Game Theory (continued)

Our discussion so far has been based on the table representation (also called "normal") form of strategies and resulting utilities. One may wonder how games that involve a timing element can be represented. The graph (also called "extensive") form is the answer. As illustration, consider Figure 1 below

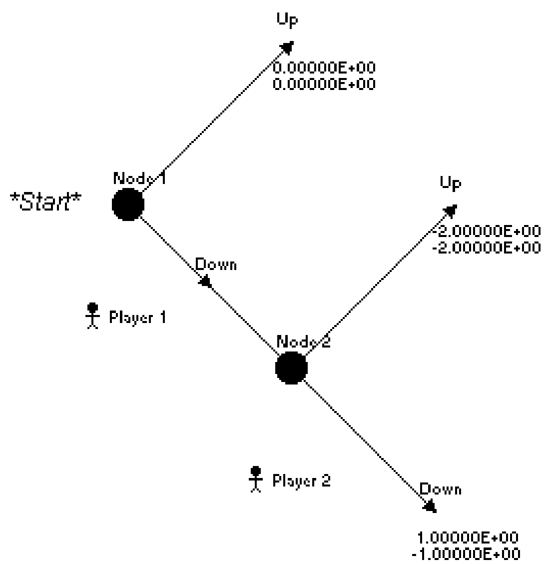

Figure 1

In this case, Player 1 decides first (at Node 1) and makes one of two choices: Up or Down. Up ends the game in the outcome valued 0 by both sides. Down leads to Player 2's turn who may also choose Up or Down with the outcomes listed in the order of the players. What should rationality lead the players to do? Clearly, should Player 2 ever have a choice, she should choose Down. But expecting her to be rational, Player 1 should then choose Down as well. This indeed provides a Nash equilibrium. This extensive form can be converted back to the previous table form as follows

U_1, U_2	Up	Down
Up	$0, 0$	$0, 0$
Down	$-2, -2$	$1, -1$

Clearly, the choice of Up by Player 1 (Row) preempts any kind of decision by Player 2 (Column). But the table translates the idea that Player 2 may pre-commit himself to choosing Up although this won't have any effect on Player 1 should he choose Up. However, this has a subtle and annoying effect: In table form, this game has two Nash equilibria: (Up,Up) and (Down,Down) whereas the game of Figure 1 doesn't seem to have any other than (Down,Down). The (Up,Up) equilibrium seems to involve the non-credible threat by Player 2 to choose Up (at her turn) in order to force Player 1 to choose Up to her benefit. This kind of difficulty has created many developments in game theory and made the extensive form a more appealing vehicle for the analysis.

Sometime, the existence of more than one equilibrium does not create a credibility problem but is just as annoying For instance, the following game is known as the Battle of the Sexes. The story goes as follows: A husband and a wife are planning to go out on Saturday night and must choose between a baseball game and a ballet. Following the usual stereotypes, the husband prefers the baseball and the wife the ballet. However, they prefer to go out together. If they are to buy their tickets independently, four outcomes are possible with the following ratings

U_1, U_2	Baseball	Ballet
Baseball	$2, 1$	$-1, -1$
Ballet	$-2, -2$	$1, 2$

One finds that there are two obvious Nash equilibria to this game (there is also a less obvious mixed one). Evidently, one side will be exploiting the other if they attempt to coordinate their choices. One cannot find a completely satisfactory solution to this game without some additional assumptions on equilibrium *selection*.

In reality, however, a game is often played repeatedly. The husband and wife, for instance, could coordinate their choices and build some fairness by alternating between his and her favorite choice. Indeed, doing so would give them an *average* utility that is strictly better than what they can achieve through any mixed strategy pair of the one-shot game. But this raises the question of how players should value a repeated implementation of a given game. There are two main approaches that I will discuss now.

The Discounted Utility Approach

"In the long run, we'll all be dead" John Maynard Keynes (a prominent economist) used to say. But of course, we are usually concerned about tomorrow's consequences of today's decisions, and also about the day after tomorrow. However, the far enough future tends to be discounted somewhat more than the near future. This decreasing concern for the future can be modeled by a discounted sum of utilities. In our game situation, if $U_i(X)$ is Player i's utility of the choices (pure or mixed) denoted X, one may consider a whole sequence of states $\sigma = (X_t)_{t=0,1,..}$ and give each future $U_i(X_t)$ a weight that decreases with t. A standard approach is to define (for some discount factor $\delta \in (0,1)$)[1]

$$W_i[\sigma = (X_t)_{t=0,1,..}] = \sum_{t=0}^{\infty} \delta^t U_i(X_t) \tag{1}$$

