Notes for Economics 200c*

S. Nageeb Ali†

UCSD

April 3, 2011

*These notes draw heavily upon a number of textbooks and unpublished lecture notes by Susan Athey, Dirk Bergemann, Doug Bernheim, Navin Kartik, Jon Levin, Paul Milgrom, David Miller, Steve Morris, Ilya Segal, Joel Sobel, and Steve Tadelis. In other words, one of many individuals could be responsible for any error but the fault is mine in creating errors when they did not exist (the probable event) or propagating them when they did exist before (improbable). I hope you will bring errors to my or Aislinn's attention.

†Assistant Professor of Economics. E-mail: snali@ucsd.edu; Web page: dss.ucsd.edu/~snali; Office: Economics Bldg., Rm 214; Address: 9500 Gilman Dr. #0508, La Jolla, CA 92093–0508.

Contents

Preface

These lecture notes describe the material that will be taught during the first five weeks in ECON 200c in 2010-2011. During these weeks, we shall study:

1. Repeated games with perfect monitoring.
2. Static games with Incomplete information.
3. Dynamic games with incomplete information.
4. Reputation effects.

I am not planning to discuss all of these notes in lecture, but expect that you understand them well for the problem sets and exams; you should think of my lectures as the "highlights" of the material, and that a thorough learning shall come from reading these notes, solving problems, and discussing the material with your classmates. There are several good general textbooks in game theory that you may like to consult for greater discussion: Fudenberg and Tirole (1991), Mas-Colell, Whinston, and Green (1995), Myerson (1997), and Osborne and Rubinstein (1994).

Part I

Repeated games with Perfect Monitoring

Chapter 1

Preliminaries

1.1 Introduction

You have learned quite a few things about repeated games already. This chapter re-introduces you to them, perhaps slightly differently, discusses the one-stage deviation principle and the self-generation technique to construct equilibria.

1.2 The Setting

Consider this section to be an exercise of learning a language; there are no formal results herein. I adopt the formulation of the stage and repeated game in Mailath and Samuelson (2006).[1]

[1]This may or may not correspond to what you have seen in the previous quarter; if it doesn't, seeing a variety of approaches and formalizations probably helps rather than hurts you.

Stage Game

There are n players, and a generic player i chooses action $a_i \in A_i$, and we denote the set of action profiles by A and a generic action profile by a; assume that A_i is finite, and let Player i's payoffs, u_i, be extended to mixed actions by taking expectations: i.e., let α_i be a mixed action in $\triangle(A_i)$ and the set of (possibly correlated) mixed action profiles be in $\triangle(A)$. Let u be a vector of payoffs for the n players. The set of stage-game payoffs generated by A is

$$\mathscr{F} \equiv \left\{ v \in \mathbb{R}^n : v = u(a) \text{ for some } a \in A \right\} \tag{1.1}$$

and the set of *feasible* payoffs is

$$\mathscr{F}^\dagger \equiv \text{co}\mathscr{F}, \tag{1.2}$$

the convex hull of $\mathscr{F}$.[2] In much of repeated games, payoffs are supported through the use of punishments. As such, it's typically important to assess just how much a player i can be punished. This role is played by the *minmax*:

1. Suppose that all other players choose a profile a_{-i},
2. and Player i best responds to it,
3. the (pure action) minmax corresponds to the lowest payoff that player i obtains given that he behaves optimally wrt to their punishment action profile.

Definition 1.1. Player i's *pure action minmax payoff* is

$$v_i^p \equiv \min_{a_{-i} \in A_{-i}} \max_{a_i \in A_i} u_i(a_i, a_{-i}). \tag{1.3}$$

Exercise 1. Contrast conceptually $\min_{a_{-i} \in A_{-i}} \max_{a_i \in A_i} u_i(a_i, a_{-i})$ from $\max_{a_i \in A_i} \min_{a_{-i} \in A_{-i}} u_i(a_i, a_{-i})$. Which is (weakly) higher and why? Prove it.

[2] I explain why we are interested in the convex hull in Section 3.2.

One can extend the notion of minmax to allow other players to *independently* randomize when punishing player i.

Definition 1.2. Player i's *mixed action minmax payoff* is

$$\underline{v_i} \equiv \min_{\alpha_{-i} \in \times_{j \neq i} \Delta(A_j)} \max_{a_i \in A_i} u_i(a_i, \alpha_{-i}). \tag{1.4}$$

A payoff, v_i is individually rational for player i if $v_i \geq \underline{v_i}$; for the purposes of Folk Theorems, it is useful to consider payoffs that are strictly individually rational.

Definition 1.3. The set of *feasible* and *strictly individually rational* payoffs is

$$\mathscr{F}^* \equiv \left\{ v \in \mathscr{F}^\dagger : v_i > \underline{v_i} \text{ for every } i \right\}. \tag{1.5}$$

Exercise 2. Find a game in which the mixed action minmax payoff is strictly lower than the pure action minmax.

Exercise 3. The mixed action minmax assumes that other players independently randomize. Define a notion of minmax in which opponents randomize in a *correlated* manner. Is the *correlated action minmax payoff* for player i (weakly) lower, same, or higher than her mixed action minmax payoff? Prove your assertion.

The Repeated Game

The repeated game (also sometimes referred to as the *super-game*) is a repetition of the stage game for each non-negative integer $t < T$ in which if T is finite and strictly positive, this is a *finitely repeated* game, and if $T = \infty$, then it is an *infinitely repeated* game.

We study *perfect monitoring* games in which all behavior in prior rounds is observed: a history h^t is simply a list of t action profiles: $h^t = \{a^0, ..., a^{t-1}\}$ denotes a history of actions at period t in which $a^\tau = (a_1^\tau, ..., a_n^\tau)$ is the action profile at time τ. The space of possible histories at time $t > 0$ is $H^t \equiv A^t$ and $H^0 \equiv \emptyset$, and the space of all possible histories is $H \equiv \cup_{t=0}^{\infty} H^t$.

A *pure strategy* is a mapping from the set of all possible histories into the set of pure actions, $\sigma_i : H \to A_i$, a *mixed strategy* is a mixture over the set of all pure strategies, and a *behavioral strategy* is a mapping $\sigma_i : H \to \triangle(A_i)$. (I trust that you recall the conceptual distinctions between mixed and behavioral strategies from Watson's half of 200B).

