I. The Prisoner's Dilemma
A game with the following actions and payoffs:

<table>
<thead>
<tr>
<th>Player 1</th>
<th>Player 2</th>
</tr>
</thead>
<tbody>
<tr>
<td>Cooperate</td>
<td>Cooperate</td>
</tr>
<tr>
<td></td>
<td>2nd worst, 2nd worst</td>
</tr>
<tr>
<td></td>
<td>worst, best</td>
</tr>
<tr>
<td>Defect</td>
<td>Cooperate</td>
</tr>
<tr>
<td></td>
<td>best, worst</td>
</tr>
<tr>
<td></td>
<td>2nd best, 2nd best</td>
</tr>
</tbody>
</table>

*Notes: “x,y” means payoff x for player 1 and y for player 2. Higher numbers are better.

- What should each player do?
  - Player 1 should defect. Defection dominates cooperation.
  - Player 2 should likewise defect.

- Problem: “defect, defect” outcome is worse for both than “cooperate, cooperate” outcome. Hence the “dilemma.”
- Important: Game theory “games” are not competitions. Goal is to maximize own payoff, not to beat other player.

II. Some Examples
- Real life example: Trade
  - You and I agree to trade my cow for your keg of beer. Somehow, each of us has to give what he has before he knows that the other party has done their part.
  - Outcomes:
    1. I cooperate, you defect: I lose my cow & get nothing. (worst for me)
    2. Both cooperate: Each has something we prefer to what we started with (good)
    3. Both defect: Both keep what we started with (less good)
    4. I defect, you cooperate: I get both the cow & the beer. (Best for me)
- Birds picking ticks off each other’s heads (Dawkins, 207)
  1. Worst outcome: I pick your ticks, you do nothing.
  2. Second worst: No one picks ticks.
  3. Second best: Both pick ticks off each other’s heads.
  4. Best outcome: You pick ticks off my head, I don’t do anything.

III. Sidebar: The PD with known iterations
- Two parties play PD (e.g.) 10 times, # of plays known in advance.
- Your action in one round may affect other player’s actions in future rounds.
- Theoretical prediction: Both parties defect every time.

IV. The indefinitely iterated PD
- PD is to be played many times (unknown #).
- Your action may affect other player’s actions in future rounds.
- What is the best strategy?
  - Nice strategies: are never first to defect
  - Provocable strategies: can be caused to defect by other player’s behavior
Forgiving strategies: do not hold a grudge
Best strategy: Tit for tat
  • Cooperate in round 1.
  • In round \( n+1 \), do what the other player did in round \( n \).
This was tested & confirmed by Robert Axelrod (The Evolution of Cooperation)

Why is this best?
- Avoids extended runs of sucker’s payoffs. Cost: potential 1 round of sucker’s payoff.
- Obtains extended runs of cooperation with other nice players.
- Beats nasty strategies: The payoffs from cooperation outweigh the 1 round of sucker’s payoff.
- Beats pacifist strategy (“always cooperate“): avoids multiple rounds of sucker’s payoffs.

Important points:
- Best strategy is relative to the field of competitors. However, TFT works in many contexts.
- Variation: PD with errors: Suppose a player periodically misinterprets the other player’s action.
  - Then TFT can lead to long runs of mutual recrimination, CD DC CD ...
  - Can be improved by a more forgiving strategy.
- Sidebar: Other ways to increase cooperation
  - Reputation: Players have access to other players’ previous moves in interactions with third parties.
  - Selection: Players can choose whom to play with.
  - Reproduction: Frequency of a program increases with each round that it is successful.

V. Lessons for understanding society
- To be selfish, be nice, forgiving, but provocable.
- Don’t be envious. Goal is to maximize your score, not ‘beat’ your partner. TFT cannot beat its partner.
- Strategies may be conscious or unconscious. “Strategies” are defined purely behaviorally. Hence,
  - Real-life players may be moved by thoughts of “justice”, the virtue of “forgiveness”, etc.
  - But they may behave like rational utility-maximizers.

