A Whirlwind Tour of Game Theory

(Mostly from Fudenberg & Tirole)

Players choose actions, receive rewards based on their own actions and those of the other players.

Example, the Prisoner’s Dilemma:

<table>
<thead>
<tr>
<th></th>
<th>Cooperate</th>
<th>Defect</th>
</tr>
</thead>
<tbody>
<tr>
<td>Cooperate</td>
<td>+3, +3</td>
<td>0, +5</td>
</tr>
<tr>
<td>Defect</td>
<td>+5, 0</td>
<td>+1, +1</td>
</tr>
</tbody>
</table>
Strategies and Nash Equilibrium

A strategy is a specification for how to play the game for a player. A pure strategy defines, for every possible choice a player could make, which action the player picks. A mixed strategy is a probability distribution over strategies.

A Nash equilibrium is a profile of strategies for all players such that each player’s strategy is an optimal response to the other players’ strategies. Formally, a mixed-strategy profile $\sigma^*_i$ is a Nash equilibrium if for all players $i$:

$$u^i(\sigma^*_i, \sigma^{-i}_*) \geq u^i(s^i, \sigma^{-i}_*) \forall s^i \in S^i$$

Nash equilibrium of Prisoner’s Dilemma: Both players defect!
Matching Pennies

<table>
<thead>
<tr>
<th></th>
<th>H</th>
<th>T</th>
</tr>
</thead>
<tbody>
<tr>
<td>H</td>
<td>+1, −1</td>
<td>−1, +1</td>
</tr>
<tr>
<td>T</td>
<td>−1, +1</td>
<td>+1, −1</td>
</tr>
</tbody>
</table>

No pure strategy equilibria

Nash equilibrium: Both players randomize half and half between actions.
More on Equilibria

Dominated strategies: Strategy $s_i$ (strictly) dominates strategy $s_i'$ if, for all possible strategy combinations of opponents, $s_i$ yields a (strictly) higher payoff than $s_i'$ to player $i$.

Iterated elimination of strictly dominated strategies: Eliminate all strategies which are dominated, relative to opponents’ strategies which have not yet been eliminated.

If iterated elimination of strictly dominated strategies yields a unique strategy $n$-tuple, then this strategy $n$-tuple is the unique Nash equilibrium (and it is strict).

Every Nash equilibrium survives iterated elimination of strictly dominated strategies.
Multiple Equilibria

A coordination game:

<table>
<thead>
<tr>
<th></th>
<th>L</th>
<th>R</th>
</tr>
</thead>
<tbody>
<tr>
<td>U</td>
<td>9, 9</td>
<td>0, 8</td>
</tr>
<tr>
<td>D</td>
<td>8, 0</td>
<td>7, 7</td>
</tr>
</tbody>
</table>

$U, L$ and $D, R$ are both Nash equilibria. What would be reasonable to play? With and without coordination?

While $U, L$ is pareto-dominant, playing $D$ and $R$ are “safer” for the row and column players respectively...
Existence of Equilibria

Nash’s theorem, translated: every game with a finite number of actions for each player where each player’s utilities are consistent with the (previously discussed) axioms of utility theory has an equilibrium in mixed strategies.

Idea 1: Reaction correspondences. Player \( i \)’s reaction correspondence \( r_i \) maps each strategy profile \( \sigma \) to the set of mixed strategies that maximize player \( i \)’s payoff when her opponents play \( \sigma_{-i} \). Note that \( r_i \) depends only on \( \sigma_{-i} \), so we don’t really need all of \( \sigma \), but it will be useful to think of it this way. Let \( r \) be the Cartesian product of all \( r_i \). A fixed point of \( r \) is a \( \sigma \) such that \( \sigma \in r(\sigma) \), so that for each player, \( \sigma_i \in r_i(\sigma) \). Thus a fixed point of \( r \) is a Nash equilibrium.

Kakutani’s FP theorem says that the following are sufficient conditions for \( r : \Sigma \rightarrow \Sigma \) to have a FP.
1. $\Sigma$ is a compact, convex, nonempty subset of a finite-dimensional Euclidean space. 
Satisfied, because it's a simplex

2. $r(\sigma)$ is nonempty for all $\sigma$
Each player's playoffs are linear, and therefore continuous, in her own mixed strategy. Continuous functions on compact sets attain maxima.

3. $r(\sigma)$ is convex for all $\sigma$
Suppose not. Then $\exists \sigma', \sigma''$ such that $\lambda \sigma' + (1 - \lambda) \sigma'' \notin r(\sigma)$ But for each player $i$,
\[
u_i(\lambda \sigma'_i + (1 - \lambda) \sigma''_i, \sigma_{-i}) = \\
\lambda \nu_i(\sigma'_i, \sigma_{-i}) + (1 - \lambda) \nu_i(\sigma''_i, \sigma_{-i})
\]
so that if both $\sigma'$ and $\sigma''$ are best responses to $\sigma_{-i}$, then so is their weighted average.
4. $r(\cdot)$ has a closed graph

The correspondence $r(\cdot)$ has a closed graph if the graph of $r(\cdot)$ is a closed set. Whenever the sequence $(\sigma^n, \hat{\sigma}^n) \to (\sigma, \hat{\sigma})$, with $\hat{\sigma}^n \in r(\sigma^n) \forall n$, then $\hat{\sigma} \in r(\sigma)$ (same as upper hemi-continuity)