For any history h^t, the *continuation game* is the infinitely repeated game that begins in period t following history h^t. For a strategy profile σ, player i's *continuation strategy induced by h^t*, denoted $\sigma_i|_{h^t}$ is

$$\sigma_i|_{h^t}(h^\tau) = \sigma_i(h^t h^\tau), \tag{1.6}$$

in which $h^t h^\tau$ is the concatenation of the history h^t followed by h^τ. We represent the *continuation strategy profile induced by h^t* as $\sigma|_{h^t} = \left(\sigma_1|_{h^t}, ..., \sigma_n|_{h^t}\right)$.

Notice that $\sigma_i|_{h^t}$ is a strategy in the original repeated game; therefore, the continuation game associated with each history is a *subgame* that is strategically identical to the original game. In other words, repeated games have a straightforward recursive structure.

Towards determining payoffs in this environment, if T is finite, then payoffs correspond to the terminal nodes of the game. In an infinitely repeated game, payoffs instead have to be assigned to an *outcome path*, i.e., an infinite sequence of action profiles, $\mathbf{a} \equiv \left(a^0, a^1, a^2, ...\right)$. Any pure strategy profile maps into a unique outcome path, whereas a behavioral strategy profile generates a distribution over outcome paths; in other words, the former yields a *deter-*

ministic flow payoff and the latter an *expected* flow payoff. In both cases, the payoff obtained over time can be aggregated via *exponential discounting*: in a finite horizon game with T rounds, a sequence of expected payoffs $\{u_i^t\}_{t=0}^T$ generates aggregate payoffs of

$$\sum_{t=0}^{T} \delta^t u_i^t,$$

in which $\delta < 1$ is the discount factor. For the infinite horizon game, it is more useful to normalize payoffs into the *average discounted payoffs*

$$(1-\delta)\sum_{t=0}^{\infty} \delta^t u_i^t.$$

The specific reason that these are called average or normalized discounted payoffs is that if a player obtains a *constant* payoff of u in each period, then the player obtains $(1-\delta)\sum_{t=0}^{\infty}\delta^t u$, which is u.

There are two features of discounting to note. First, assuming that players discount ($\delta < 1$) treats periods asymmetrically: a player cares more about the payoffs in an early period than in a later one. In a finite horizon game, there is no challenge in setting $\delta = 1$, but in an infinite horizon game, total payoffs can be unbounded with $\delta = 1$. There are alternatives to discounting, but these are seldom used. Second, discounting can represent both time preferences, as we have done so implicitly here, or uncertainty about when the game will end, e.g., suppose that conditional on reaching a period t, the game continues to the next period with a probability $\delta < 1$.

Exercise 4. Suppose that Ann and Bob face the game below:

Ann and Bob met at a Led Zeppelin concert in the 70s, and fell in, and then out of love (and had a rough break-up). Both of them prefer Led Zeppelin to Rush, but really would prefer to not run into each other (in which case, the

		Bob	
		Led Zeppelin	Rush
Ann	Led Zeppelin	0,0	3,1
	Rush	1,3	0,0

Figure 1.1: The Battle of the Exes

music becomes irrelevant). They first sought a "love doctor" (a sociologist) to help them with their woes, and eventually turn to you, a practictioner of the dismal arts with the following questions:

1. Please help us identify the set of feasible and strictly individually rational payoffs in the static game. What are the pure and mixed action minmax payoffs?
2. We believe in repeated games. Both of us have discount factor δ: if we decide to alternate bands in each period, with Ann going to see Led Zeppelin and Bob going to see Rush at $t = 0$, what are our average payoffs?
3. Now suppose that Ann has discount factor δ_A, and Bob has discount factor $\delta_B < \delta_A$. Can you suggest a strategy profile that obtains average payoffs outside the set of feasible payoffs from the static game?

1.3 Definition of Equilibria

Definition 1.4. A strategy profile σ is a *Nash equilibrium of the repeated game* if for all players i and strategies σ'_i, $U_i(\sigma) \geq U_i(\sigma'_i, \sigma_{-i})$.

Exercise 5. Prove that if σ is a pure (alt. mixed) strategy equilibrium, then for all i, $U_i(\sigma) \geq v_i^p$ (alt. $\underline{v_i}$).

The challenge of Nash equilibria in dynamic games is that, as we know, they permit too much, including irrational behavior off the equilibrium path. Accordingly, it is more appropriate to restrict attention to subgame perfect equilibria (SPE).

Definition 1.5. A strategy profile σ is a SPE of the repeated game if for all histories h^t, $\sigma|_{h^t}$ is a Nash equilibrium of the repeated game.

Existence is immediate: the stage game has a Nash equilibrium, and repetition of it in every period and following every history is necessarily an equilibrium of the repeated game. A more challenging exercise is verifying that a strategy profile is a SPE: checking whether countably infinite number of strategy profiles are Nash equilibria is simply infeasible. Furthermore, when checking whether a strategy profile is Nash, a player could choose to deviate in any of infinitely many periods. The following section discusses tools to simplify this task.

Exercise 6. Suppose that $T < \infty$ and the stage game has a unique (possibly mixed) Nash equilibrium, α^*. Prove that the unique SPE of the repeated game is the history-independent strategy profile σ^* such that for all t and h^t, $\sigma^*(h^t) = \alpha*$.

Exercise 7. Consider an overlapping generations model (not a repeated game): agents $t = 0, 1, 2, ...$ choose an action $a_t \in \{0, 1\}$ in sequence observing the actions of all agents who moved prior to them. Agent $t's$ payoff is given by $va_{t+1} - ca_t$, where $v > c > 0$.

1. What is the space of histories that agent t observes? How many pure strategies does agent t have?

2. Show that there is a subgame perfect equilibrium in which 0 is played in every period. Is there a subgame perfect equilibrium in which 1 is played in every period?

3. Suppose that rather than observing the actions of all previous players, agent t observes only the action of his predecessor, agent $t-1$. How many strategies does agent t now have? Show that in any pure strategy SPE, 0 is played in every period.

4. What if agent t can observe the actions of upto the M agents directly before him where $1 < M < \infty$. Is there a pure strategy SPE in which all agents choose 1 on the equilibrium path?

