Game theory (Ch. 17.5)

<table>
<thead>
<tr>
<th></th>
<th>joke</th>
<th>serious</th>
</tr>
</thead>
<tbody>
<tr>
<td>joke</td>
<td>😊</td>
<td>😞</td>
</tr>
<tr>
<td>serious</td>
<td>😊</td>
<td>😞</td>
</tr>
</tbody>
</table>
Announcements

Midterm Wednesday

HW2 solutions posted on moodle
Game theory

Today we will look at a hard combination of environment factors:

1. Partially observable
2. Multi-agent (competitive primarily)
3. Stochastic
4. Sequential (a little)
5. Static (one easy one!)
6. Discrete (and another!)
Game theory

Typically game theory uses a payoff matrix to represent the value of actions.

The first value is the reward for the left player, right for top (positive is good for both).
A pure strategy is one where a player always picks the same strategy (deterministic).

A mixed strategy is when a player chooses actions probabilistically from a fixed probability distribution (i.e. the percent of time they pick an action is fixed).

If one strategy is better or equal to all others across all responses, it is a dominant strategy.
Dominance & equilibrium

Here is the famous “prisoner's dilemma”

Each player chooses one action without knowing the other's and the is only played once.

<table>
<thead>
<tr>
<th></th>
<th>PRISONER 2</th>
</tr>
</thead>
<tbody>
<tr>
<td>Confess</td>
<td>-8, -8</td>
</tr>
<tr>
<td>Lie</td>
<td>0, -10</td>
</tr>
</tbody>
</table>

<table>
<thead>
<tr>
<th>PRISONER 1</th>
<th></th>
</tr>
</thead>
<tbody>
<tr>
<td>Confess</td>
<td>-10, 0</td>
</tr>
<tr>
<td>Lie</td>
<td>-1, -1</td>
</tr>
</tbody>
</table>
**Dominance & equilibrium**

What option would you pick?

Why?

<table>
<thead>
<tr>
<th></th>
<th>Confess</th>
<th>Lie</th>
</tr>
</thead>
<tbody>
<tr>
<td>Confess</td>
<td>-8, -8</td>
<td>0, -10</td>
</tr>
<tr>
<td>Lie</td>
<td>-10, 0</td>
<td>-1, -1</td>
</tr>
</tbody>
</table>

[Diagram showing outcomes for Dave and Henry]
Dominance & equilibrium

What would a rational agent pick?

If prisoner 2 confesses, we are in the first column... -8 if we confess, or -10 if we lie
--> Thus we should confess

If prisoner 2 lies, we are in the second column, 0 if we confess, -1 if we lie
--> We should confess
Dominance & equilibrium

It turns out regardless of the other player's action, it is in our personal interest to confess.

This is the Nash equilibrium, as any deviation of our strategy (i.e. lying) can result in a lower score (i.e. if opponent confesses).

The Nash equilibrium looks at the worst case and is greedy.
Dominance & equilibrium

Alternatively, a **Pareto optimum** is a state where no other state is strictly better for all players.

If the PD game, \([-8, -8]\) is a Nash equilibrium, but is not a Pareto optimum (as \([-1, -1]\) better for both players)

However \([-10, 0]\) is also a Pareto optimum...
Every game has at least one Nash equilibrium and Pareto optimum, however...

- Nash equilibrium might not be the best outcome for all players (like PD game, assumes no cooperation)

- A Pareto optimum might not be stable (in PD the [-10,0] is unstable as player 1 wants to switch off “lie” and to “confess” if they play again or know strategy)
Find the Nash and Pareto for the following: (about lecturing in a certain csci class)

<table>
<thead>
<tr>
<th>Teacher</th>
<th>Student</th>
</tr>
</thead>
<tbody>
<tr>
<td>prepare well</td>
<td>pay attention</td>
</tr>
<tr>
<td></td>
<td>sleep</td>
</tr>
<tr>
<td>slack off</td>
<td>pay attention</td>
</tr>
<tr>
<td></td>
<td>sleep</td>
</tr>
</tbody>
</table>

Teacher

Student
Find best strategy

How do we formally find a Nash equilibrium?

