Repeated Games

Introduction

One striking feature of many one-shot games we study (e.g., the Prisoners' Dilemma) is that the Nash equilibria are so noncooperative: each player would prefer to fink than to cooperate. Real life is not a one-shot game; in fact, it's not even a disconnected series of one-shot games. Real life is a bigger game in which what a player does early on can affect what others choose to do later on. Repeated games can incorporate phenomena which we believe are important but which aren't captured when we restrict our attention to static, one-shot games. In particular we can strive to explain how cooperative behavior can be established as a result of rational behavior.

We will develop a useful formalism, the semiextensive form, for analyzing repeated games, i.e. those which are repetitions of the same one-shot game (called the stage game). We will describe strategies for such repeated games as sequences of history-dependent stage-game strategies. The payoffs to the players in this repeated game will be functions of the stage-game payoffs. We will define the concept of Nash equilibrium and—after identifying the subgames in this formalism—the concept of a subgame-perfect equilibrium for a repeated game.

We will prove one result which applies to both finitely and infinitely repeated games, which makes it easy to identify at least some of the Nash equilibria of a repeated game: any sequence of stage-game Nash equilibria is a subgame-perfect repeated-game equilibrium.

We will then focus our attention on finitely-repeated games. We will exploit the existence of a final period—which wouldn't exist in an infinitely-repeated game—to establish a necessary condition for a repeated-game strategy profile to be a Nash equilibrium: the last period's play must be Nash on the equilibrium path.

We then turn to subgame perfection within the finitely-repeated game context and prove two theorems about subgame-perfect equilibria. The first is a necessity theorem which is much stronger than the one we obtained for Nash equilibria: the last period's play must be Nash even off the equilibrium path. The second pertains to the special case in which the stage game has a unique Nash equilibrium payoff.[1] In this case the subgame-perfect equilibria of the repeated game are any period-by-period repetitions of the stage-game Nash equilibria.

Repeating the Stage Game

Consider a game G (which we'll call the *stage game* or the *constituent game*). As usual we let the player set be $I=\{1,\ldots,n\}$. In our present repeated-game context it will be clarifying to refer to a player's stage-game choices as *actions* rather than strategies. (We'll reserve "strategy" for choices in the repeated game.) So each player has a pure-action space A_i. The space of action profiles is $A=\times_{i\in I}A_i$. Each player has a von Neumann-Morgenstern utility function defined over the outcomes of G, $g_i\colon A\to\mathbb{R}$. (I want to reserve "u" for the payoff to the entire repeated game.)

Let G be played several times (perhaps an infinite number of times) and award each player a payoff which is the sum (perhaps discounted) of the payoffs she got in each period from playing G.[2] Then this sequence of stage games is itself a game: a *repeated game* or a *supergame*.[3]

1 For example, the Prisoners' Dilemma or a Cournot duopoly with linear demand and constant marginal costs.

2 This is the typical assumption, although a weaker assumption is sufficient for most results: What is really necessary is that every player's ultimate payoff be an additively separable function of her per-period payoffs.

3 I have conflicting reports on whether supergame is a more general concept than repeated game or vice versa. Friedman [1990: 108] says that "A repeated game is a supergame in which the same (ordinary) game is played at each iteration." On the other side, Mertens [1987: 551] says "'Supergame' is the original name for situations where the same game is played repetitively.… Repeated game is used for more general models.…"

Two statements are implicit when we say that in each period we're playing the same stage game:

a For each player the set of actions available to her in any period in the game G is the same regardless of which period it is and regardless of what actions have taken place in the past.

b The payoffs to the players from the stage game in any period depend only on the action profile for G which was played in that period, and this stage-game payoff to a player for a given action profile for G is independent of which period it is played.[4]

Statements a and b are saying that the environment for our repeated game is *stationary* (or, alternatively, independent of time and history).[5] This does *not* mean the actions themselves must be chosen independently of time or history.

We'll limit our attention here to cases in which the stage game is a one-shot, simultaneous-move game. Then we interpret a and b above as saying that the payoff matrix is the same in every period. Our stationarity assumption rules out some economically important situations, such as those which involve investment or learning. Modeling these situations requires a more general dynamic-game framework.

We make the typical "observable action" or "standard signalling" assumption that the play which occurred in each repetition of the stage game is revealed to all the players before the next repetition. Therefore even if the stage game is one of imperfect information (as it is in simultaneous-move games)—so that during the stage game one of the players doesn't know what the others are doing/have done that period—each player *does* learn what the others did before another round is played. This allows subsequent choices to be conditioned on the past actions of other players.

Example: A two-period repetition of the tubular coordination game

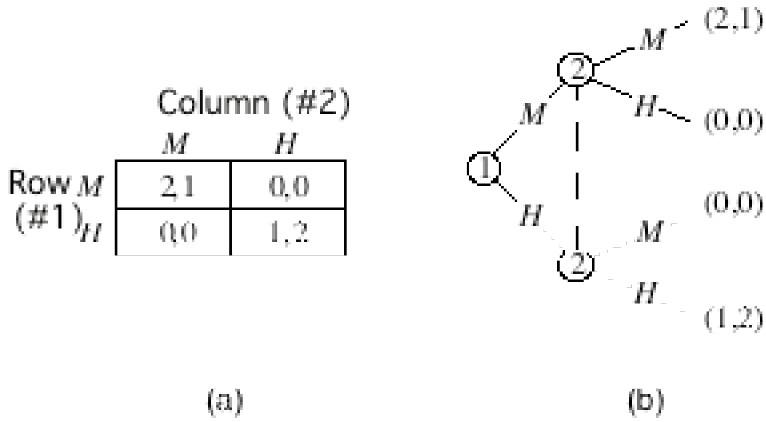
As an example, consider the tubular coordination game in normal form in Figure 1a.[6] Figure 1b depicts an extensive-form game which corresponds to that normal form.

---

4 This doesn't rule out the discounting of payoffs from the stage game. Discounting is performed after the stage game's payoffs are calculated.

5 A dependence on time can be different from a dependence on history. The payoffs of the stage game could change with time in a way that didn't depend on what actions had been taken in the past. This would be time dependence. (For example, the payoff to a player to a particular action profile could be proportional to e^{-rt}.) Alternatively, the payoffs in a period could become double that of the previous period's payoffs if a particular action (applying for a patent, for instance) were taken the previous period *regardless of which particular period the action was taken*. This would be a dependence on history but not on time.