Exercise 8. Suppose the following game is played thrice with no discounting. Show that if $x \geq 1/2$, then there is a SPE in which (A, A) is played in the first period.

	A	B	C
A	$3,3$	$-1,4$	$0,0$
B	$4,-1$	$0,0$	$0,-1$
C	$0,0$	$-1,0$	x,x

1.4 The One-Shot Deviation Principle

One of the most important results to know when studying repeated games is the *one-shot deviation principle*: its essence is that when evaluating profitable deviations from an SPE, it suffices to consider strategies that deviate at exactly one information set, i.e., the behavior in this deviating strategy is identical to that of the SPE in every other history. If a strategy profile has no profitable *one-shot deviation*, then it is an SPE.

Definition 1.6. A *one-shot deviation* for player i from strategy σ_i is a strategy σ_i' such that there exists a unique history h^t such that

$$\sigma_i\left(h^t\right) \neq \sigma_i'\left(h^t\right)$$
$$\sigma_i\left(h^\tau\right) = \sigma_i'\left(h^\tau\right) \text{ for all } h^\tau \neq h^t.$$

Definition 1.7. Fix a strategy profile σ: a one shot deviation σ_i' is *profitable* if for some h^t,

$$U_i(\sigma_i'|_{h^t}, \sigma_{-i}|_{h^t}) > U_i(\sigma|_{h^t}).$$

The payoffs from a deviation are evaluated *ex interim*, i.e. based on payoffs realized once histories in which the deviation differs are realized, and not evaluated *ex ante* at the beginning of time. Thus, the history that makes a deviation profitable may be off the path of play. This should make you reflect on how the one shot deviation principle relates to Nash equilibria and SPE: a Nash equilibrium can have profitable one-shot deviations if these deviations occur off the equilibrium path; in contrast, because an SPE must be immune to deviations on and off the equilibrium path, there can be no profitable one-shot deviations. In other words, the absence of profitable one-shot deviations is *necessary* for a profile to be SPE; it is also sufficient.

Theorem 1.8. *A strategy profile σ is a SPE if and only if there are no profitable one-shot deviations.*

Proof. Necessity ("only if") is immediate as outlined above: a strategy profile with a profitable one-shot deviation cannot be a SPE. Sufficiency is more subtle, and the proof below works only for pure strategy SPE.

Suppose that σ is not a SPE: then there exists a history h^t, player i, and a strategy σ_i' that satisfies

$$U_i(\sigma_i'|_{h^t}, \sigma_{-i}|_{h^t}) - U_i(\sigma|_{h^t}) > 0. \tag{1.7}$$

Let ϵ be the difference expressed in the LHS above. First notice that without loss of generality, we can assume that $h^t = \emptyset$ because we can treat the subgame and everything that follows as the entire game. If σ_i' were a one shot deviation, we would be done; we will suppose otherwise. The first step is to show that there must exist a profitable deviation in a finite number of histories, and the second step is to use that deviation to construct a profitable one shot deviation.

Step 1: Since the game is finite, there exists some $M < \infty$ such that

$$\max_{a,a' \in A} |u_i(a) - u_i(a')| < M,$$

and so there exists T sufficiently large that $\delta^T M < \epsilon/2$. Re-examine Eq. 1.7: since the difference in payoffs is ϵ, and a difference strictly less than $\epsilon/2$ can be obtained from any behavioral changes that happen after period T, it must be that deviations in the first T periods accounts for more than half the gain that Player i obtains from deviating from σ_i to σ_i', assuming that others play σ_{-i}.[3] Consider a strategy σ_i'' that is identical to σ_i' in the first T periods and to σ_i thereafter; by the previous sentence, this is a profitable deviation that deviates at only a finite number of histories.

Step 2: This step follows an inductive argument. The strategy σ_i'' departs from σ_i in the first T stages: consider the $T-1$ period history h^{T-1} induced by

[3] Recall that we assumed wlog that $h^t = \emptyset$ and zoomed into the subgame as if this were the entire game.

$(\sigma_i'', \sigma_{-i})$, and think about the difference

$$U_i(\sigma_i''|_{h^{T-1}}, \sigma_{-i}|_{h^{T-1}}) - U_i(\sigma|_{h^{T-1}}).$$

If this difference is strictly positive, then we have a profitable one-shot deviation for player i which is at history h^{T-1}. If this difference is weakly negative, redefine σ_i'' to coincide with σ_i at history h^{T-1}, but don't change it otherwise; consider the $T-2$ period history h^{T-2} induced, and evaluate the difference above (replacing h^{T-1} with h^{T-2}). If this difference is positive, we have a profitable one-shot deviation; otherwise, continue this process iteratively. Eventually, the difference is positive, otherwise, there is a contradiction to Step 1. □

The essence of the proof is that because payoffs are discounted, any strategy that is a profitable deviation to the proposed profile σ must be profitable within a finite number of periods. Once we can restrict attention to finite steps, backward induction proves the existence of a profitable one-shot deviation. Discounting is not necessary for Theorem 1.8: Fudenberg and Tirole (1991) shows that it holds for any extensive-form game with perfect monitoring in which payoffs are *continuous at infinity*, which is guaranteed by discounting.

Naturally, you may be wondering about the relationship between the one-shot deviation principle and Nash Equilibria that are not sub-game perfect. Since Nash equilibria do not demand incentives off the equilibrium path, profitable (one or multi-shot) deviations off the path of play do not prevent a strategy profile from being a Nash equilibrium. A more reasonable hope is that restricting attention to histories that arise on the equilibrium path would relate Nash equilibria to one-shot deviations: i.e., you may conjecture:

Conjecture 1. *A strategy profile σ is a Nash Equilibrium if and only if there are no profitable one-shot deviations from histories that arise on the path of play."*

The exercise below should help you evaluate the validity of this conjecture.

