Extensions of the Concept of a Rational Solution

- Every maximin solution is a dominant strategy equilibrium
- Every dominant strategy equilibrium is a Nash equilibrium

Cooperation Better for Both: A Dilemma

<table>
<thead>
<tr>
<th>Price Competition</th>
<th>Alpha</th>
<th>Beta</th>
</tr>
</thead>
<tbody>
<tr>
<td></td>
<td>$p = 1$</td>
<td>$p = 2$</td>
</tr>
<tr>
<td>$p = 1$</td>
<td>00</td>
<td>50, -10</td>
</tr>
<tr>
<td>$p = 2$</td>
<td>-10, 50</td>
<td>20, 20</td>
</tr>
<tr>
<td>$p = 3$</td>
<td>-20, 40</td>
<td>10, 90</td>
</tr>
</tbody>
</table>

Dilemmas

- Dilemma: “A situation that requires choice between options that are or seem equally unfavorable or mutually exclusive” --- Am. Her. Dict.
- In game theory: each player acts rationally, but the result is undesirable (less reward)

The Prisoners’ Dilemma

- Devised by Melvin Dresher & Merrill Flood in 1950 at RAND Corporation
- Further developed by mathematician Albert W. Tucker in 1950 presentation to psychologists
- It “has given rise to a vast body of literature in subjects as diverse as philosophy, ethics, biology, sociology, political science, economics, and, of course, game theory.” --- S.J. Hagenmayer
- “This example, which can be set out in one page, could be the most influential one page in the social sciences in the latter half of the twentieth century.” --- R.A. McCain
Prisoners’ Dilemma: The Story

- Two criminals have been caught
- They cannot communicate with each other
- If both confess, they will each get 10 years
- If one confesses and accuses other:
  - confessor goes free
  - accused gets 20 years
- If neither confesses, they will both get 1 year on a lesser charge

Prisoners’ Dilemma Payoff Matrix

<table>
<thead>
<tr>
<th></th>
<th>Ann</th>
<th>Bob</th>
</tr>
</thead>
<tbody>
<tr>
<td>cooperate</td>
<td></td>
<td>defect</td>
</tr>
<tr>
<td>Ann</td>
<td>(-1, -1)</td>
<td>(-20, 0)</td>
</tr>
<tr>
<td>defect</td>
<td>(-20, 0)</td>
<td>(-10, -10)</td>
</tr>
</tbody>
</table>

- defect = confess, cooperate = don’t
- payoffs < 0 because punishments (losses)

Ann’s “Rational” Analysis (Dominant Strategy)

- If cooperates, may get 20 years
- If defects, may get 10 years
- \(\therefore\), best to defect

Bob’s “Rational” Analysis (Dominant Strategy)

- If he cooperates, may get 20 years
- If he defects, may get 10 years
- \(\therefore\), best to defect
Suboptimal Result of “Rational” Analysis

<table>
<thead>
<tr>
<th></th>
<th>Bob</th>
</tr>
</thead>
<tbody>
<tr>
<td>Ann</td>
<td></td>
</tr>
<tr>
<td>cooperate</td>
<td>−1, −1</td>
</tr>
<tr>
<td>defect</td>
<td>0, −20</td>
</tr>
<tr>
<td>cooperate</td>
<td>−20, 0</td>
</tr>
<tr>
<td>defect</td>
<td>−10, −10</td>
</tr>
</tbody>
</table>

• each acts individually rationally ⇒ get 10 years (dominant strategy equilibrium)
• “irrationally” decide to cooperate ⇒ only 1 year

Summary

• Individually rational actions lead to a result that all agree is less desirable
• In such a situation you cannot act unilaterally in your own best interest
• Just one example of a (game-theoretic) dilemma
• Can there be a situation in which it would make sense to cooperate unilaterally?
  – Yes, if the players can expect to interact again in the future

Classification of Dilemmas

General Payoff Matrix

<table>
<thead>
<tr>
<th></th>
<th>Bob</th>
</tr>
</thead>
<tbody>
<tr>
<td>Ann</td>
<td></td>
</tr>
<tr>
<td>cooperate</td>
<td>CC (R) Reward</td>
</tr>
<tr>
<td>defect</td>
<td>DD (P) Punishment</td>
</tr>
<tr>
<td>cooperate</td>
<td>CD (S) Sucker</td>
</tr>
<tr>
<td>defect</td>
<td>DC (T) Temptation</td>
</tr>
</tbody>
</table>
General Conditions for a Dilemma

- You always benefit if the other cooperates:
  - $CC > CD$ and $DC > DD$
- You sometimes benefit from defecting:
  - $DC > CC$ or $DD > CD$
- Mutual coop. is preferable to mut. def.
  - $CC > DD$
- Consider relative size of $CC$, $CD$, $DC$, $DD$
  - think of as permutations of $R$, $S$, $T$, $P$
  - only three result in dilemmas

