu_n/(\mu_n + (1-\mu_n)(1-p_{n-1}))$ by Bayes's rule; this simplifies to $\mu_{n-2} = \mu_n/(1-(1-\mu_n)p_{n-1}) = 2\mu_n$. We already know that $\mu_1 = 1$ since the normal player 1 goes down with certainty at node 2. Therefore $\mu_3 = 1/2$, $\mu_5 = 1/4$, and so forth; and in particular $n^* \leq 20$, since otherwise $\mu_{21} = 1/1024 < 0.001$, but clearly the posterior probability of the crazy type at any node cannot be lower than the prior. This shows that for all but the last 20 periods, both players are going across with probability 1 in equilibrium.

(One can in fact continue to solve for the complete description of the sequential equilibrium: now that we know player 2's posterior at each period, we can compute player 1's mixing probabilities from Bayes's rule, and we can also compute player 2's mixing probabilities given that 1 must be indifferent whenever he mixes. But we've already gotten the punch line of this model.)

The papers by the Gang of Four consider repeated interactions between the same players, with one-sided incomplete information. Inspired by this work, Fudenberg and Levine (1989) consider a model in which a long-run player faces a series of short-run players, and where there are many possible "crazy" types of the long-run player, each with small positive probability. They show that if the long-run player is sufficiently patient, he will get close to his Stackelberg payoff in any *Nash* equilibrium of the repeated game.

Let's lay out the model. There are two players, playing the normal-form game (N, A, u) (with $N = \{1, 2\}$) in each period. Player 1 is a long-run player. Player 2 is a short-run player (which we can think of as a series of players who play for one period each, or one very impatient player), who plays a best reply to player 1's anticipated action in each stage.

Let

$$u_1^* = \max_{a_1 \in A_1} \min_{\sigma_2 \in BR_2(a_1)} u_1(a_1, \sigma^2).$$

This is player 1's **Stackelberg payoff**; the action a_1^* that achieves this maximum is the **Stackelberg action**. The main paper of Fudenberg and Levine only allows a_1 to be a pure action; in a follow-up paper three years later they allow for mixed actions by player 1, which is more complicated.

A strategy for player 1 consists of a function $\sigma_1^t : H^{t-1} \to \Delta(A_1)$ for each $t \geq 0$. A strategy for the player 2 who plays at time t consists of a function $\sigma_2^t : H^{t-1} \to \Delta(A_2)$. With the usual payoff formulation, we have the **unperturbed game**. Fudenberg, Kreps, and Maskin (1988) proved a version of the folk theorem for this game. Let $\underline{u}_1$ be the minimum payoff that player 1 can get given that player 2 is playing a best reply. Fudenberg, Kreps, and Maskin show that any payoff above $\underline{u}_1$ can be sustained in SPE for high enough δ. The main reputation result of Fudenberg and Levine shows that if there is a rich space of crazy types of player 1, each with positive probability, this folk theorem is completely overturned — player 1 is guaranteed to get close to u_1^* (or more) in any Nash equilibrium for high δ. We don't even have to impose subgame-perfection.

Accordingly, we consider the perturbed game, where there is a countable state space Ω. Player 1's payoff depends on the state $\omega \in \Omega$; thus write $u_1(a_1, a_2, \omega)$. Player 2's payoff does not depend on ω. There is some common prior μ on Ω, but the true state is known only to player 1. When the state is $\omega_0 \in \Omega$, player 1's payoffs are given by the original u_1; we call this the "rational" type of player 1.

Suppose that for every $a_1 \in A_1$, there is a state $\omega(a_1)$ for which playing a_1 at every history is a strictly dominant strategy in the *repeated* game. (Assuming it is strictly dominant in the stage game is not enough.) Thus, at state $\omega(a_1)$, player 1 is guaranteed to play a_1 at every history. Write $\omega^* = \omega(a_1^*)$. We assume also that the probability $\mu^* = \mu(\omega^*)$ is strictly positive. That is, with positive probability, player 1 is a type who is guaranteed to play a_1^* in every period.