Exercise 9. Consider the Prisoner's Dilemma below (feel free to write E for Effort, and S for Shirk in your answer):

		Bob	
		Effort	Shirk
Ann	Effort	$3,3$	$-1,4$
	Shirk	$4,-1$	$1,1$

Figure 1.2: The Prisoner's Dilemma

1. What is the set of feasible and individually rational payoffs?
2. If the game is repeated infinitely often, what conditions on δ would have to be satisfied so that both choose to work in each period in an SPE? Describe the SPE strategy profile.
3. Consider a "tit-for-tat" strategy profile in which both players work at time $t = 0$, and for $t \geq 1$, each player mimics the action that the opponent picked in the previous period.
   a) Suppose that Ann shirks forever while Bob follows "tit-for-tat". What are her average expected payoffs from this "infinite-shot" deviation?
   b) Consider the one-shot deviation from the equilibrium path: suppose that Ann chooses to deviate only at $t = 0$, but then follows tit-for-tat forever after. What are her and Bob's average expected

payoffs? Under what condition on δ is this not a profitable one-shot deviation?

c) Under what condition on δ is tit-for-tat a Nash profile? Compare the condition with that above to evaluate Conjecture 1.

d) Prove that tit-for-tat is not a SPE, regardless of the value of δ.

Chapter 2

Repeated Game Payoffs

Disclaimer: restrict attention throughout this chapter to pure actions and strategies.

2.1 Introduction

This is the coolest technique in repeated games; it is one that is typically not taught in the first-year core, but given its greater use throughout microeconomics and macroeconomics, we feel that it should be introduced to all students earlier.

We know that the future is a powerful incentive mechanism, but its difficult to understand what repeated games can or cannot achieve when the space of strategy profiles themselves are infinite-dimensional spaces. Abreu et al. (1986, 1990) recognized that one can decompose repeated games into straightforward dynamic programming problems in which behavior today is implemented with self-enforcing “utility payments” tomorrow; this problem is analagous to a principal-agent setting in which the agent is rewarded or punished for certain actions. The caveat is that these promises of utility have to be"self-enforcing," or in other words, generated from equilibria in the repeated game. This decomposition of SPE payoffs into flow payoffs today and promised utility tomorrow simplifies the study of repeated interaction, but is also subtle.

2.2 Enforceability

Consider functions of the form $\gamma : A \to W$ in which W is a non-empty subset of $\mathbf{R}^n$: you should think of $\gamma_i(a)$ as the payment that player i obtains when the action profile is a. That the range of γ is W means that only payments in W can be used. These payments do not happen immediately but in the future; their values determines the strength of the incentives that they create.

Definition 2.1. The action profile a^* is *enforceable* on W if there exists some specification of continuation promises $\gamma : A \to W$ such that for all players i and actions a_i,

$$(1-\delta)\,u_i\left(a^*\right)+\delta\gamma_i(a^*) \geq (1-\delta)\,u_i(a_i, a^*_{-i})+\delta\gamma_i(a_i, a^*_{-i}). \tag{2.1}$$

We say that γ *enforces* a^* on W if it is a continuation promise that satisfies above.

An action profile is enforceable using promised utilities from a set W if there exists continuation promises that make the action profile incentive compatible for every player. We now iterate this logic: we have a set of promised utilities W, and actions that are enforceable given those promised utilities; why not think about payoffs that emerge from combining actions that are enforceable on W and the enforcing continuation promises?

2.3 Decomposability

Definition 2.2. A payoff v is *decomposable* on W if there exists an action a^* that is enforceable on W such that

$$v_i = (1-\delta)u_i(a^*) + \delta\gamma_i(a^*) \tag{2.2}$$

for some $\gamma : A \to W$ that enforces a^*. The set of payoffs that are decomposable on W when the discount factor is δ is denoted by $\mathscr{B}(\delta, W)$.

Exercise 10. Consider a payoff v such that for some player i, $v_i < \underline{v_i}$: is v decomposable on any set $W \subset \mathscr{F}^\dagger$? Offer an example or prove otherwise.

Exercise 11. Prove the following statements:

1. $\mathscr{B}(\delta, W)$ is a *monotone* operator: $W \subset W' \Rightarrow \mathscr{B}(\delta, W) \subset \mathscr{B}(\delta, W')$.
2. $\mathscr{B}(\delta, \mathscr{F}^\dagger) \subset \mathscr{F}^\dagger$.
3. If $\mathscr{B}^k(\delta, W)$ for integers $k > 1$ is defined recursively by

$$\mathscr{B}^k(\delta, W) \equiv \left\{ v : v \text{ is decomposable on } \mathscr{B}^{k-1}(\delta, W) \right\}.$$

   then, for every integer $k \geq 1$, $\mathscr{B}^{k+1}(\delta, \mathscr{F}^\dagger) \subset \mathscr{B}^k(\delta, \mathscr{F}^\dagger)$

To make matters concrete, return to the Prisoner's Dilemma in Figure 1.2: let $W = \{(3,3),(1,1)\}$ and let us see whether (E,E) is enforceable on W. Consider γ such that

$$\begin{aligned} \gamma(E,E) = &\quad (3,3), \\ \gamma(E,S) = \gamma(S,E) = \gamma(S,S) = &\quad (1,1). \end{aligned} \tag{2.3}$$

Then (E,E) is enforceable on W if

$$(1-\delta)(3) + \delta(3) \geq (1-\delta)(4) + \delta. \tag{2.4}$$

If the above condition is satisfied, then because $3 = (1-\delta)(3) + \delta(3)$, the payoff $(3,3)$ is decomposable on W. Notice that $(1,1)$, being the payoff from the stage game Nash, is also decomposable on W [use a constant γ], and so W is decomposable on itself. This is our first example of a *self-generating set* of payoffs.

2.4 Self-Generation

Definition 2.3. A set $W \subseteq \mathscr{F}^{\dagger}$ is *self-generating* if every payoff in W is decomposable on W, i.e., $W \subset \mathscr{B}(\delta, W)$.

Notice that W may be self-generating, but $\mathscr{B}(\delta, W)$ may contain payoffs not in W: i.e., more payoffs are decomposable on W than those in W. This should naturally motivate you to wonder about the largest self-generating set.

Exercise 12. Suppose that $W = \mathscr{B}(\delta, W)$ and W is non-empty and a strict subset of W': can W' also be self-generating? Construct an example or prove otherwise.

One reason to be interested in self-generating sets is their tight connection to SPE payoffs. Let $\mathscr{E}^p(\delta)$ be the set of SPE payoffs when players have discount factor δ.

Theorem 2.4. *Any set of payoffs that is self-generating is a set of SPE payoffs; a fortiori, any set of payoffs that are decomposable on a self-generating set is a set of SPE payoffs: i.e.,* $W \subset \mathscr{B}(\delta, W) \Rightarrow \mathscr{B}(\delta, W) \subset \mathscr{E}^p(\delta)$.

