The Folk Theorem

November 30th, 2011
Repeated Games

What would happen if we repeatedly play a game many times? Could we get equilibria with higher average payoffs per game?

Consider the Prisoner’s Dilemma.

<table>
<thead>
<tr>
<th></th>
<th>(u₁, u₂)</th>
<th>C</th>
<th>D</th>
</tr>
</thead>
<tbody>
<tr>
<td>C</td>
<td>(-1,-1)</td>
<td>(-10,0)</td>
<td></td>
</tr>
<tr>
<td>D</td>
<td>(0,-10)</td>
<td>(-9,-9)</td>
<td></td>
</tr>
</tbody>
</table>

If we play once, “Defect" is a dominating strategy for both players, so (D, D) is the only NE.
Repeated Games - Fixed Number of Plays

If we play a fixed number of times $T$, then we cannot do better. In fact, the only NE will have each player play $D$ in every game.

We can prove this by backwards induction: In the last game $T$, the players have no incentive to deviate from their dominant strategy. Since the previous games have been played out, they are determining their play based only on the payoffs for the last game, and therefore choose the dominant strategy $D$.

But knowing this, there is no incentive to co-operate in game $T - 1$, since both players will play $D$ in game $T$. So they choose $D$ for game $T - 1$. With repeated application of this argument, we get that each player will play $D$ in every game.
Repeated Games - Infinite Play / Unknown End Time

However, there are ways to get around this. For example

- We can end the game after any round with probability $p \in (0, 1)$.
- We can play the game an infinite number of times, but add a discount factor $\lambda \in (0, 1)$ to the payoffs, so the payoffs for game $t$ are reduced by a factor $\lambda^{t-1}$.

Note that, for the same strategy profile, the expected payoffs in the first case are equal to the payoffs in the second case when $1 - p = \lambda$. 
Extensive-Form Games

Before proving a result in this setting, it is useful to talk about what we mean by “strategies" for repeated games.

- Up until now, we have been considering games with no concept of “turns", "moves" or any sort of time element. Players choose a strategy and then all play that strategy at the same time. This is the *normal form* representation of games.

- But for many games, it makes sense to introduce “moves". We can represent the game as a tree where nodes at level $k$ correspond to the $k$th move. The leaves of the tree mark the end of the game and include a payoff for those choices of moves. This is the *extensive form* representation of a game.
Extensive-Form Games

- We won’t go into the details of the definition of the extensive form. What’s important is that we can represent any extensive form game in the normal form.

- The normal form representation of our game will assign one outcome for every possible choice of all moves in the game.

- So a single strategy $s_i$ encodes a choice for all of player $i$’s turns for each possible choice of the other players’ moves (for perfect information games).

- This can lead to exponentially large payoff matrices. For example, a naive representation (ignoring symmetries) of tic-tac-toe gives a payoff matrix with 255,168 entries.
The Folk Theorem

For our purposes, it is enough to know that we can represent repeated 2x2 games in the normal form, and that a given strategy tells us what to do for all possible responses by the other player.

Now, consider an infinite game with discount factor $\lambda$. We will prove the following theorem

**Folk Theorem:** For sufficiently large $\lambda$, any feasible (undiscounted per-game) payoff vector $\mathbf{x}$ that strictly Pareto-dominates some NE payoff $\mathbf{y}$ is obtainable as an expected outcome of a NE strategy for the repeated game.
A feasible payoff vector is any payoff in the convex hull of payoffs points for the game.

Not every feasible \( \mathbf{x} \) may be obtainable by use of mixed strategies, so the players can either use correlated equilibria or coordinate over several rounds to get expected payoff per game \( \mathbf{x} \).

If the undiscounted per-game payoffs are \( \mathbf{x} \), the the total payoff for each player \( i \) is \( u_i = x_i \sum_{t=1}^{\infty} \lambda^{t-1} = x_i/(1 - \lambda) \).

\( \mathbf{x} \) strictly Pareto-dominates \( \mathbf{y} \) iff \( x_i > y_i \ \forall \ i \).
Left: Feasible payoffs for PD. Right: Feasible payoffs for PD that strictly Pareto-dominate the Nash Equilibrium.
Our NE strategy for the repeated game is the following: 

*Play the single-game strategy which gives payoff \( x \) for as long as the other players also play this strategy. If another player deviates at any time, play the NE which gives \( y \) from that point onward.*

This strategy is sometimes called a *grim trigger* because it punishes any deviation (the “trigger”) for the remainder of the game. \( y_i \) is the threat value by which other players force \( i \) to play along.