6 To make a long story short…. Tubular refers to the boob tube. M and H refer to MacNeil-Lehrer and Homer Simpson, respectively. Both of these programs are broadcast in Tucson at 7 p.m. Thursday nights.



Figures 1: A tubular coordination game in normal and extensive forms

We consider this game to be our stage game G. If we repeated G twice we would get the extensive-form game shown in Figure 2.[7] Here we have assumed there is no discounting, so the payoffs at each terminal node are the sum of the payoffs from each period.[8] You can see that even just a two-period repetition of the simplest normal-form game imaginable results in a rather complicated extensive-form representation.

The Semiextensive Form

Before we can talk about equilibrium strategies in repeated games, we need to get precise about what a strategy in a repeated game is. As we saw in the above example, even a few repetitions of a simple stage game would yield a very cumbersome game tree. We'll find it useful when studying repeated games to consider the *semiextensive form*.[9] This is a representation in which we accept the normal-form description of the stage game but still want to retain the temporal structure of the repeated game.

---

7 Note that each player has a total of five information sets in this two-period game. One of the information sets is from the first period and four are from the second period. There are four information sets from the second period because there are four different ways the first round could have been played [viz. (U,L), (U,R), (D,L), and (D,R)].

8 Note the italicized stage game payoff vectors with the "+" signs. These indicate the contribution to the players' final payoffs from each stage. Along each path two of these are encountered. Their sum is the payoff vector for the repeated game.

9 This terminology is from Friedman [1990: 109].

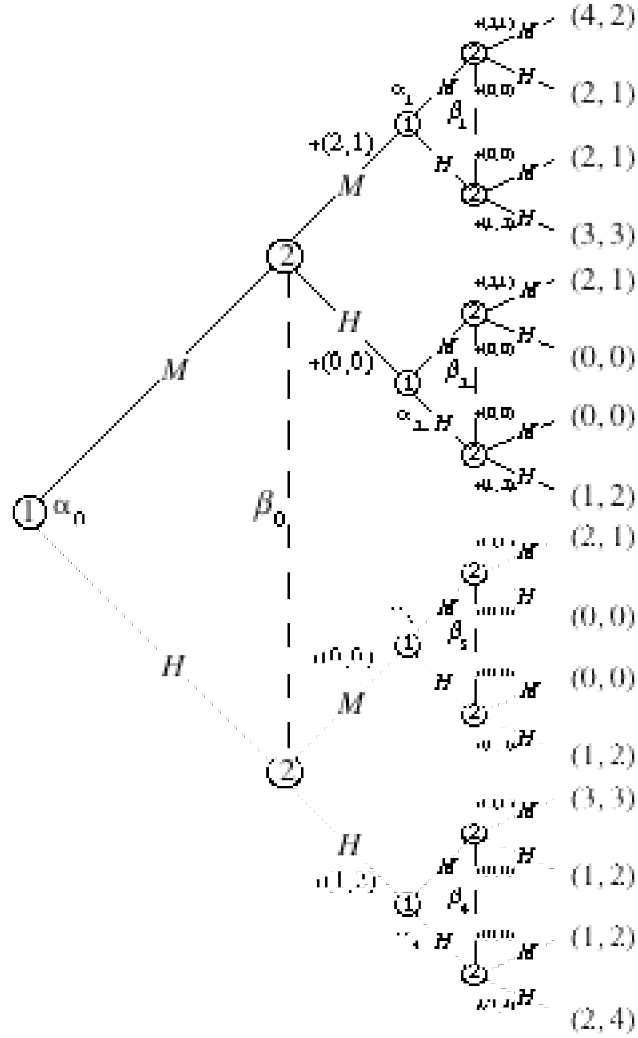


Figure 2: The extensive form for a twice-repeated tubular coordination game.

History and period-t stage-game strategies

I'll try to minimize the notation here, but a certain amount is unavoidable in the pursuit of clarity. We let the first period be labeled $t=0$.[10] The last period, if one exists, is period T, so we have a total of $T+1$ periods in our game. We allow the case where $T=\infty$, i.e. we can have an infinitely repeated game.

We'll refer to the action of the stage game G which player i executes in period t as a_i^t. The *action profile* played in period t is just the n-tuple of individuals' stage-game actions

$$a^t=\left(a_1^t, \ldots, a_n^t\right). \tag{1}$$

We want to be able to condition the players' stage-game action choices in later periods upon actions

10 This simplifies some of the discounting expressions we'll encounter later. However, this choice varies from paper to paper, with some adopting $t=1$ as the first period.

taken earlier by other players. To do this we need the concept of a *history*: a description of all the actions taken up through the previous period. We define the history at time t to be

$$h^t = (a^0, a^1, \ldots, a^{t-1}).^{11} \tag{2}$$

In other words, the history at time t specifies which stage-game action profile (i.e., combination of individual stage-game actions) was played in each previous period. Note that the specification of h^t includes within it a specification of all previous histories $h^0, h^1, \ldots, h^{t-1}$. For example, the history h^t is just the concatenation of h^{t-1} with the action profile a^{t-1}; i.e. $h^t = (h^{t-1}; a^{t-1})$. The history of the entire game is $h^{T+1} = (a^0, a^1, \ldots, a^T)$. Note also that the set of all possible histories h^t at time t is just

$$A^t = \mathop{\times}_{j=0}^{t-1} A, \tag{3}$$

the t-fold Cartesian product of the space of stage-game action profiles A.

To condition our strategies on past events, then, is to make them functions of history.[12] So we write player i's period-t stage-game strategy as the function s_i^t, where $a_i^t = s_i^t(h^t)$ is the stage-game action she would play in period t if the previous play had followed the history h^t. A player's stage-game action in any period and after any history must be drawn from her action space for that period, but because the game is stationary her stage-game action space A_i does not change with time. Therefore we write: $(\forall i \in I)$ $(\forall t)$ $(\forall h^t \in A^t)$ $s_i^t(h^t) \in A_i$. Alternatively, we can write $(\forall i \in I)$ $(\forall t)$ s_i^t: $A^t \to A_i$. The period-t stage-game strategy profile s^t is

$$s^t = \left(s_1^t, \ldots, s_n^t\right). \tag{4}$$

This profile can be described by s^t: $A^t \to A$. I.e. $\forall h^t \in A^t$, $s^t(h^t) = (s_1^t(h^t), \ldots, s_n^t(h^t))$.