VI. More real-life examples
- German & British soldiers “cooperating” with mutual non-aggression during WWI.
- The British soccer example.
  - Teams A and B will each move to the next round if they draw against each other. If one loses, the loser is eliminated.
  - Result: Teams start cooperating for a tied score.
- Chess tournaments can have similar cases. Both players agree to a draw.

VII. Postscript
- In a later tournament, other strategies won. The strategy:
  - People from Southampton University submitted a total of 60 programs.
  - Southampton programs were designed to recognize each other in the first 5-10 rounds against each other.
  - After recognition, the programs assume a “master-slave” relationship.
  - The programs defect against non-Southampton programs.
- This result has little theoretical interest. (Why?)
I. Background: IEDS
- Step 1: Look for any strategy by any player that is dominated by some other strategy. Eliminate that strategy.
- Step 2: From the remaining matrix, eliminate any dominated strategies.
- ... continue until no dominated strategies exist.
Example:

<table>
<thead>
<tr>
<th>Player 1</th>
<th>A4</th>
<th>A5</th>
</tr>
</thead>
<tbody>
<tr>
<td>A1</td>
<td>1,1</td>
<td>0,1</td>
</tr>
<tr>
<td>A2</td>
<td>1,0</td>
<td>0,2</td>
</tr>
<tr>
<td>A3</td>
<td>0,-1</td>
<td>0,0</td>
</tr>
</tbody>
</table>

- IEDS solution = A4, A1. (Why?)

II. Nash Equilibrium
- A Nash equilibrium is a set of choices for all the players in a game, such that no individual could do better by switching his strategy, given what all the other players are doing. In other words:
  - Each player plays a best response against a conjecture about other players’ strategies.
  - The conjectures are all correct.
- Any IEDS solution is a Nash Equilibrium. Some Nash equilibria are not IEDS solutions.
Example 1: What are the Nash Equilibria in the game above?

<table>
<thead>
<tr>
<th>Player 1</th>
<th>A1</th>
<th>A2</th>
<th>A3</th>
</tr>
</thead>
<tbody>
<tr>
<td>A4</td>
<td>1,1</td>
<td>1,0</td>
<td>0,-1</td>
</tr>
<tr>
<td>A5</td>
<td>0,1</td>
<td>0,2</td>
<td>0,0</td>
</tr>
</tbody>
</table>

Example 2: This shows that Nash Equilibrium is weaker than IEDS solution:

<table>
<thead>
<tr>
<th>Player 1</th>
<th>A1</th>
<th>A2</th>
<th>A3</th>
</tr>
</thead>
<tbody>
<tr>
<td>A4</td>
<td>-8,-6</td>
<td>-1,-4</td>
<td>7,-4</td>
</tr>
<tr>
<td>A5</td>
<td>-4,-1</td>
<td>4,1</td>
<td>4,-4</td>
</tr>
<tr>
<td>A6</td>
<td>1,2</td>
<td>1,-1</td>
<td>1,-4</td>
</tr>
</tbody>
</table>

- Find the IEDS solution.
- Find the Nash equilibria.
III. Motivating the Nash Equilibrium solution concept
Why would rational parties choose a Nash equilibrium?
• A reasonable prescription for play for all parties. Deviation from the prescription worsens anyone's payoff.
• Preplay communication: Players might try to agree on a set of strategies to play. Any non-Nash set would be vetoed.
• Assume no one makes a “mistake” ± Nash equilibrium
• It's a Schelling point.
• Trial and error (repeated games) ± Nash equilibrium. Other points are unstable.

To reflect on: How successful are these arguments?