The value of δ indicates the players' concern for the future. δ close to zero means a very shortsighted player while δ close to one shows a long sighted one. But if the above W_i is indeed what matters to Player i, he ought to formulate a whole *sequence* of choices $\sigma_i = (X_{it})_{t=0,1,..}$ in order to optimize it. Indeed, he must also formulate *expectations* of his opponents' own sequence of choices $\sigma_{-i} = (X_{-it})_{t=0,1,..}$. Of course, our standard definition of equilibrium extends easily to this case: a strategy σ_i is best against the others' strategies σ_{-i} if for any alternative $\sigma_i^{'}$ we have

$$W_i[\sigma_i^{'}, \sigma_{-i}] \leq W_i[\sigma_i, \sigma_{-i}] \tag{2}$$

and σ is a Nash equilibrium if this holds for each player i of the game. Interestingly, we immediately have an existence result: if X^* is a Nash equilibrium of the *one-shot* game (on which the repeated game is constructed) then the strategies defined by $X_t \equiv X^*$ form an equilibrium of the discounted repeated game.

But there is much more. Usually, such repeated games have a plethora of equilibria, thus multiplying the selection problem. I will illustrate this on the most well known example. Recall the Prisoner's Dilemma of the last class for which we found that $(Dfct, Dfct)$ was a Nash equilibrium. Consider now the following simple plan for each side: "I will initially cooperate and will continue doing so until I see a defection. Then I will defect forever." This is often called the Grim strategy. It turns out to be a very good Nash equilibrium[2] at least when the players have enough concern for the future.

There are in fact numerous other Nash equilibria in this game. One of great interest is the famous Tit-for-Tat: "I will cooperate at first and then duplicate what my opponent did at the last turn." To see why this is an equilibrium, consider any alternative strategy for Player i. Let's consider i's discounted utility assuming that j plays according to Tit-for-Tat. If we find that it can't be better than zero, we will have our proof.

If we denote by $w_i(X)$ the *best* possible discounted utility for i, at any point in time, when play is currently at state $X \in \{C_iC_j, C_iD_j, D_iC_j, D_iD_j\}$, we may write

$$w_i(C_iC_j) \leq 0 + \delta \times \max\{w_i(C_iC_j), w_i(D_iC_j)\} \tag{a}$$

$$w_i(D_iC_j) \leq 1 + \delta \times \max\{w_i(C_iD_j), w_i(D_iD_j)\} \tag{b}$$

---

[1]Some like to divide (1) by the factor $(1-\delta)$ in order to define weights that add up to 1.
[2]It is also something we like to call a "subgame perfect equilibrium."

$$w_i(C_iD_j) \leq -2 + \delta \times \max\{w_i(C_iC_j), w_i(D_iC_j)\} \quad \text{(c)}$$

$$w_i(D_iD_j) \leq -1 + \delta \times \max\{w_i(C_iD_j), w_i(D_iD_j)\} \quad \text{(d)}$$

If $w_i(D_iC_j) \leq w_i(C_iC_j)$ one easily finds from (a) that $w_i(C_iC_j) \leq 0$ and that all three other estimates are strictly negative provided $\delta > \frac{1}{2}$. If $w_i(D_iC_j) > w_i(C_iC_j)$ then from (c): $w_i(C_iD_j) \leq -2 + \delta w_i(D_iC_j)$. Replacing in (b) and (d) yields

$$w_i(D_iC_j) \leq 1 + \delta \times \max\{-2 + \delta w_i(D_iC_j), w_i(D_iD_j)\} \quad \text{(b')}$$

$$w_i(D_iD_j) \leq -1 + \delta \times \max\{-2 + \delta w_i(D_iC_j), w_i(D_iD_j)\} \quad \text{(d')}$$

Again, for $\delta > \frac{1}{2}$, these two estimates must be negative and all must also be except (a) that may be zero.

The Limit Average Utility Approach

Let us now imagine that the players are very patient and are mostly concerned about the very long run. Or imagine that players eventually survive on the basis of their average scores over the long run. A natural objective would instead be

$$V_i[\sigma = (X_t)_{t=0,1,..}] = \lim_{n \to \infty} \frac{1}{n} \sum_{t=0}^{n} U_i(X_t)$$

This objective will be particularly useful in the so-called *evolutionary* approach. Tit-for-tat and Grim can also be examined under this kind of objective and have similar properties.

Homework

1) Show that the Grim strategy forms a Nash equilibrium in the repeated Prisoner's Dilemma. Hint: any alternative strategy for Player i may defect at one point in time t only to be followed by perpetual defection by the opponent, against which it is best (for i) to perpetually defect. Calculate the resulting discounted utilities and show they are worse than sticking to Grim provided δ is close enough to 1.