Strategies in the Repeated Game

So far we have been referring to stage-game strategies for a particular period. Now we can write, using these stage-game entities as building blocks, a specification for a player's strategy for the repeated game. We write player i's strategy for the repeated game as

$$s_i = \left(s_i^0, s_i^1, \ldots, s_i^T\right), \tag{5}$$

i.e. a $(T+1)$-tuple of history-contingent player-i stage-game strategies. Each s_i^t takes a history $h^t \in A^t$ as its argument. The space S_i of player-i repeated-game strategies is the set of all such $(T+1)$-tuples of player-i stage game strategies s_i^t: $A^t \to A_i$.

We can write a strategy profile s for the repeated game in two ways. We can write it as the n-tuple

11 Again, this can vary from paper to paper. Some would call this history h^{t-1}.

12 In the first period, when $t = 0$, there is no history on which to condition, so the functional dependence is degenerate.

profile of players' repeated-game strategies

$$s = (s_1, \ldots, s_n), \tag{6}$$

as defined in (5). Alternatively, we can write the repeated-game strategy profile s as

$$s = (s^0, s^1, \ldots, s^T), \tag{7}$$

i.e., as a collection of stage-game strategy profiles, one for each period, as defined in (4). These two alternative perspectives of repeated-game strategy profiles are shown schematically in Figure 3.[13]

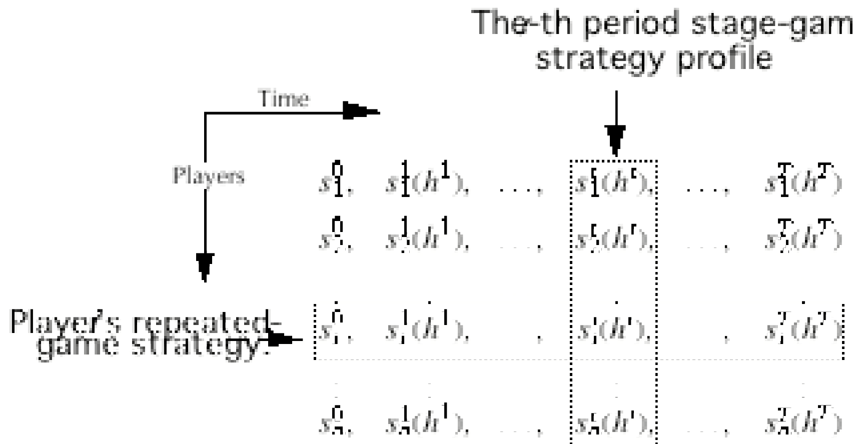


Figure 3: A repeated-game strategy profile viewed alternatively a as players' repeated-game strategies or b as a sequence of stage-game strategy profiles.

Playing the repeated game

In (6) and (7) we have written a repeated-game strategy profile as a collection of functions of histories. It is interesting to note that a repeated-game strategy profile is not itself a function of histories. The reason is simple: Once every one of the individual players' repeated-game strategies s_i is specified (or alternatively once each time period's history-contingent stage-game strategy profile s^t is specified), the sequence of actions, and therefore the history of the entire game, is determined. Perhaps this will become clearer when we see below explicitly how the game is played out.

Let's see how this repeated game is played out once every player has specified her repeated-game strategy s_i. It is more convenient at this point to view this repeated-game strategy profile as expressed in (7), i.e. as a sequence of $T+1$ history-dependent stage-game strategy profiles. When the game starts, there is no past play, so the history h^0 is degenerate: Every player executes her $a_i^0 = s_i^0$ stage-game strategy from (5). This zero-th period play generates the history $h^1 = (a^0)$, where $a^0 = (a_1{}^0, \ldots, a_n{}^0)$. This history is then revealed to the players so that they can condition their period-1 play upon the period-0 play. Each player then chooses her $t=1$ stage-game strategy $s_i{}^1(h^1)$. Consequently, in the $t=1$ stage

13 We will see that the second form, viz. a sequence of stage-game profiles, is more useful when defining the players' payoffs in the repeated game and in proving some theorems characterizing equilibrium in finitely-repeated games. The first form, viz. as a collection of individual repeated-game strategies, is more useful for defining equilibrium in the repeated game.

game the stage-game strategy profile $a^1 = s^1(h^t) = (s_1^1(h^1), \ldots, s_n^1(h^1))$ is played. In order to form the updated history this stage-game strategy profile is then concatenated onto the previous history: $h^2 = (a^0, a^1)$. This new history is revealed to all the players and they each then choose their period-2 stage-game strategy $s_i{}^2(h^2)$. And so it goes.… This process is shown schematically in Figure 4. We say that h^{T+1} is the *path* generated by the repeated-game strategy profile s.

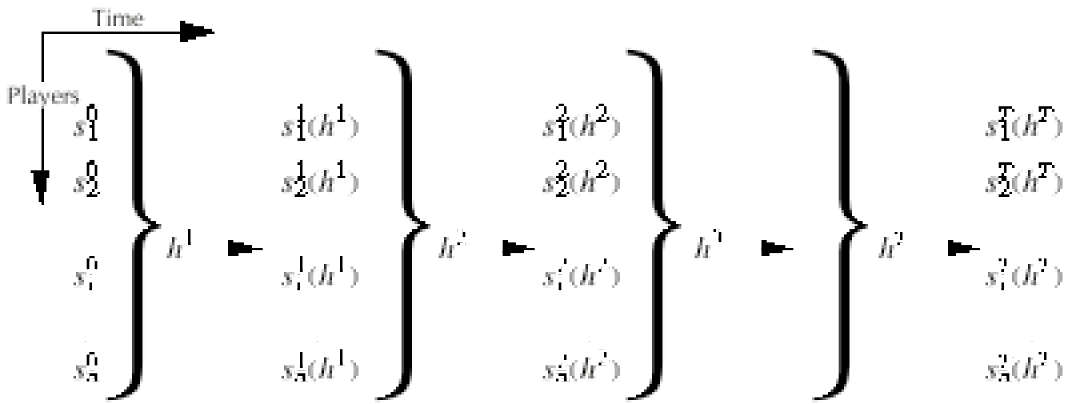


Figure 4: Playing out the repeated game according to a repeated-game strategy.