I won't present the formal argument here (nor expect you to know it) but the intution should be compelling: each payoff v in $\mathscr{B}(\delta, W)$ can be associated with an action a^* and an enforcing promised utility function γ such that the selected action profile is a^* (no player has an incentive to deviate). In the continuation game, whatever action profile a is played, the promised utility $\gamma(a)$ is in W; because W is self-generating, the promised utility is also decomposable on W, and one can iterate this argument.

Theorem 2.5. *The following are true:*

1. *$\mathscr{E}^p$ is self-generating: $\mathscr{E}^p \subset \mathscr{B}(\delta, \mathscr{E}^p)$.*
2. *Any payoff that is decomposable on $\mathscr{E}^p$ is an SPE payoff: $\mathscr{B}(\delta, \mathscr{E}^p) \subset \mathscr{E}^p$. Therefore, $\mathscr{E}^p = \mathscr{B}(\delta, \mathscr{E}^p)$.*
3. *$\mathscr{E}^p$ is the largest self-generating set.*

Exercise 13. Prove the above result. (This may be hard: try your best to outline the logic).

Viewing SPE payoffs through the lens of decomposability and self-generation makes it feasible to construct algorithms to compute the set of SPE payoffs, based on the statements in Exercise 11. The algorithm begins with the set of all feasible payoffs, $\mathscr{F}^\dagger$, finds the set of payoffs that are decomposable on it, $\mathscr{B}(\delta, \mathscr{F}^\dagger)$, then the set of payoffs decomposable on that, and so on and so forth. Given that one has a sequence of decreasing sets, it is of interest to look at payoffs in all of them:

$$\mathscr{F}^\dagger_\infty \equiv \bigcap_{k=1}^{\infty} \mathscr{B}^k(\delta, \mathscr{F}^\dagger).$$

The following result is true, useful, and offered without a proof here:

Theorem 2.6. $\mathscr{F}^\dagger_\infty = \mathscr{E}^\delta$.

Exercise 14. Prove the easy part of the above theorem: $\mathscr{E}^\delta \subset \mathscr{F}^\dagger_\infty$.

Exercise 15. Consider an infinitely repeated version of the game

	L	M	R
U	$1,1$	$3,0$	$-2,0$
M	$0,3$	$2,2$	$-2,0$
D	$0,-2$	$0,-2$	$-4,-4$

1. Suppose $\delta \geq \frac{1}{2}$. Show that the set $\{(1,1),(2,2)\}$ is self-generating.
2. Suppose $\delta = \frac{1}{3}$. Find a self-generating set that contains the point $(2,2)$.
3. Show that there is no self-generating set that contains the point $(-3,-3)$.

Chapter 3

Folk Theorems

3.1 Introduction

Folk theorems are staples in repeated games: one must consume or be aware of them in thinking about interactions that persist over time. What a folk theorem asserts is that every feasible and strictly individually rational payoff can be associated with some SPE so long as players are sufficiently patient. The idea of a folk theorem is to use variations in payoffs to create incentives towards particular kinds of behavior; the variations in any single stage-game may be small, but are amplified when they exist in each period as $\delta \to 1$. I will present three different versions of this result:

1. *Nash Threats Folk Theorem*: A weaker version in which payoffs above some stage-game Nash are supportable using stage-game Nash equilibria as punishments; the set of payoffs supportable here is generally less than the feasible and strictly individually rational payoff set.

2. *Folk Theorem for Nash Equilibrium*: This folk theorem supports payoffs that are feasible and strictly individually rational, but only as *Nash equi-*

libria.

3. *Folk Theorem for SPE*: This is the strongest version (in these notes) that supports payoffs that are feasible and strictly individually rational as *SPE*.

I return to questions of how to interpret such results after presenting them.

3.2 Convexity

The set of payoffs that I want to support in equilibrium is a subset of $\mathscr{F}^{\dagger}$, the convex hull of stage-game payoffs, but there is a subtlety here: the set of feasible payoffs in the stage game, and therefore, also in the repeated game for small discount factors, need not be convex. For example, in the Battle of the Exes in Figure 1.1, the payoff $(2,2)$ is unattainable by independent randomizations of the players. There are two approaches to tackle this issue, one that is tricky and the other that is tractable:

1. *Tricky*: Sorin (1986) and Fudenberg and Maskin (1991) show that this nonconvexity disappears as $\delta \to 1$ by using nonstationary sequences of outcomes. So long as δ is near 1, a time-varying deterministic path can attain payoffs close to any convex combination of pure-strategy payoffs. For example, payoffs arbitrarily close to $(2,2)$ in Figure 1.1 can be attained as $\delta \to 1$ by playing one pure Nash equilibrium of the state-game in every odd period and the other in every odd period. This approach is elegant but beyond the material of this course.

2. *Tractable*: Fudenberg and Maskin (1986) show that these nonstationarities can be avoided through *public correlation devices* in which all players publicly observe the outcome of a random device. This public observation of the random device allows the grand coalition of all players to

correlate their strategies such that they can play any mixed action profile in $\triangle(A)$ rather than $\triangle(A_1) \times ... \times \triangle(A_n)$.[1] We will use this approach throughout this course (and this has become a standard assumption in repeated games).

Formally, let $\{\omega^0, ..., \omega^t, ...\}$ be a sequence of independent draws from, without loss of generality, a uniform distribution from $[0, 1]$ and assume that players observe ω^t at the beginning of period t. With public correlation, the history relevant in period t is

$$h^t \equiv \{a^0, ..., a^{t-1}, \omega^0, ..., \omega^t\},$$

so it includes all prior actions, all prior realizations of the public random variable, and the current realization of it.

Exercise 16. Does a public correlation device obviate the need for players to randomize in a mixed action minmax? Or with public correlation, is a mixed action minmax always attainable by players playing pure actions conditioned on the realization ω?

3.3 Nash Threats Folk Theorem

This result is often also called the Nash Reversion Folk Theorem as well, and is most commonly used in applications: the idea is that for each player i, the punishment from not engaging in equilibrium behavior is a repetition of the stage game Nash that offers the lowest payoff to him. By reverting to a stage-game Nash equilibrium, once players are in a punishment phase, dynamic

[1] Notice that public correlation differs from that introduced in Exercise 3: in the former, even Player i who is being minmaxed observes the random variable used for correlation, whereas in the exercise, Player i is not privy to this information.

incentives are no longer necessary to ensure that all players cooperate in the punishment. Joel Watson presented this folk theorem to you so I will be brief.