We must prove that the grim trigger strategy is a Nash Equilibrium for the repeated game.
Proof

- Note that if all players play grim trigger, then they will get payoffs $x$.

- If player $i$ deviates in game $j$, then the player will get payoff $\lambda^{j-1}z_i$ for that turn for some $z_i$ (The only interesting case is when $z_i > x_i$).

- However, in each future game $k$, that player will receive a payout less than or equal to $\lambda^{k-1}y_i$, because the other players will all play the NE which gives payout $y$, and thus, by definition of a NE, $i$ can do no better than $y_i$.

- Remembering the discount factor, the total change in $i$’s utility is $\lambda^{j-1}[(z_i - x_i) - (x_i - y_i)\lambda/(1 - \lambda)]$ which is less than 0 for large enough $\lambda$. ◼
We can get an even larger space of equilibria for the repeated game by taking a stronger threat point.

Let $S_i$ be the strategy set for player $i$. Define

$$v_i = \min_{s_{-i} \in S_{-i}} \max_{s_i \in S_i} u_i(s_i, s_{-i})$$

The argmin, $s^*_{-i}$, is the worst threat the other players can make against $i$, giving him value $v_i$. Note that $s^*_{-i}$ may be different for each $i$. 

Stronger Threats
We then have the result:

**Folk Theorem II:** For sufficiently large $\lambda$, any feasible (per-game) payoff vector $\mathbf{x}$ that strictly Pareto-dominates payoff $\mathbf{v}$ is obtainable as an expected outcome of a NE strategy for the repeated game.

**Proof:** Same as for Folk Theorem 1.
 Forgiveness

In practice, there is also an incentive not to punish indefinitely. In games of many rounds, with many players, there are other algorithms that do very well, particularly when the other players may be employing different strategies.

Some guidelines for a good strategy:

- Be naturally cooperative (don’t be the first to hurt others)
- Don’t be a pushover (punish strategies which hurt you)
- Show forgiveness (reward strategies that help you)
Tit-For-Tat

For the Prisoner’s Dilemma game, a very simple strategy satisfies all these criteria. It is called *Tit-For-Tat* and is extremely easy to describe

- Play C on the first turn
- Play whatever your opponent last played on every subsequent turn

It’s easy to see that TFT obeys all of our guidelines. In fact, it has shown to be extremely successful in Iterated PD tournaments [see *The Evolution of Cooperation* by Axelrod, ’84]. It also has applications in the study of evolution.
For two player games, we know how to compute threat strategies. Unfortunately, it is NP-complete to find the optimal threats for games with three or more players.

Even worse, we have the following Theorem:

**Theorem:** Finding an equilibrium in a 3-player repeated game is PPAD-complete.
Proof of Theorem

We reduce to a non-repeated (*single-shot*) 2-player game. Let *I* be the row player for the 2-player game and *II* be the column player, with corresponding utilities and strategies.

Introduce the third player and let his strategy set $S_{III}^{rep} = S_I \cup S_{II}$, so he can either choose a row or a column as a strategy. Define the payoffs differently depending on if *III* takes a row or a column. The payoffs $u^{rep} = (u_{1}^{rep}, u_{2}^{rep}, u_{3}^{rep})$ are, in terms of the 2-player payoffs $u_1$ and $u_2$

\[
\begin{align*}
  u^{rep}(r, c, r') &= (u_1(r, c) - u_1(r', c), 0, u_1(r', c) - u_1(r, c)) \\
  u^{rep}(r, c, c') &= (0, u_2(r, c) - u_2(r, c'), u_2(r, c') - u_2(r, c))
\end{align*}
\]
Proof of Theorem

This is a zero-sum game.

But note that I can guarantee a non-negative payoff by playing a best response to II and vice-versa. III can also guarantee a non-negative payoff by playing a best response to either player (and gets positive payoff when the other players fail to play best responses to each other).

So the pareto-dominant NE for this game is for I and II to play an NE in the two player game, and for III to play one of their best responses. Therefore solving for an NE in the 3-player repeated game is at least as hard as finding an NE in the original 2-player game. ■

(Note: The best threat for I and II against III is to play an NE)
The NEs from Folk Theorem I are subgame perfect when we have a dominant NE at $y$: They are also Nash Equilibria for every subgame of the whole game (even on those parts of the tree that are not played).

However, the NEs from Folk Theorem II are not subgame perfect. If a player deviates, the threat strategies are not best responses to each other, so they do not form an NE. Note that this doesn’t cause a problem with the Theorem, because each player assumes the others will follow the trigger strategy and only they can deviate.