Payoffs in the Repeated Game

To complete our description of the semiextensive form we need to define the payoffs for this repeated game. Consider some repeated-game strategy profile s and some time period t. As the game is played out according to this repeated-game strategy profile a history h^t of actions up to this period is compiled. The history-contingent stage-game strategy profile to be played this period is $s^t(h^t)$. The payoff to player i from the period-t stage game when each player executes her component of the stage-game strategy profile $s^t(h^t)$ is $g_i(s^t(h^t))$.[14] We define the payoff to player i for the repeated game, when the repeated-game strategy profile s [as expressed in (7)] is played, to be

$$u_i(s) = \sum_{t=0}^{T} \delta^t g_i(s^t(h^t)), \tag{8}$$

where δ is the common discount factor.[15] Note the appearance of the histories h^t in the summand of (8). These are the histories which are generated sequentially as the game is played, as described in Figure 4. I.e. $h^t = (h^{t-1}; s^t(h^{t-1}))$.

14 Note that this g_i doesn't require a superscript "t" because the game is stationary and hence the stage-game payoff functions don't change with time.

15 Assuming a common discount factor is done for simplicity only. The analysis can be carried out allowing the discount factor to vary between players and also across time for each player.

Equilibrium in Repeated Games

Nash Equilibrium in the Repeated Game

Now that we know what strategies and payoffs are in the repeated game, we are in position to make clear what a Nash equilibrium is in this semiextensive form of a repeated game. Now we will use the expression of a repeated-game strategy profile as a collection of players' repeated-game strategies, as in (6). As usual, we say that a strategy profile $\bar{s}$ for the repeated game is a Nash equilibrium if each player's part of $\bar{s}$—i.e., the strategy $\bar{s}_i$, as given by (5)—is a best response given that the other players are playing their parts of $\bar{s}$. More formally, we say that the repeated-game strategy profile $\bar{s}$ is a Nash equilibrium if for all players i

$$\bar{s}_i \in \underset{s_i \in S_i}{\arg\max}\, u_i(s_i, \bar{s}_{-i}). \tag{9}$$

Let $\bar{h}^{T+1}$ be the history generated by the repeated-game strategy profile $\bar{s}$; i.e. let $\bar{h}^{T+1}$ be the path associated with $\bar{s}$. If $\bar{s}$ is a Nash-equilibrium strategy profile, then $\bar{h}^{T+1}$ is its associated *equilibrium path*.

Note from (8) that $u_i(\bar{s})$ depends only upon stage-game strategy profiles played along the equilibrium path. Therefore in a Nash equilibrium each player's repeated-game strategy need only be optimal along the equilibrium path.

Subgame-Perfect Equilibrium in the Repeated Game

We know that a subgame-perfect equilibrium strategy profile $\bar{s}$ is one such that the restriction of $\bar{s}$ to any subgame is a Nash-equilibrium strategy profile in that subgame. We know what a subgame is in the familiar extensive form of a game, but what is a subgame in this semiextensive form?

A subgame, as we know, is a piece of the bigger game which—if we detach it—still makes sense as a game by itself and, if reached, it would be common knowledge that this was the game remaining to be played. Clearly, we could begin our repeated game at some other period, say the k-th period, and the shorter repeated game would still make sense as a game by itself. It is tempting to say, then, that we can find subgames by simply jumping in at any period and playing the game from there. Although tempting, this solution to the "what are the subgames?" problem isn't quite right. We'll see that a subgame of a repeated game is identified by a beginning period *plus* a specification of the history of actions up to that period.

To see why our definition of a subgame of a repeated game must include a specification of the history we can look back at Figure 2, which depicts the extensive form for our twice-repeated tubular coordination game. Clearly each of the second-period repetitions of the stage game (the parts of the tree beginning at the nodes labeled $\alpha_1 \to \alpha_4$) is a subgame. Even though all four instances of this second-period repetition of the stage game look very similar—because they *are* exactly the same stage game—the four do not constitute four clones of a single subgame but rather they are four distinguishable subgames. (For example, they lead to different payoffs.) Note that the history of actions leading to each one is different. For example, the history leading to the repetition corresponding to the information set α_1 is $h^1=(a^0)=(M,M)$, but the history leading to the α_4 repetition is $h^1=(a^0)=(H,H)$. Therefore we see that we need to specify the history of actions in order to identify a particular subgame within the set of those subgames which begin at a particular period. Or another way to see this: When a subgame is reached in the game, the players know the history of play. Therefore when we analyze the subgame, we must allow the players all the information they would have in the original game if they encountered the subgame.

Consider some strategy profile $\bar{s}$ of the original full-length repeated game in extensive form. It specifies an action at each information set and can be written in the form:

$$\bar{s}=(\alpha_0,\alpha_1,\alpha_2,\alpha_3,\alpha_4;\beta_0,\beta_1,\beta_2,\beta_3,\beta_4). \tag{10}$$

In general $\bar{s}$ can specify a different action by player #1 at each of the information sets $\alpha_1\rightarrow\alpha_4$. Similarly, it can specify a different action by player #2 at each of the information sets $\beta_1\rightarrow\beta_4$.

If we want to determine whether $\bar{s}$ is a subgame-perfect equilibrium strategy profile of the full-length repeated game, we need to examine its restrictions to every subgame. The restriction of $\bar{s}$ to the subgame beginning with α_k is of the form (α_k,β_k), $k\in\{1,\ldots,4\}$. Because these (α_k,β_k) pairs are not necessarily the same (because they could be specified independently in $\bar{s}$), the restriction of $\bar{s}$ to each subgame is a possibly different pair of actions. Asking whether $\bar{s}$ is a subgame-perfect strategy profile is the same as asking whether the restriction of $\bar{s}$ to each subgame is a Nash equilibrium of that subgame. But this is the same as asking whether the pair of actions (α_k,β_k) is a Nash-equilibrium strategy profile for the stage game G. Because all four of these (α_k,β_k) pairs can be different, we need to ask this question four times, once for each $k\in\{1,\ldots,4\}$. Therefore it is possible that the restriction of $\bar{s}$ to one subgame will be a Nash equilibrium in the stage game, but that the restriction to a different subgame will not be a Nash equilibrium in that stage game.