Theorem 3.1. *Let α^* be an equilibrium of the stage game with expected payoffs v^{NE}. Then for any $v^* \in \mathscr{F}^\dagger$ such that $v_i^* > v_i^{NE}$, there exists some $\underline{\delta}$ such that for all $\delta > \underline{\delta}$, there is an SPE with payoffs v^*.*

Proof. First, assume that there is an action profile a^* such that $u(a^*) = v^*$, and consider the following strategy profile:

1. *Cooperation Phase*: If $t = 0$ or $t \geq 1$ and a^* was played in every prior period, play a^*.
2. *Punishment Phase*: If any other action profile is played in any prior period, then play α^* for every subsequent period.

Let us argue that this is a SPE: once a deviation occurs, players are repeating a stage game Nash equilibrium regardless of history; since this is a Nash equilibrium of the repeated game at every history, there is nothing more to be established about histories off the equilibrium path. So it suffices to argue that no player has a unilateral profitable one-shot deviation in the Cooperation Phase. The relevant incentive condition is

$$(1-\delta)\max_{a_i} u_i\left(a_i, a_{-i}^*\right) + \delta v_i^{NE} < v_i^*. \tag{3.1}$$

Since the inequality holds strictly at $\delta = 1$, and the left-hand side is continuous in δ, it is satisfied for δ sufficiently close to 1. Now, assume that there is no action profile such that $u(a^*) = v^*$, but $v^* \in \mathscr{F}^\dagger$. The argument is a bit trickier, but not overly so: replace the action profile a^* with a public randomization stage-game strategy $a^*(\omega)$ that yields expected payoff v^*. The punishment phase incentives are unaffected, but the cooperation phase incentives change in the following sense: depending on ω^t, player i may not receive exactly v^* in the current period[2], and his deviation payoff also depends on the realization of the public correlation device. However, the cooperation phase incentive condition is satisfied so long as

$$(1-\delta)\max_a u_i(a) + \delta v_i^{NE} < (1-\delta) u_i(a(\omega)) + \delta v_i^*. \tag{3.2}$$

As before, this inequality holds at $\delta = 1$, and so is satisfied so long as δ is sufficiently close to 1. □

The following exercise highlights how focusing on Nash reversion can exclude certain repeated game payoffs.

Exercise 17. Give an example of a stage-game that has a strictly dominant strategy for each player (and hence a unique Nash equilibrium) but in which there is an SPE of the repeated game in which each player obtains strictly lower payoffs than the payoffs in the Nash equilibrium of the stage game. Prove that the strategy profile you construct is an SPE.

3.4 Folk Theorem for Nash Equilibrium

Theorem 3.2. *For every payoff $v \in \mathscr{F}^*$, there exists some $\underline{\delta}$ such that for all $\delta > \underline{\delta}$, there is a Nash Equilibrium with payoffs v^*.*

The set of payoffs that are supportable is larger than that of the Nash Threats Folk Theorem, but the solution-concept is more permissive. The proof is very straightforward and mirrors the previous proof (because punishment phase incentives are trivial).

[2] The payoff v^* is only the expected *ex ante* payoff in each period in the Cooperation Phase but may not be the *ex post* payoff.

Proof. First, assume that there is an action profile a^* such that $u(a^*) = v^*$, and consider the following strategy profile:

1. *Cooperation Phase*: Play a^* if (1) $t = 0$ or (2) $t \geq 1$ and a^* was played in the previous period or (3) $t \geq 1$ and the realized action profile differs from a^* in two or more components.

2. *Punishment Phase*: If player i was the only one to not follow profile a^*, then in each period, each player j plays her component of a mixed strategy that makes player i attain his minmax payoffs.

Only behavior in the Cooperation Phase that corresponds to (1) and (2) need to be incentivized; the other histories are off the equilibrium. Incentives are satisfied if

$$(1-\delta)\max_{a_i} u_i\left(a_i, a_{-i}^*\right) + \delta \underline{v_i} < v_i^*. \tag{3.3}$$

As before, this inequality holds at $\delta = 1$, and so is satisfied so long as δ is sufficiently close to 1. The case in which there is no action profile such that $u(a^*) = v^*$ is tackled as in the previous theorem. □

This folk theorem is seldom applied because the punishments used may be implausible. For example, consider the game below:

		Bob L	Bob R
Ann	U	5, 5	0, −10
	D	6, 2	0, −10

Figure 3.1: Destroy the World

Theorem 3.2 informs us that $(5,5)$ is a possible payoff although the punishment for Ann is for Bob to play R, which hurts him. The lack of incentives at the punishment stage motivates the search for a Folk Theorem in SPE.

3.5 Folk Theorem for SPE

Theorem 3.3. *Suppose that* $\dim \mathscr{F}^* = n$*; i.e., the dimension of the set of feasible and individually rational players equals the number of players. For every payoff* $v \in \mathscr{F}^*$*, there exists some* $\underline{\delta}$ *such that for all* $\delta > \underline{\delta}$*, there is a SPE with payoffs* v^**.*

As in the previous results, players are punished when they deviate from a particular action. To ensure incentives in the punishment phase, the dimensionality condition plays an important role: an issue that sometimes arises when it fails is that punishing a deviator also punishes someone else (a punisher). In such a circumstance, it would be difficult to incentivize someone else to punish the deviator. The dimensionality condition ensures that a punisher can be rewarded. Prior to the proof, let us highlight with an example from Fudenberg and Maskin (1986) of how its failure can cause a failure of the Folk Theorem.

Carol

		Bob: L	Bob: R		Bob: L	Bob: R
Ann	U	1, 1, 1	0, 0, 0	U	0, 0, 0	0, 0, 0
	D	0, 0, 0	0, 0, 0	D	0, 0, 0	1, 1, 1

Figure 3.2: Curse of Common Values

In the above game, Carol moves left or right, which selects the payoffs determined by Ann's and Bob's moves. Each player's minmax in this game is 0, but there is no profile that simultaneously minmaxes all three players (since one of the three can always deviate and secure himself a payoff of 1). This creates a challenge: punishing one player necessarily entails punishing all three (because their payoffs are perfectly aligned), but all three players cannot be simultaneously punished.