Three Possible Orders

The three dilemmas: $TRSP$, $RTPS$, $TRPS$

The Three Dilemmas

- Chicken ($TRSP$)
  - $DC > CC > CD > DD$
  - characterized by mutual defection being worst
  - two Nash equilibria ($DC$, $CD$)
- Stag Hunt ($RTPS$)
  - $CC > DC > DD > CD$
  - better to cooperate with cooperator
  - Nash equilibrium is $CC$
- Prisoners’ Dilemma ($TRPS$)
  - $DC > CC > DD > CD$
  - better to defect on cooperator
  - Nash equilibrium is $DD$

The Iterated Prisoners’ Dilemma

and Robert Axelrod’s Experiments
Assumptions

- No mechanism for enforceable threats or commitments
- No way to foresee a player’s move
- No way to eliminate other player or avoid interaction
- No way to change other player’s payoffs
- Communication only through direct interaction

Axelrod’s Experiments

- Intuitively, expectation of future encounters may affect rationality of defection
- Various programs compete for 200 rounds
  - encounters each other and self
- Each program can remember:
  - its own past actions
  - its competitors’ past actions
- 14 programs submitted for first experiment

IPD Payoff Matrix

<table>
<thead>
<tr>
<th></th>
<th>B</th>
</tr>
</thead>
<tbody>
<tr>
<td></td>
<td>cooperate</td>
</tr>
<tr>
<td>A</td>
<td>3, 3</td>
</tr>
<tr>
<td></td>
<td>5, 0</td>
</tr>
</tbody>
</table>

N.B. Unless DC + CD < 2 CC (i.e. $T + S < 2R$), can win by alternating defection/cooperation

Indefinite Number of Future Encounters

- Cooperation depends on expectation of indefinite number of future encounters
- Suppose a known finite number of encounters:
  - No reason to C on last encounter
  - Since expect D on last, no reason to C on next to last
  - And so forth: there is no reason to C at all
Analysis of Some Simple Strategies

- Three simple strategies:
  - ALL-D: always defect
  - ALL-C: always cooperate
  - RAND: randomly cooperate/defect
- Effectiveness depends on environment
  - ALL-D optimizes local (individual) fitness
  - ALL-C optimizes global (population) fitness
  - RAND compromises

Expected Scores

<table>
<thead>
<tr>
<th>↓ playing</th>
<th>ALL-C</th>
<th>RAND</th>
<th>ALL-D</th>
<th>Average</th>
</tr>
</thead>
<tbody>
<tr>
<td>ALL-C</td>
<td>3.0</td>
<td>1.5</td>
<td>0.0</td>
<td>1.5</td>
</tr>
<tr>
<td>RAND</td>
<td>4.0</td>
<td>2.0</td>
<td>0.5</td>
<td>2.166…</td>
</tr>
<tr>
<td>ALL-D</td>
<td>5.0</td>
<td>3.0</td>
<td>1.0</td>
<td>3.0</td>
</tr>
</tbody>
</table>

Result of Axelrod’s Experiments

- Winner is Rapoport’s TFT (Tit-for-Tat)
  - cooperate on first encounter
  - reply in kind on succeeding encounters
- Second experiment:
  - 62 programs
  - all know TFT was previous winner
  - TFT wins again

Characteristics of Successful Strategies

- Don’t be envious
  - at best TFT ties other strategies
- Be nice
  - i.e. don’t be first to defect
- Reciprocate
  - reward cooperation, punish defection
- Don’t be too clever
  - sophisticated strategies may be unpredictable & look random; be clear
Tit-for-Two-Tats

- More forgiving than TFT
- Wait for two successive defections before punishing
- Beats TFT in a noisy environment
- E.g., an unintentional defection will lead TFTs into endless cycle of retaliation
- May be exploited by feigning accidental defection

Effects of Many Kinds of Noise Have Been Studied

- Misimplementation noise
- Misperception noise – noisy channels
- Stochastic effects on payoffs
- General conclusions:
  - sufficiently little noise ⇒ generosity is best
  - greater noise ⇒ generosity avoids unnecessary conflict but invites exploitation

More Characteristics of Successful Strategies

- Should be a generalist (robust) – i.e. do sufficiently well in wide variety of environments
- Should do well with its own kind – since successful strategies will propagate
- Should be cognitively simple
- Should be evolutionary stable strategy – i.e. resistant to invasion by other strategies

Kant’s Categorical Imperative

“Act on maxims that can at the same time have for their object themselves as universal laws of nature.”