We see that it's not enough to say that we're interested in the subgame which begins with period k. There are many subgames corresponding to the repeated game beginning with period k; there are as many such subgames as there are histories up until that point.

How then do we decide whether some repeated-game strategy profile $\bar{s}$ is a subgame-perfect equilibrium? First we must proceed period by period to begin the process of decomposing the repeated game into its subgames. For each period t, then, we must consider every possible history h^t, because there exists a subgame which corresponds to that history. Now that we have identified a subgame beginning at period t with history h^t, we must ask whether the restriction to this subgame of the original repeated-game strategy profile $\bar{s}$ is a Nash equilibrium of this subgame.

A sequence of stage-game Nash equilibria is a repeated-game subgame-perfect equilibrium

It will be useful to distinguish a particularly simple class of repeated-game strategies: open-loop strategies. When a player's strategy depends on the history h^t, we say that it is a *closed-loop* strategy. If for every t her stage game strategy at time t depends only on t but not on previously taken actions, then her strategy is an *open-loop* strategy. To be more explicit… a repeated-game strategy $s_i^t\colon A^t \to A_i$ for player i is an open-loop strategy if

$$(\forall t)\ (\exists a_i^t \in A_i)\ (\forall h^t \in A^t)\ \square\ s_i^t(h^t) = a_i^t.\text{[16]} \tag{11}$$

In this case we can simply write player i's repeated-game strategy as a $(T+1)$-tuple of stage-game actions $s_i = (a_i^0, a_i^1, \ldots, a_i^T)$. If player i is playing an open-loop strategy, then no action of player j's will ever influence a later action by player i. You should be able to show that if all other players are implementing open-loop strategies, then the remaining player might as well play an open-loop strategy as well. (Because all other players' actions are known, and these actions won't be changed as a result of your choices, then you are free to optimize against others' actions on a myopic, period-by-period basis.)

Our first theorem states a sufficient condition for a strategy profile of a repeated game to be a subgame-perfect equilibrium:

Theorem 1 A sequence of history-*independent* (i.e. open-loop) stage-game Nash-equilibrium strategy profiles is a subgame-perfect equilibrium in the (possibly infinitely) repeated game. More formally: Let $\bar{s} = (\bar{a}^0, \ldots, \bar{a}^T)$ be a history-independent strategy profile for the repeated game.[17] If each of the stage-game strategy profiles $\bar{a}^t$ of $\bar{s}$ is a Nash-equilibrium strategy profile for the stage game, then $\bar{s}$ is a subgame-perfect equilibrium strategy profile for the repeated game.

Proof (By contradiction): First we show that $\bar{s}$ is a Nash equilibrium of the repeated game. Assume not. Then for some player i there is an alternative open-loop repeated-game strategy $\hat{s}_i = (\hat{a}_i^0, \ldots, \hat{a}_i^T)$ which is different (in at least one period) from her part $\bar{s}_i = (\bar{a}_i^0, \ldots, \bar{a}_i^T)$ of $\bar{s}$ such that her payoff in the repeated game is higher, given that everyone else plays their parts of $\bar{s}$.[18] In order that her repeated-game payoff be higher, there must be at least one period t in which her stage-game payoff using $\hat{s}_i^t$ is higher than the stage-game payoff she would get from playing $\bar{s}_i^t$. But if that is true, then $\bar{s}_i^t$ could not have been a part of a Nash-equilibrium strategy profile of the stage game, because some other strategy for player i, viz. $\hat{s}_i^t$, would have been better for i. This contradicts the hypothesis that every component of $\bar{s}$ is a Nash equilibrium of the stage game.

Now consider any subgame of this repeated game. For any t and any history h^t, the restriction of $\bar{s}$ to the subgame associated with history h^t still prescribes a history-independent sequence of stage-game Nash-equilibrium profiles. We have already proved that such a sequence is a Nash equilibrium of a repeated game. Therefore the restriction of $\bar{s}$ to any subgame is a Nash equilibrium in that subgame. Therefore $\bar{s}$ is subgame perfect. ☺

---

16 Alternatively… $\forall t,\ \forall h^t, \tilde{h}^t \in A^t,\ s_i^t(h^t) = s_i^t(\tilde{h}^t)$.

17 In other words this strategy profile calls upon each player to take the specified action in each period without regard to what any other player actually did in previous periods.

18 Note that I wrote the alternative strategy $\hat{s}_i$ as an open-loop strategy (i.e. history independent), which is less general than a closed-loop strategy. I gave a sketch of the justification above.

Finitely-repeated games

We now turn to finitely-repeated games. Here, when we say that a game is repeated a finite number of times we not only mean that the game eventually ends but usually also that the stage game is played a given, fixed number of times T and that this number of repetitions is common knowledge. We can exploit the existence of a commonly known last period in order to obtain some results dramatically different from infinitely repeated games.

Nash Equilibria

We now establish a necessary condition for a repeated-game strategy profile $\bar{s}$ to be a Nash equilibrium of the repeated game.

On the equilibrium path the last period's play must be a stage-game Nash equilibrium

Theorem 2 Let $\bar{s}=(\bar{s}^0,\bar{s}^1,\ldots,\bar{s}^T)$ be a Nash-equilibrium strategy profile for the repeated game and let $\bar{h}^T$ be its associated equilibrium path (i.e. the history which is generated when all players play their parts of $\bar{s}$). Then on the equilibrium path the last period's play must be a Nash equilibrium of the stage game G; i.e. the stage-game strategy profile $\bar{s}^T(\bar{h}^T)$ must be a Nash equilibrium of the stage game G.