If it is zero-sum game, can use minimax! (our PD example was not zero sum)

Let's play a simple number game: two players write down either 1 or 0 then show each other. If the sum is odd, player one wins. Otherwise, player 2 wins (on even sum)
Find best strategy

This gives the following payoffs:

<table>
<thead>
<tr>
<th></th>
<th>Pick 0</th>
<th>Pick 1</th>
</tr>
</thead>
<tbody>
<tr>
<td>Pick 0</td>
<td>-1</td>
<td>1</td>
</tr>
<tr>
<td>Pick 1</td>
<td>1</td>
<td>-1</td>
</tr>
</tbody>
</table>

(player 2's value is negative player 1's)

We will run minimax on this tree twice:
1. Once with player 1 knowing player 2's move (i.e. choosing after them)
2. Once with player 2 knowing player 1's move
Find best strategy

Player 1 to go first (max):

If player 1 goes first, it will always lose
Find best strategy

Player 2 to go first (min):

If player 2 goes first, it will always lose
Find best strategy

This is not useful, and only really tells us that the best strategy is between -1 and 1 (which is fairly obvious)

This minimax strategy can only find pure strategies (like PD, except not a zero-sum game)

To find a mixed strategy, we need to turn to linear programming
Find best strategy

First we parameterize this and make the tree stochastic:

Player 1 will choose 0 with probability $p$, and 1 with $(1-p)$

If player 2 always picks 0, this will payoff: $(-1)p + 1(1-p)$

If player 2 always picks 1, this will payoff: $(1)p + (-1)(1-p)$
Find best strategy

Plot these two lines:

\[ U = (-1)p + 1(1-p) \]
\[ U = (1)p + (-1)(1-p) \]

As we maximize, the opponent gets to pick which line to play.

Thus we choose the intersection.
Find best strategy

Thus we find that our best strategy is to play 0 half the time and 1 the other half.

The result is we win as much as we lose on average, and the overall game result is 0.

Player 2 can find their strategy in this method as well, except they want to minimize and player 1 gets to pick the line (in the graph, this is the top triangle and reach same strategy).
Find best strategy

We have two actions, so one parameter (p) and thus we look for the intersections of lines.

If we had 3 actions (rock-paper-scissors), we would have 2 parameters and look for the intersection of 3 planes (2D).

This can generalize to any number of actions (but not a lot of fun).

<table>
<thead>
<tr>
<th></th>
<th>Player 1</th>
<th>Player 2</th>
</tr>
</thead>
<tbody>
<tr>
<td></td>
<td>Stone</td>
<td>Paper</td>
</tr>
<tr>
<td>Stone</td>
<td>(0, 0)</td>
<td>(−1, 1)</td>
</tr>
<tr>
<td>Paper</td>
<td>(1, −1)</td>
<td>(0, 0)</td>
</tr>
<tr>
<td>Scissors</td>
<td>(−1, 1)</td>
<td>(1, −1)</td>
</tr>
</tbody>
</table>
Find best strategy

How does this compare on PD?

Player 1: $p = \text{prob confess}...$

P2 Confesses: $-8p + -10*(1-p)$

P2 Lies: $0p + (-1)*(1-p)$

No cross, but slope positive!
Should pick $p=1$ -> confess
Best strategy in this game?

<table>
<thead>
<tr>
<th></th>
<th>S</th>
<th>C</th>
</tr>
</thead>
<tbody>
<tr>
<td>S</td>
<td>-10,-10</td>
<td>-1,1</td>
</tr>
<tr>
<td>C</td>
<td>1,-1</td>
<td>0,0</td>
</tr>
</tbody>
</table>

Game of Chicken
Repeated games

In repeated games, things are complicated.

For example, in the basic PD, there is no benefit to “lying”.

However, if you play this game multiple times, it would be beneficial to try and cooperate and stay in the [lie, lie] strategy.
Repeated games

One way to do this is the **tit-for-tat strategy**:
1. Play a cooperative move first turn
2. Play the type of move the opponent last played every turn after (i.e. answer competitive moves with a competitive one)

This ensure that no strategy can “take advantage” of this and it is able to reach cooperative outcomes
Repeated games

Two “hard” topics (if you are interested) are:

1. We have been talking about how to find best responses, but it is very hard to take advantage if an opponent is playing a sub-optimal strategy.

2. How to “learn” or “convince” the opponent to play cooperatively if there is an option that benefits both (yet dominated).