Any strategy profile will lead to a joint probability distribution π over play paths and states, $\pi \in \Delta((A_1 \times A_2)^\infty \times \Omega)$. Let h^* be the event (in this path-state space) that $a_1^t = a_1^*$ for all t. Let $\pi_t^* = \pi(a_1^t = a_1^* | h_{t-1})$, the probability of seeing a_1^* at period t given the previous history; this is a random variable (defined on path-state space) whose value depends on h_{t-1}. For any number $\overline{\pi} \in (0,1)$, let $n(\pi_t^* \leq \overline{\pi})$ denote the number of periods t such that $\pi_t^* \leq \overline{\pi}$. This is again a random variable, whose value may be infinite.

The crucial lemma is the following:

Lemma 3. *Let σ be a strategy profile such that $\pi(h^*|\omega^*) = 1$. Then*

$$\pi\left(n(\pi_t^* \leq \overline{\pi}) > \frac{\ln(\mu^*)}{\ln \overline{\pi}} \;\middle|\; h^*\right) = 0.$$

That is, conditional on the play path being one where a_1^* is seen in every period, there are guaranteed (almost surely) to be at most $\ln(\mu^*)/\ln \overline{\pi}$ periods in which the probability of a_1^* at the next period, given the previous history, is at most $\overline{\pi}$.

The proof is straightforward. Given that $\pi(h^*|\omega^*) = 1$, if the true state is ω^*, then player 1 will always play a_1^*. Each time the probability of seeing a_1^* next period is less than $\overline{\pi}$, if a_1^* is in fact played, the posterior probability of ω^* must increase by a factor of at least $1/\overline{\pi}$. The posterior probability starts out at μ^*, and it can never exceed 1, so it can increase no more than $\ln(\mu^*)/\ln(\overline{\pi})$ times.

Formally, consider any finite history h^t at which a_1^* has been played every period, and such that $\pi(h^t) > 0$. Then we actually have by Bayes's rule:

$$\pi(\omega^*|h^t) = \frac{\pi(\omega^*|h^{t-1})\pi(h^t|\omega^*, h^{t-1})}{\pi(h^t|h^{t-1})} = \frac{\pi(\omega^*|h^{t-1})}{\pi(h^t|h^{t-1})}.$$

(Here the second equality holds because if ω^* occurs then $a_1^t = a_1^*$.) Repeatedly expanding, we have

$$\pi(\omega^*|h^t) = \frac{\pi(\omega^*|h^0)}{\pi(h^t|h^{t-1})\pi(h^{t-1}|h^{t-2}) \cdots \pi(h^1|h^0)}.$$

But the numerator of the right-hand side is exactly μ^* while the left-hand side is at most 1. So at most $\ln(\mu^*)/\ln(\overline{\pi})$ of the denominator terms can be less than or equal to $\overline{\pi}$. Since the denominator term $\pi(h^s|h^{s-1})$ is at most π_s^*, the result follows: we cannot, with positive probability, see a history at which $n(\pi_t^* < \overline{\pi}) > \ln \mu^* / \ln \overline{\pi}$.

This lemma can be read as saying "a_1^* can be a surprise only so many times."

Now we get to the main theorem. Let $u_m = \min_a u_1(a, \omega_0)$, the worst possible stage payoff for player 1. Denote by $\underline{v}_1(\delta, \mu, \omega_0)$ the infimum across all Nash equilibria of the rational player 1's payoffs in the repeated game, for given discount factor δ and prior μ.

Theorem 24. *For any value μ^*, there exists a number $\kappa(\mu^*)$ with the following property: for all δ and all (μ, Ω) with $\mu(\omega^*) = \mu^*$, we have*

$$\underline{v}_1(\delta, \mu, \omega_0) \geq \delta^{\kappa(\mu^*)} u_1^* + (1 - \delta^{\kappa(\mu^*)}) u_m.$$

As $\delta \to 1$, then, this payoff bound converges to u_1^*.

Proof. First, we will show that there exists a $\overline{\pi} < 1$ such that, in every play path of every Nash equilibrium, at every stage t where $\pi_t^* > \overline{\pi}$, player 2 plays a best reply to a_1^*. This is straightforward: if it is not true, there must be a sequence of mixed stage-game actions of player 1, converging to a_1^*, such that for each of them, player 2 has a best reply that is not in $BR_2(a_1^*)$. By finiteness, some best reply a_2 of player 2 occurs infinitely often. But then the theorem of the maximum implies this a_2 is a best reply to a_1^*, a contradiction. So the desired $\overline{\pi}$ exists.