Proof (By contradiction): Assume that $\bar{s}^T(\bar{h}^T)$ is not a Nash equilibrium of the stage game. Then for some player i, there is a stage-game action $\hat{a}_i\in A_i$ which is different in the last period from her part of $\bar{s}$ along the equilibrium path, i.e., $\hat{a}_i\neq\bar{s}_i{}^T(\bar{h}^T)$, such that she would increase her payoff in period T by playing this other strategy $\hat{a}_i$, rather than playing $\bar{s}_i{}^T(\bar{h}^T)$, given that the other players play $\bar{s}_{-i}{}^T(\bar{h}^T)$. But we see from the definition of the payoff for the repeated game in (8) that if player i increases her payoff in the T-th period (without changing any of her other periods' payoffs), she increases her repeated-game payoff.[19] Therefore, she would do better in the repeated game to play the strategy

$$\hat{s}_i=(\bar{s}_i{}^0,\bar{s}_i{}^1,\ldots,\bar{s}_i{}^{T-1},\hat{a}_i), \tag{12}$$

rather than the entirety of her part of the alleged equilibrium strategy profile. Therefore, contrary to assumption, $\bar{s}$ must not have been a Nash equilibrium of the repeated game. ☺

Subgame-Perfect Equilibria

We learned above that in a Nash equilibrium of a repeated game the play in the last period along the equilibrium path must be a Nash equilibrium in the stage game. If our strategy profile $\bar{s}$ of the repeated

[19] How do we know that her defection from the equilibrium doesn't affect any of her other payoffs? There are no later payoffs to consider, and previous payoffs depend on previous actions which are conditioned at most on *previous* histories. Therefore no action in period T can affect an earlier stage-game payoff.

game is not only a Nash equilibrium but is also a subgame-perfect equilibrium of the repeated game, then we can prove a theorem which states a stronger necessary condition for $\bar{s}$.

Subgame perfection requires that the last period's play be a stage-game Nash equilibrium for every possible history

Theorem 3 In a subgame-perfect equilibrium of the repeated game the last-period play for *any* history (i.e., not just along the equilibrium path) must be a Nash equilibrium of the stage game. More formally: If $\bar{s}=(\bar{s}^0, \bar{s}^1, \ldots, \bar{s}^T)$ is a subgame-perfect equilibrium strategy profile of the repeated game, then, for any conceivable history $h^T \in A^T$, the last period's stage-game strategy profile, $\bar{s}^T(h^T)$, must be a Nash equilibrium of the stage game.

Note the difference between Theorem 3 and Theorem 2. Theorem 2 only required—in order that a repeated-game strategy profile was a Nash equilibrium—that the last period's stage-game strategy profile *along the equilibrium path*—i.e. given that everyone played what she was supposed to play according to $\bar{s}$—was a Nash-equilibrium stage-game strategy profile. Theorem 3 requires—in order that $\bar{s}$ be a subgame-perfect equilibrium strategy profile for the repeated game—that no matter what actually occurs in periods $0 \rightarrow (T-1)$ every player is compelled by $\bar{s}$ to play their part of a stage-game Nash equilibrium in the last period. This theorem is one consequence of subgame perfection's requirement that behavior be optimal even off the equilibrium path.

Proof (By definition): As we have seen, if we specify the history by which we reach the last period, then we have completely specified a subgame which consists solely of one play of the stage game. A subgame-perfect equilibrium of the repeated game must induce Nash-equilibrium behavior in every subgame, therefore in particular its restriction to the last period for any possible history must be a Nash equilibrium of the stage game. ☺

A unique stage-game Nash equilibrium payoff vector is repeated in every repeated-game subgame-perfect equilibrium

A special case arises when the stage game has a unique Nash-equilibrium payoff vector.[20] (Of course, this includes the case where the stage game has a unique Nash equilibrium.) Then we can use a backward-induction argument to show that every subgame-perfect equilibrium strategy profile of the repeated game involves a repetition every period of a Nash-equilibrium stage-game strategy profile and hence a repetition of the unique Nash-equilibrium stage-game payoff vector.

20 I.e. there may be multiple Nash equilibria, but each one yields the same payoff vector.

Theorem 4 Let $\bar{A}$ be the set of stage-game Nash equilibria. Further, let $\bar{A}$ have the property that there exists a payoff vector $\bar{u}\in\mathbb{R}^n$ such that, for all stage-game Nash equilibria $\bar{a}\in\bar{A}$ and for all players $i\in I$, $u_i(\bar{a})=\bar{u}_i$.

Let $\bar{s}=(\bar{s}^0,\bar{s}^1,\ldots,\bar{s}^T)$ be any subgame-perfect equilibrium strategy profile of the repeated game. Then for all t and all $h^t\in A^T$,

$$\bar{s}^t(h^t)\in\bar{A}. \tag{13}$$

Proof Since we know that $\bar{s}$ is subgame perfect, we know that regardless of how we reach the last period, i.e. for whatever $h^T\in A^T$, when we get there a stage-game Nash-equilibrium strategy profile $\bar{s}^T(h^T)\in\bar{A}$ will be played. This yields the payoff vector $\bar{u}$. Therefore the payoffs to each player from the final stage game will be independent of the history, and in particular the payoffs will be independent of the play in the penultimate, $T-1$, period.

For any history h^{T-1}, the restriction of $\bar{s}$ must be a Nash equilibrium in the subgame h^{T-1} defines. Therefore a stage-game Nash-equilibrium strategy profile will be played in this subgame in period $T-1$. (Assume not. Then some player would be able to increase her stage-game payoff in that period. But that would increase her repeated-game payoff as well, because her final-period payoff is independent of anything she chooses now. So if something other than a stage-game Nash equilibrium were played in any subgame beginning in period $T-1$, then the restriction of $\bar{s}$ would not be a Nash equilibrium of the subgame.) So for any history h^{T-1}, the players receive the unique stage-game Nash-equilibrium payoff vector $\bar{u}$ in period $T-1$.

Now their payoffs in the last two periods are determined independently of their previous actions. So the same argument goes through for the $T-2$ period, and so on, establishing that every period's play after any history is a stage-game Nash-equilibrium profile. ☺

Example: A two-period repeated game

Consider the two-player 3×3 stage game in Figure 5. The stage game has two pure-strategy Nash equilibria, viz. (U,l) and (M,m), and importantly these Nash equilibria result in different payoff vectors.[21] Let's determine the set of repeated-game subgame-perfect equilibrium payoffs of the repeated game in which this stage game is repeated twice. (We will simply use the sum of the two per-period payoffs, without discounting, as the repeated-game payoff.)

---

21 If there were only one stage-game Nash equilibrium payoff vector, Theorem 4 would tell us that, in any subgame-perfect repeated-game equilibrium, a stage-game Nash equilibrium would be played in every period for every history.

	l	m	r
U	**5,5**	−1,−1	**−2**,−2
M	−1,−1	**0,0**	**−2**,−2
D	−2,**−2**	−2,**−2**	−6,−6

Figure 5: A two-player game with two pure-strategy Nash equilibria.