Proposition 3.4. *The minimum payoff attainable in any SPE for any player and any discount factor* δ *is (weakly) greater than* $1/4$.

Proof. Let $\alpha_i(1)$ be the probability that player i plays his first action (i.e. row, column, or matrix); notice that for mixed action profile, there is at least one player i for whom $\alpha_j(1), \alpha_k(1) \geq 1/2$ or $(1-\alpha_j(1)), (1-\alpha_k(1) \geq 1/2$.[3] Therefore, there is at least one player who can obtain a payoff of $1/4$. Therefore, for an equilibrium, let player i be the one who given the first-period actions of others can obtain at least $1/4$ in the first period. Doing so ensures a payoff of at least $(1-\delta)\frac{1}{4} + \delta v^*$, in which v^* is the lowest payoff attainable in any SPE for any player and any discount factor. If there exists an equilibrium in which v^* is obtained by all three players, it must be that this payoff is better than that earlier derived:

$$v^* \geq (1-\delta)\frac{1}{4} + \delta v^*$$

which implies that $v^* \geq \frac{1}{4}$. □

Theorem 3.3 evades such difficulities by breaking the alignment in preferences: if $\dim \mathscr{F}^* = n$, this ensures that players need not be simultaneously

[3]The argument is identical to how if one has a drawer with only black and brown socks, one is guaranteed to find a pair that matches in color when picking out three.

minmaxed but that player-specific punishments are possible. The proof for this theorem is below.

Proof. Fix $v \in F^*$. Suppose for simplicity that:

1. There exists a pure action profile a such that $u(a) = v$,
2. For each player i, assume that a profile m^i that minmaxes player i is in pure actions.

Both of the above ensure that deviations are detectable, whereas with mixed action profiles, deviations are not detectable with probability 1. Tackling mixed action requires more care that I describe after the proof.

Choose some $v' \in \text{Interior}(F^*)$ such that $v'_i \in \left(\underline{v_i}, v_i\right)$, and choose $\epsilon > 0$ such that the vector,

$$v^i = (v'_1 + \epsilon, \ldots, v'_{i-1} + \epsilon, v'_i, v'_{i+1} + \epsilon, \ldots, v'_n + \epsilon) \in F^*.$$

The full dimension assumption ensures that the above exists for some ϵ and v'. The profile v^i is an ϵ-reward for other players relative to v' but not for player i. To avoid public randomizations, assume that there exists an action profile a^i such that $u(a^i) = v^i$. Choose T such that for all i,

$$\max_a u_i(a) + T\underline{v_i} < \min_a u_i(a) + T v'_i \tag{3.4}$$

In other words, the punishment length T is sufficiently long that player i is worse off from deviating and then being minmaxed for T periods than obtaining the lowest possible payoff once and then T periods of v'_i. Finally, the strategy profile is:

1. *Phase I*: Play a so long as no player deviates or at least two players deviate. If player i alone deviates, go to Phase II_i.

2. *Phase II_i*: Play m^i for T periods so long as no one deviates or two or more players deviate. Switch to Phase III_i after T successive periods in II_i; if player j alone deviates, go to II_j.
3. *Phase III_i*: Play a^i so long as no one deviates or two or more players deviate. If a player j alone deviates, go to II_j.

Notice that the above construction makes explicit what actions should follow when two or more players deviate; such a specification is necessary because strategies are complete contingent plans. It may seem odd that when two or more players deviate, no one is punished for it; this does not pose an issue to equilibrium because players compute the expected payoffs of deviating assuming that no one else deviates.

We use the one-shot deviation principle to check that this is an SPE:

Phase I:

1. Follow strategy: v_i
2. Deviate once: At most $(1-\delta)\max_a u_i(a) + \delta(1-\delta^T)\underline{v_i} + \delta^{T+1} v_i'$
3. Conclusion: Since $v_i > v_i'$, the top line is greater for δ sufficiently large.

Phase II_i: (Suppose there are $T' \leq T$ periods left)

1. Follow strategy: $(1-\delta^{T'})\underline{v_i} + \delta^{T'} v_i'$
2. Deviate once[4]: At most $(1-\delta)\underline{v_i} + \delta(1-\delta^T)\underline{v_i} + \delta^{T+1} v_i'$.
3. Conclusion: Deviation offers no short-term reward, and increases the length of punishment so isn't profitable.

Phase II_j: (Suppose there are $T' \leq T$ periods left)

1. Follow strategy: $(1-\delta^{T'})u_i(m^j) + \delta^{T'}(v_i' + \epsilon)$
2. Deviate once: At most $(1-\delta)\max_a u_i(a) + \delta(1-\delta^T)\underline{v_i} + \delta^{T+1} v_i'$.

[4] Recall that i is being minmaxed.

3. Conclusion: For δ close to 1, the ϵ difference between what she gets in Phase III_j vs. that which she gets in Phase III_i makes the deviation unprofitable.[5]

Phase III_i:

1. Follow strategy: v'_i
2. Deviate once: At most $(1-\delta)\max_a u_i(a) + \delta(1-\delta^T)\underline{v_i} + \delta^{T+1}v'_i$.
3. Conclusion: For δ close to 1, the inequality in Eq. 3.4 ensures that this deviation is unprofitable.

Phase III_j:

1. Follow strategy: $v'_i + \epsilon$
2. Deviate once: At most $(1-\delta)\max_a u_i(a) + \delta(1-\delta^T)\underline{v_i} + \delta^{T+1}v'_i$.
3. Conclusion: For δ close to 1, the epsilon difference ensures that the deviation is unprofitable.

Thus, this profile is an SPE.

Throughout the proof, we have assumed that all payoffs are attained using pure action profiles (in which case deviations are immediately detectable). Matters are more subtle when the minmax requires randomization: then a player must receive the same normalized payoff for each action in the support of her mixed strategy, although she may realize different payoffs in the stage game. Thus, continuation values must be adjusted in subtle ways based on the realized actions of the punishers. See Fudenberg and Maskin (1986) for more details. □

Exercise 18. In a repeated game, show that if for each player, there is a SPE where that player's payoff is his minmax value, then any payoff of a Nash equilibrium is also the payoff of a SPE.