Thus, by the lemma, we have a number $\kappa(\mu^*)$ of periods such that $\pi(n(\pi^* \leq \overline{\pi}) > \kappa(\mu^*) \mid h^*) = 0$. Now, whatever player 2's equilibrium strategy is, if the rational player 1 deviates to simply playing a_1^* every period, there are at most $\kappa(\mu^*)$ periods in which player 2 will not play a best reply to a_1^* — since player 2 is playing a best reply to player 1's expected play in each period. Thus the rational player 1 gets a stage payoff of at least u_m in each of these periods, and least u_1^* in all the other periods. This immediately gives that player 1's payoff from deviating is at least $\delta^{\kappa(\mu^*)} u_1^* + (1 - \delta^{\kappa(\mu^*)}) u_m$. Since we have a Nash equilibrium, player 1's payoff in equilibrium is at least his payoff from deviating. The theorem follows. □

Notice that the payoff u_1^* we have looked at is a "lower Stackelberg payoff." There is also an "upper Stackelberg payoff" in the stage game,

$$\overline{u}_1 = \max_{a_1} \max_{\sigma_2 \in BR_2(a_1)} u_1(a_1, a_2),$$

and evidently player 1 cannot get more than the upper Stackelberg payoff (if he is playing a pure strategy; otherwise the outer max should be taken over mixed strategies). In generic games, the lower and upper Stackelberg payoffs coincide, and so we get a unique equilibrium payoff for the normal player 1 in the limit as $\delta \to 1$.

25. Reputation and Bargaining

Abreu and Gul (2000) consider reputation in the context of bargaining. Suppose there are two players $1, 2$ who have to split a pie of total size 1. Each player i can be either rational or a crazy type who just always demands a share α^i of the pie.

Abreu and Gul consider very general bargaining protocols; the details of the protocol turn out not to make a difference. There is a function $g : [0, \infty) \to \{0, 1, 2, 3\}$; $g(t)$ indicates what happens at time t. If $g(t) = 0$ nothing happens; if $g(t) = 1$ then player 1 makes an offer and player 2 can accept or reject; if $g(t) = 2$ then player 2 makes an offer and player 1 can accept or reject; if $g(t) = 3$ then both players simultaneously offer, and if the offers are compatible (the shares the players claim sum to at most 1) then they get their shares and the surplus is split equally, otherwise the bargaining continues.

Let $T_i = \{t \mid g(t) \in \{i, 3\}\}$, the set of times when i can make an offer. They consider a protocol such that each T_i is finite. They then consider a continuous bargaining limit by looking at a sequence g^n of finite protocols such that for any ε and t, for all sufficiently large n, each player gets to offer between t and $t + \varepsilon$.

For any such protocol, whenever either player i has revealed himself to be rational by doing anything other than demanding α^i, there will be almost immediate agreement, and j can get himself a share close to α^j by continuing to use his reputation. This is similar to the Fudenberg-Levine reputation result, but it turns out to be complicated to prove. So what happens in equilibrium if both players are rational? They play a war of attrition — each player pretends to be irrational but has some probability of conceding at each period (by revealing rationality), and as soon as one concedes the ensuing payoffs are those given by the reputation story. These concession probabilities must make each player indifferent between conceding and not; from this we can show that the probabilities are stationary, up to some finite time, and then both players simply fail to concede after that time even if they are rational.

With that overview, let's specify the model in detail. There are two players $1, 2$. Player i has discount rate r_i. If an agreement $(x, 1 - x)$ is reached at time t, the payoffs (if the players are rational) are $(x_1 e^{-r_1 t}, x_2 e^{-r_2 t})$. Each player i, in additional to his rational type, has an irrational type, whose behavior is fixed: this type always demands α_i, and always accepts offers that give him at least α_i and rejects lower offers. We assume $\alpha_1 + \alpha_2 > 1$. The probability that player i is irrational is z_i.