A simple punishment structure

We will only consider repeated-game strategy profiles which can be expressed in a particularly simple form: as four pairs of stage-game action profiles of the form

$$s = \langle (A,b),(C,d);(E,f),(G,h)\rangle, \tag{14}$$

where $A,C,E,G \in S_R = \{U,M,D\}$ and $b,d,f,g \in S_C = \{l,m,r\}$. We interpret this repeated-game strategy profile in the following way: The first two stage-game action profiles, viz. $\langle (A,b),(C,d)\rangle$ represent the path of the strategy profile. (In other words, in the first period Row should play A and Column should play b. If both players conformed in the first period, then in the second period Row should play C and Column should play d.)

The third and fourth stage-game action profiles are *punishment* profiles to be played in the second period if one player unilaterally deviates in the first period from the prescribed action profile (A,b); the punishment (E,f) is directed against Row if she deviates unilaterally from the prescribed first-period action profile (A,b); (G,h) is directed against Column if he deviates unilaterally from the prescribed first-period action profile (A,b).

To be perfectly clear… Let the observed first-period action profile be $(X,y) \in S_R \times S_C$. We say that Row unilaterally deviated from (A,b) if ($X \neq A$ and $y=b$); Column unilaterally deviated from (A,b) if ($X=A$ and $y \neq b$). If ($X \neq A$ and $y \neq b$), then neither player unilaterally deviated. [In this case we interpret the strategy profile s as instructing the players to play their on-the-path action profile (C,d) in the second period. In other words, play continues as if they had both conformed in the first period.[22]] If Row unilaterally deviates in the first period (by choosing some $X \neq A$ when Column plays b), then s dictates that in the second period Row plays E and Column plays f. If Column unilaterally deviates in the first period, then in the second period Row should play G and Column should play h.

Equilibrium implications

What would an open-loop repeated-game strategy profile look like in this game? Being open loop implies that the second-period play would be independent of history (i.e. independent of first-period play and therefore independent of whether the players conformed to the first-period prescription). Therefore s is an open-loop repeated-game strategy profile if and only if

[22] It is unimportant what specification we make for the second period following a bilateral deviation in the first period. When we check whether a strategy profile is a Nash equilibrium, we only compare a player's conformity payoff with her payoffs to possible unilateral deviations. Nowhere does her payoff to a multilateral deviation appear in the equilibrium-verification computations.

$$(C,d)=(E,f)=(G,h). \tag{15}$$

Requiring s to be a repeated-game Nash-equilibrium strategy profile puts an immediate restriction on one of the action profiles, viz. the second-period on-the-path action profile (C,d). From Theorem 2, we know that this action profile must be a stage-game Nash-equilibrium action profile. In our example, then, if s is a repeated-game Nash-equilibrium,

$$(C,d)\in\{(U,l),(M,m)\}. \tag{16}$$

Requiring s to be a repeated-game subgame-perfect equilibrium imposes restrictions on the two punishment action profiles, viz. (E,f) and (G,h), as well. [Of course, (C,d) is still constrained by (16) because any subgame-perfect equilibrium is also a Nash equilibrium. Alternatively, (C,d) is still constrained because it is a last-period action profile.] Both of these are action profiles for the last period, albeit corresponding to off-the-path information sets. Theorem 3 tells us that in order for s to be subgame perfect both of these action profiles must be stage-game Nash-equilibrium action profiles. So for s to be subgame perfect it is necessary that

$$(C,d),(E,f),(G,h)\in\{(U,l),(M,m)\}. \tag{17}$$

Subgame-perfect equilibrium payoffs via open-loop strategy profiles

Recall that Theorem 1 gave us a way to identify some of the repeated game's subgame-perfect equilibria: Construct open-loop strategy profiles which always choose stage-game Nash-equilibrium action profiles. For example, we can achieve the repeated-game payoff vector $(10,10)$ as a subgame-perfect equilibrium via the open-loop strategy profile

$$\langle(U,l),(U,l);(U,l),(U,l)\rangle. \tag{18}$$

This is an open-loop sequence of stage-game Nash-equilibrium action profiles according to (15).

Similarly we can support the repeated-game payoff vector $(0,0)$ as a subgame-perfect equilibrium via the open-loop strategy profile

$$\langle(M,m),(M,m);(M,m),(M,m)\rangle. \tag{19}$$

We can achieve the repeated-game payoff vector $(5,5)$ as a subgame-perfect equilibrium via either of the following two open-loop strategy profiles:

$$\langle(M,m),(U,l);(U,l),(U,l)\rangle\quad\text{or}\quad\langle(U,l),(M,m);(M,m),(M,m)\rangle. \tag{20}$$

These three repeated-game payoff vectors, viz. $(10,10)$, $(5,5)$, and $(0,0)$, exhaust the repeated-game payoff vectors which can be supported as subgame-perfect equilibrium payoffs using open-loop strategy profiles.

Non–open-loop subgame-perfect equilibria: general principles

Consider a repeated-game strategy profile which is not an open-loop sequence of stage-game Nash-equilibrium action profiles. By choosing each second-period action profile to be a stage-game Nash equilibrium [i.e. $(C,d),(E,f),(G,h)\in\{(U,l),(M,m)\}$ in accordance with (17)] we ensure that no player will choose to deviate in the second period no matter what transpires in the first period. (When the second period is reached, the remaining game is simply a one-shot play of the stage game. No player has an incentive to unilaterally deviate from the stage-game Nash equilibrium in the second period.)

The challenge, however, is to construct the four action profiles describing a strategy profile so that no player will want to deviate from the prescribed *first*-period action profile. Let's see under what conditions a player would choose to unilaterally deviate from her first-period prescription. A player will deviate in the first period if she prefers a the combination of first-period deviance followed by second-period punishment to b conformity with the path in both periods.