[5]This is a place in which rewarding people for punishing others helps with incentives.

Exercise 19. Consider a trading favors environment in which Ann and Bob interact repeatedly at dates $t = 0, 1, 2, ...$ in a *sequential* move game: Ann moves first choosing an assistance $a \in [0, 1]$ to provide to Bob. Having observed Ann's action, Bob chooses to reciprocate by providing $b \in [0, 1]$ in return to Ann. The resulting stage game payoffs are

$$u_A(a, b) = b - \frac{a^2}{4}, \; u_B(a, b) = a - b.$$

Alice and Bob share a common discount factor δ.

1. Identify the static Nash equilibrium payoffs.
2. Identify the set $\mathscr{F}^*$ of feasible and individually rational expected payoffs, and Bob's highest payoff in $\mathscr{F}^*$.
3. Derive conditions on δ under which there is an SPE in which Bob's average payoff equals his highest payoff in $\mathscr{F}^*$.

Exercise 20. Consider a repeated all-pay auction in which at each date $t = 0, 1, 2, ...$, an auctioneer has an object of value $v > 0$ for sale. There are two bidders, Players 1 and 2. When Player i bids b_i^t in period t, and her opponent bids b_j^t, her stage game payoffs are

$$u_i(b_i^t, b_j^t) = \begin{cases} v - b_i^t & \text{if } b_i^t > b_j^t, \\ \frac{v}{2} - b_i^t & \text{if } b_i^t = b_j^t, \\ -b_i^t & \text{if } b_i^t < b_j^t. \end{cases}$$

Both players share a common discount factor δ.

1. Derive the pure action minmax of the stage game. Can the pure action minmax be attained by a pure action Nash equilibrium of the stage game?

2. Derive the mixed action minmax of the stage game. Can the mixed action minmax be attained by a mixed action Nash equilibrium of the stage game?

3. Identify the set $\mathscr{F}^*$ of feasible and individually rational expected payoffs.

4. In the repeated game, assume that each player observes the offers made by both in all prior periods. Under what conditions on δ is there an SPE in which on the equilibrium path, both players bid 0 and each wins the prize with probability 1/2? Describe the off-path behavior of this SPE.

5. Is there a SPE of the repeated game (for any discount factor δ) in which on the equilibrium path, one player obtains zero payoffs in every period whereas the other obtains strictly positive payoffs in every period?

Exercise 21. Consider the following dynamic game. There are an infinite number of periods, $t = 0, 1, 2, \ldots, \infty$, and the players have a common discount factor $\delta < 1$. In each period t the stage game $\langle N, A, \pi^t \rangle$ is played; that is, in each period the set of players and action profiles is the same, but the payoff functions change over time. In particular, $\langle N, A, \pi^t \rangle$ is the following game, where $\gamma \in (0, 1]$:

		Player 2 C	Player 2 D
Player 1	C	$4\gamma^{t-1}, 4\gamma^{t-1}$	$-\gamma^{t-1}, 5\gamma^{t-1}$
	D	$5\gamma^{t-1}, -\gamma^{t-1}$	$0, 0$

1. Suppose $\gamma = 1$. Prove that any feasible and individually rational payoff vector can be approximated arbitrarily closely as the average payoffs of a *pure strategy* subgame perfect equilibrium, given a sufficiently high discount factor $\delta < 1$.

2. Now fix $\gamma < 1$. Suppose the players can use any arbitrary public correlation device. Let V^δ be the set of average payoff vectors that can be supported by subgame perfect equilibria given δ. Characterize $\lim_{\delta\to 1} V^\delta$.

3. Let $V^{\gamma,\delta}$ be the set of average payoff vectors that can be supported by subgame perfect equilibria given γ and δ. Use your earlier answers to understand if

$$\lim_{\delta\to 1}\left(V^{1,\delta}\right) = \lim_{\gamma\to 1}\left(\lim_{\delta\to 1} V^{\gamma,\delta}\right).$$

3.6 Epilogue

The Folk Theorem is often believed to be troubling because it asserts that anything can be an equilibrium. For this reason, non-theorists often abstract from repeated games on the grounds that allowing for repetition will generate multiple equilibria. I personally find this to be very strange: if real-life interaction is indeed repeated, and we as modelers obtain unique predictions because we are abstracting from the repeated nature of interaction, then it is difficult to have any faith in our predictions. In my opinion, a more thoughtful approach is to embrace the repeated nature of interaction, but to have principled methods of selection, which could be selecting Pareto or utilitarian efficient equilibria, those equilibria that are simple, or those that are "renegotiation-proof."

Part II

Games with Incomplete Information

Chapter 4

Introduction

See p. 33-39 of Navin Kartik's lecture notes available here:

http://www.econ.ucsd.edu/~jsobel/200Cs09/Kartik_Notes.pdf.

Bibliography

Dilip Abreu, David Pearce, and Ennio Stacchetti. Optimal cartel equilibria with imperfect monitoring. *Journal of Economic Theory*, 39(1):251–69, June 1986.

Dilip Abreu, David Pearce, and Ennio Stacchetti. Toward a theory of discounted repeated games with imperfect monitoring. *Econometrica*, 58(5): 1041–1063, September 1990.

D. Fudenberg and E. Maskin. On the dispensability of public randomization in discounted repeated games. *Journal of Economic Theory*, 53(2):428–438, 1991.

Drew Fudenberg and Eric S. Maskin. The folk theorem in repeated games with discounting or with incomplete information. *Econometrica*, 54(3):533–554, May 1986.

Drew Fudenberg and Jean Tirole. *Game Theory*. MIT Press, Cambridge, MA, 1991.

George J. Mailath and Larry Samuelson. *Repeated Games and Reputations: Long-Run Relationships*. Oxford University Press, New York, 2006.

Andreu Mas-Colell, Michael D. Whinston, and Jerry R. Green. *Microeconomic Theory*. Oxford University Press, New York, 1995.

R.B. Myerson. *Game Theory: Analysis of Conflict.* Harvard University Press, 1997.

Martin J. Osborne and Ariel Rubinstein. *A Course in Game Theory.* MIT Press, Cambridge, Massachusetts, 1994.

S. Sorin. On repeated games with complete information. *Mathematics of Operations Research*, pages 147–160, 1986.