We consider bargaining protocols that are a generalization of the Rubinstein alternating-offers protocol. A protocol is given by a function $g : [0, \infty) \to \{0, 1, 2, 3\}$. If $g(t) = 0$, then nothing happens at time t. If $g(t) = 1$ then player 1 makes an offer, and 2 immediately decides whether to accept or reject. If $g(t) = 2$ then the same happens with players 1 and 2 reversed. If $g(t) = 3$ then both players simultaneously offer. If their offers are incompatible (the amount player 1 demands plus the amount player 2 demands exceeds 1) then both offers are rejected and the game continues; otherwise each player gets what he demands and the remaining surplus is split equally.

The protocol is discrete, meaning that for every t, $g^{-1}(\{1, 2, 3\}) \cap [0, t)$ is finite. A sequence of such protocols (g_n) **converges to the continuous limit** if, for all $\varepsilon > 0$, there exists n^* such that for all $n > n^*$, and for all t, $\{1, 2\} \subseteq g_n([t, t + \varepsilon])$. For example, this is satisfied if g_n is the Rubinstein alternating protocol with time increments of $1/n$ between offers. As Abreu and Gul show, each g_n induces a game with a unique equilibrium outcome, and these equilibria converge to the unique equilibrium outcome of the continuous-time limit game.

To make sense of this, we need a description of the continuous-time limit game. This game is a war of attrition: Each player initially demands α_i. At any time, each player can concede or not. Thus, rational player i's strategy is a probability distribution over times $t \in [0, \infty]$ at which to concede (given that j has not already conceded). ($t = \infty$ corresponds to never conceding.) When player i concedes at time t, the payoffs are $(1 - \alpha_j) e^{-r_i t}$ for i and $\alpha_j e^{-r_j t}$ for j. With probability z_i, player i is the irrational type who never concedes. (If there is no concession, both players get payoff 0.)

Without going through all the results in detail, we will sketch the relationship between these bargaining games and the reputation machinery we have developed, and will outline the analysis of the continuous-time game.

Abreu and Gul show that in the discrete games, once player i has conceded, there must be agreement in equilibrium, and it is almost immediate. More precisely, they show (their Lemma 1):

Lemma 4. *For any $\varepsilon > 0$, if n is sufficiently high, then after any history in g_n where i has revealed rationality and j has not, in equilibrium play of the continuation game, i obtains at most $1 - \alpha_j + \varepsilon$ and j obtains at least $\alpha_j - \varepsilon$.*

Proof. Consider the equilibrium continuation play starting from some history at which i has revealed rationality and j has not as of time t. Let $\hat{t}$ be any time increment such that, with positive probability (in this continuation), the game still has not ended at time $t + \hat{t}$. We will first show that all there is an upper bound on $\hat{t}$.

Let π be the probability that j does not reveal rationality in $[t, t + \hat{t})$. Then i's expected payoff as of time t satisfies $v_i \leq 1 - \pi + \pi e^{-r_i \hat{t}}$. We also have $v_i \geq (1 - \alpha_j) z_j^t$ where z_j^t is the posterior that j is irrational as of time t. Combining,

$$1 - \pi + \pi e^{-r_i \hat{t}} \geq (1 - \alpha_j) z_j^t \geq (1 - \alpha_j) z_j.$$

By taking $\hat{t} \to \infty$ (for which π must be increasing) we can see that π is bounded above by some $\bar{\pi} < 1$, for large enough $\hat{t}$.

Now we apply the reasoning from Fudenberg and Levine (1989). Assume $\hat{t}$ is large enough that j always has a chance to offer in any interval of length $\hat{t}$. Each time an interval of length $\hat{t}$ goes by without j conceding, the posterior probability that j is irrational increases by a factor of at least $1/(1 - \bar{\pi}) > 1$. The number of such increases that can occur is bounded above (by $-\ln(1 - \bar{\pi})$). Thus there is an upper bound on the amount of time the game can continue, as claimed.

Next, by a refinement of this argument, one shows that there exists $\beta \in (0, 1)$ and $\zeta < 1$ with the following property: for any sufficiently small increment ε, given that i has revealed rationality and j has not *and* that the maximum length of time the game can continue in equilibrium if j continues not conceding is ε, then the probability that j will still not have revealed rationality within time $\beta\varepsilon$ (if he has the chance to do so) is at most ζ. (We omit the details of this argument.)