Under our four–action-profile framework in (14), a player is punished for a unilateral deviation independently of precisely which deviation she commits. Therefore we can without loss of generality assume that she picks her myopically most advantageous defection. For example let's consider some proposed strategy profile $s=\langle(A,b),(C,d);(E,f),(G,h)\rangle$. Consider the Row player for definiteness. (The analysis for the Column player is parallel.) Let $\bar{g}_R(b;A)$ be the first-period value of Row's most advantageous stage-game defection from some first-period prescribed action profile (A,b). I.e.

$$\bar{g}_R(b;A)=\max_{X\in\{U,M,D\}\setminus\{A\}} g_R(X,b). \tag{21}$$

If Row unilaterally deviates in the first period, she can secure herself a payoff of $\bar{g}_R(b;A)$ in the first period. Then she will be subjected to the punishment action profile (E,f) in the second period. Row will be willing to conform to the strategy profile s, then, if

$$\bar{g}_R(b;A)+g_R((E,f))\leq g_R((A,b))+g_R((C,d)). \tag{22}$$

For a given prescribed path $\langle(A,b),(C,d)\rangle$, we see from (17) and (22) that we can without loss of generality choose Row's punishment to be the stage-game Nash-equilibrium action profile which is worst for Row. I.e. for Row's punishment we are free to choose (E,f) to solve

$$(E,f)\in\operatorname*{arg\,min}_{(X,y)\in\{(U,l),(M,m)\}} g_R(X,y). \tag{23}$$

"Without loss of generality" here means the following: Let (E',f') be some Row-punishing action profile which gives Row a sufficiently small payoff $g_R((E',f'))$ so that (22) is satisfied and therefore that Row is deterred from a first-period deviation. Then (E,f) as defined in (23) is also such that (22) is satisfied and therefore also deters Row from deviating.

In this game, the same stage-game Nash-equilibrium action profile is the worst for both players, viz. (M,m). Therefore we set

$$(E,f)=(G,h)=(M,m). \tag{24}$$

We want to determine the set of repeated-game payoffs which can be supported as subgame-perfect equilibrium payoffs. To construct candidate subgame-perfect repeated-game strategy profiles we use (24) to define the punishments and (16) to constrain our choice for the second-period on-the-path action profile. Our choice for the first-period on-the-path action profile (A, b) is unconstrained.

The subgame-perfect equilibrium payoffs

We have already seen how to support $(10, 10)$, $(5, 5)$, and $(0, 0)$ as subgame-perfect two-period repeated-game payoffs using open-loop strategy profiles. We can also support as a subgame-perfect equilibrium the two-period payoff vector $(-1, -1)$. The path we choose to support is $\langle (D, r), (U, l) \rangle$ which yields $(-6, -6) + (5, 5) = (-1, -1)$. Using the punishments from (24) gives us the strategy profile $s = \langle (D, r), (U, l); (M, m), (M, m) \rangle$. To verify that this is indeed a subgame-perfect equilibrium we first verify that (17) is satisfied. To check that Row would not choose to deviate in the first period we calculate her most advantageous deviation against Column's action r to be either U or M which yields Row a first-period deviation payoff of $\bar{g}_R(r; D) = -2$. We verify that (22) holds, then, by noting that indeed $-2 + 0 \leq -6 + 5$. An exactly parallel computation confirms that Column would not choose to deviate either.

You can show that we can also support a two-period payoff of $(3, 3)$ as a subgame-perfect equilibrium via the path $\langle (M, r), (U, l) \rangle$.

However, we cannot support $(4, 4)$ because such a two-period payoff can only be achieved by a first-period payoff of $(-1, -1)$ followed by $(5, 5)$.[23] However, a payoff of $(-1, -1)$ requires either (M, l) or (U, m). From either of these action profiles one player would be able to deviate in order to receive 5 in the first period. Because this already exceeds the total conformity payoff, conformity would never be preferred.

We also cannot support a payoff of $(-2, -2)$. This would require a sequence of payoffs on the equilibrium path of $(-2, -2)$ followed by $(0, 0)$. Therefore (M, m) is the implied action profile in the second period on the path; but this is already each player's worst stage-game Nash equilibrium. Therefore there is no effective punishment to wield against a player who defects in the first period, because she's already getting her worst punishment even if she conforms. More formally… from (22), if the second-period on-the-path prescription is also a worst-for-Row Nash equilibrium, i.e. $g_R((C, d)) = g_R((E, f))$, then subgame perfection requires that $g_R((A, b)) \geq \bar{g}_R(b; A)$. I.e. A must be a best response by Row to Column's choice b. When this holds for both players, as it does here, this implies that if the second-period prescription is (M, m) then the first-period prescription must be a stage-game Nash equilibrium.

We can also show that any subgame-perfect equilibrium of the repeated game must give each player at least -2 for a two-period repeated-game payoff. Calculate the minimum value of Row's most advantageous defections for each of Column's three stage-game actions l, m, and r. For example, if Column is playing l, Row can earn at least -1 by defecting; $\min \{ \bar{g}_{\bar{R}}(l; X) : X \in \{U, M, D\} \} = -1$. (You

[23] We can't have $(5, 5)$ followed by $(-1, -1)$ because the second-period payoff must correspond to a stage-game Nash equilibrium.

might say “Row can earn at least *5* by defecting against *l*.” But if Row is supposed to be playing U already, then U would not be a defection.) Similarly $\min\{\bar{g}_{\bar{R}}(m;X):X\in\{U,M,D\}\}=-1$. And against r, $\min\{\bar{g}_{\bar{R}}(r;X):X\in\{U,M,D\}\}=-2$. So, for all $(X,y)\in\{U,M,D\}\times\{l,m,r\}$,

$$\bar{g}_R(y;X)\geq -2. \tag{25}$$

Reconsider the condition (22) for Row’s conformity. Note that the right-hand side is Row’s two-period payoff to the strategy profile s, and that Row’s payoff to her worst stage-game Nash-equilibrium action profile (M,m) is 0. Condition (22) then implies that Row’s two-period payoff in any subgame-perfect equilibrium must exceed $\bar{g}_R(b;A)\geq -2$. An exactly parallel analysis holds for Column.

So we have found all of the subgame-perfect equilibrium payoffs.

References

Friedman, James W. [1990] *Game Theory with Applications to Economics*, second edition, Oxford University Press.

Mertens, Jean-François [1987] “Supergames,” in *The New Palgrave Dictionary of Economics*, eds. John Eatwell, Murray Milgate, and Peter Newman, McMillan Press Limited, pp. 551–552.