Suppose that there is a sequence $(\sigma^n, \hat{\sigma}^n) \to (\sigma, \hat{\sigma})$ such that $\hat{\sigma}^n \in r(\sigma^n)$ for every $n$, but $\hat{\sigma} \notin r(\sigma)$. Then there exists $\epsilon > 0$ and $\sigma'$ such that

$$u_i(\sigma'_i, \sigma_{-i}) > u_i(\hat{\sigma}_i, \sigma_{-i}) + 3\epsilon$$

Then, for sufficiently large $n$,

$$u_i(\sigma'_i, \sigma_{-i}^n) > u_i(\sigma'_i, \sigma_{-i}) - \epsilon > u_i(\hat{\sigma}_i, \sigma_{-i}) + 2\epsilon > u_i(\hat{\sigma}_i^n, \sigma_{-i}^n) + \epsilon$$

which means that $\sigma'_i$ does strictly better against $\sigma_{-i}^n$ than $\hat{\sigma}_i^n$ does, contradicting our assumption.
Learning in Games*

How do players reach equilibria?

What if I don’t know what payoffs my opponent will receive?

I can try to learn her actions when we play repeatedly (consider 2-player games for simplicity).

Fictitious play in two player games. Assumes stationarity of opponent’s strategy, and that players do not attempt to influence each others’ future play. Learn weight functions

\[
\kappa_t^i(s^{-i}) = \kappa_{t-1}^i(s^{-i}) + \begin{cases} 
1 & \text{if } s_{t-1}^{-i} = s^{-i} \\
0 & \text{otherwise}
\end{cases}
\]

Calculate probabilities of the other player playing various moves as:

$$\gamma_t^i(s^{-i}) = \frac{\kappa_t^i(s^{-i})}{\sum_{\tilde{s}^{-i} \in S^{-i}} \kappa_t^i(\tilde{s}^{-i})}$$

Then choose the best response action.
Fictitious Play (contd.)

If fictitious play converges, it converges to a Nash equilibrium.

If the two players ever play a (strict) NE at time $t$, they will play it thereafter. (Proofs omitted)

If empirical marginal distributions converge, they converge to NE. But this doesn’t mean that play is similar!

<table>
<thead>
<tr>
<th>$t$</th>
<th>Player1 Action</th>
<th>Player2 Action</th>
<th>$\kappa^1_{T_t}$</th>
<th>$\kappa^2_{T_t}$</th>
</tr>
</thead>
<tbody>
<tr>
<td>1</td>
<td>T</td>
<td>T</td>
<td>(1.5, 3)</td>
<td>(2, 2.5)</td>
</tr>
<tr>
<td>2</td>
<td>T</td>
<td>H</td>
<td>(2.5, 3)</td>
<td>(2, 3.5)</td>
</tr>
<tr>
<td>3</td>
<td>T</td>
<td>H</td>
<td>(3.5, 3)</td>
<td>(2, 4.5)</td>
</tr>
<tr>
<td>4</td>
<td>H</td>
<td>H</td>
<td>(4.5, 3)</td>
<td>(3, 4.5)</td>
</tr>
<tr>
<td>5</td>
<td>H</td>
<td>H</td>
<td>(5.5, 3)</td>
<td>(4, 4.5)</td>
</tr>
<tr>
<td>6</td>
<td>H</td>
<td>H</td>
<td>(6.5, 3)</td>
<td>(5, 4.5)</td>
</tr>
<tr>
<td>7</td>
<td>H</td>
<td>T</td>
<td>(6.5, 4)</td>
<td>(6, 4.5)</td>
</tr>
</tbody>
</table>

Cycling of actions in fictitious play in the matching pennies game
Persistent miscoordination: Players start with weights of \((1, \sqrt{2})\)

<table>
<thead>
<tr>
<th></th>
<th>A</th>
<th>B</th>
</tr>
</thead>
<tbody>
<tr>
<td>A</td>
<td>0, 0</td>
<td>1, 1</td>
</tr>
<tr>
<td>B</td>
<td>1, 1</td>
<td>0, 0</td>
</tr>
</tbody>
</table>

A rule \(\rho^i\) is said to be \(\epsilon\)-universally consistent if for any \(\rho^{-i}\)

\[
\lim_{T \to \infty} \sup_{\sigma^i} \max_{\gamma^i} u^i(\sigma^i, \gamma^i) - \frac{1}{T} \sum_t u^i(\rho^i_t(h_{t-1})) \leq \epsilon
\]

almost surely under the distribution generated by \((\rho^i, \rho^{-i})\), where \(h_{t-1}\) is the history up to time \(t - 1\), available for the decision-making algorithm at time \(t\).
Back to Experts

Bayesian learning cannot give good payoff guarantees.

- Suppose the true way your opponent’s actions are being generated is not in the support of the prior – want protection from unanticipated play, which can be endogenously determined.

- The Bayesian optimal method guarantees a measure of learning something close to the true model, but provides no guarantees on received utility.

- Can use the notion of experts to bound regret!
Define *universal expertise* analogously to universal consistency, and bound regret (lost utility) with respect to the best expert, which is a strategy.

The best response function is derived by solving the optimization problem

\[
\max_{\mathcal{I}^i} \bar{\mathbf{u}}_t^i + \lambda v^i(\mathcal{I}^i)
\]

\(\bar{\mathbf{u}}_t^i\) is the vector of average payoffs player \(i\) would receive by using each of the experts

\(\mathcal{I}^i\) is a probability distribution over experts

\(\lambda\) is a small positive number.

Under technical conditions on \(v\), satisfied by the entropy:

\[-\sum_s \sigma(s) \log \sigma(s)\]

we retrieve the exponential weighting scheme, and for every \(\epsilon\) there is a \(\lambda\) such that our procedure is \(\epsilon\)-universally expert.