So suppose we have a history at time t where i has revealed rationality, j has not, and the latest possible end of the game (if j continues not conceding) is $t + \varepsilon$. If j has had the chance to reveal rationality by time $t + \beta\varepsilon$, which is $(1 - \beta)\varepsilon$ before the end of the game, and has not done so, the posterior probability that j is irrational must increase by at least a factor $1/\zeta > 1$. Then, if j has had another chance to reveal rationality by the time $(1 - \beta)^2\varepsilon$

before the end of the game, and has not done so, the posterior probability of irrationality must increase by another factor of $1/\zeta$. And so forth. There can only be some number k of such increments before the posterior belief exceeds 1. Hence, ε must be small enough so that j cannot have a chance to reveal rationality in each of these first k subintervals.

As $n \to \infty$, because the offers in the games g_n become increasingly frequent, the corresponding upper bounds on ε go to 0. Thus, once i has revealed rationality, the maximum amount of time that it can take before the game ends if j continues to act irrationally goes to 0 as $n \to \infty$. This means that by acting irrationally, j can guarantee himself a payoff arbitrarily close to α_j for n sufficiently high.

□

This leads (with a little further technical work) to the result that the continuous-game equilibrium is the limit of the discrete-game equilibria.

So it remains just to analyze the continuous-time war of attrition. This is a well-known game, but with the twist that there are irrational types. In equilibrium, let F_i denote the cdf of times when i concedes — unconditional on i's type; thus $\lim_{t\to\infty} F_i(t) \leq 1 - z_i$ because the irrational player never concedes.

What are the rational player i's payoffs from holding out until time t, then conceding? We get

$$u_i(t) = \alpha_i \int_0^{t^-} e^{-r_i y}\, dF_j(y) + \frac{1}{2}(\alpha_i + 1 - \alpha_j)(F_j(t) - F_j(t^-)) + (1 - \alpha_j)(1 - F_j(t))e^{-r_i t}$$

(these terms correspond to i winning, both players conceding at the same time, and j winning, respectively). If t belongs to the support of F_i, then $t \in \text{argmax}\ u_i(t)$.

Properties of the equilibrium are:

- At most one player concedes at time 0: if both conceded at time 0 with positive probability, then each player would prefer to wait and concede later; the loss from waiting is negligible while the gain from winning is discrete.
- There is no interval of time in which neither player concedes, but such that concessions do happen later with positive probability. There is also no interval during which only one player concedes with positive probability. Neither player's concession time

distribution has a mass point on any positive time. (All of these are shown by similar tricks.)

- After 0, each player concedes with a constant hazard rate. This hazard rate can be computed to be $\lambda_i = r_i(1-\alpha_i)/(\alpha_i+\alpha_j-1)$. The reason is that i has to concede at a rate that makes j indifferent to conceding everywhere on his support. Writing down j's local indifference condition, we see that this uniquely determines i's instantaneous hazard rate of concession. So each player has a constant hazard rate, and we can work out from the indifference conditions what those hazard rates are.
- Both players stop conceding at the same time, at which point they are both known to be irrational. This is because if player i continued to concede after j could no longer concede, then player i would prefer to deviate by conceding earlier (since he knew he was going to end up conceding anyway). So they stop conceding at the same time, and no agreement can happen after that time. If i still had positive probability of being rational at that point, then i would prefer to continue conceding rather than wait and get no agreement.

The constant-hazard-rate finding tells us that F_i must have the form $F_i(t) = 1 - c_i e^{-\lambda_i t}$ for some constant c_i. The constants c_1, c_2 can be computed from the fact that both players become known to be irrational at the same time ($F_1^{-1}(1-z_1) = F_2^{-1}(1-z_2)$) and that only one player can concede with positive probability at time 0 (so either c_1 or c_2 is 1).

If i is the player who has positive probability of conceding at time 0, then i's ex ante expected payoff in equilibrium must be $1-\alpha_j$ (if he is rational). And j's indifference to conceding at any positive time implies that his ex ante expected payoff is $F_i(0)\alpha_j + (1 - F_i(0))(1-\alpha_i)$.

Thus we get a fairly detailed understanding of how bargaining games with frequent offers play out when irrational types are present; and a relationship between reputation and the war of attrition.

14.126 Game Theory
Spring 2010