IV. Application: Collusion
• A group of airlines all want to see higher prices for air travel. They band together to form a cartel, with an agreement as to the price to be charged per mile. Assume:
  - If everyone sticks to the agreement, everyone does better than they would with the original (competitive) prices.
  - The agreement, being secret and illegal, will not be enforced by the courts. (Why is this important?)
  - If one airline charges slightly lower prices, they get more business & more money. The other airlines lose business.
  - We will consider two players, and three prices they could charge.
    *A side: This is like a Prisoner’s Dilemma, except there can be multiple players & multiple actions.
• Payoff matrix:

<table>
<thead>
<tr>
<th></th>
<th>United</th>
</tr>
</thead>
<tbody>
<tr>
<td></td>
<td>Low price</td>
</tr>
<tr>
<td>Low price</td>
<td>2,2</td>
</tr>
<tr>
<td>Medium price</td>
<td>1,6</td>
</tr>
<tr>
<td>High price</td>
<td>0,4</td>
</tr>
</tbody>
</table>

- Find the Nash equilibrium.
- The same logic applies when you introduce many price levels, and many airlines.
- The process is more certain when there are more players. (Why?)
- It is also more certain when there are more subjective or hard-to-detect ways of chiseling (e.g., giving customers something extra of value when they buy a ticket).
• To consider: What happens if the game is iterated indefinitely many times?
Notes #11: Bargaining Theory

I. Philosophical Relevance of Bargaining Theory

- A popular conception of justice: the terms of a rational social agreement (a kind of social contract), designed to resolve conflicts of interests.
- Two versions:
  1. Justice as mutual advantage: Justice must advantage both/all parties, relative to what happens in the absence of agreement. Takes bargaining power into consideration. Self-interest motivates us to follow justice.
  2. Justice as impartiality: The agreement that would be reached from an impartial standpoint, considering everyone's interests equally.
- Justice as mutual advantage: Requires
  1) A non-agreement point – what is the outcome relative to which justice is judged as "advantageous" or not?
  2) An account of how to choose a solution from among the mutually advantageous possible agreements. Two ideas about this:
     a. Rely on bargaining power & the outcome of rational, self-interested bargaining.
     b. Rely on intuitive notion of fairness, which favors an equal division of the gains.
- (2a) leads us to: Is there such a thing as the rational outcome of a bargaining process?
  - This question is also independently interesting. Identifying such a thing might help us understand many real-world events.

II. Bargaining Problems

- Two rational people need to agree on something. If they do not agree, they both get a payoff of 0. (True by stipulation – definition of the “0” point.)
- There is a range of points (possible agreements) that give both parties payoffs above 0.
- There may also be points that give one party less than zero.
- Q: What outcome will they agree on? What determines the outcome? Traditional answers:
  - They will pick a point that gives both parties >0 payoff.
  - They will not “leave money on the table”: they will pick a Pareto optimal point. Vocabulary word: Pareto optimality/efficiency:
    • This occurs when there is no way to make anyone better off without making someone worse off.
    • A “Pareto improvement” is a change that makes someone better off without making anyone worse off.
    • A state of affairs is Pareto efficient (Pareto optimal, on the Pareto frontier) if no Pareto improvements are available.
  - Beyond that, the solution is indeterminate: Every solution that satisfies those constraints is in both parties’ interests.
- Example: bargaining on the price of a car.
- Example: the game of Bilateral Monopoly:
  - Two people are in a room with $100. Each person must write down on a piece of paper a possible division of the money between them. If they write down the same number, then they get that division. If they do not agree, then both get nothing.
  - Note: Every possible division is a Nash equilibrium.
- But John Nash disagrees with this conclusion: He says there is a determinate rational outcome.
III. The Nash Bargaining Solution

- Background: Von Neumann-Morgenstern utilities
  - Set 0 as the non-agreement utility for both parties.
  - Set 1, for each party, as the maximum utility that party could gain from the interaction. (All the cooperative surplus goes to me!)
  - Important: This does not involve interpersonal utility comparisons! My “1” utility might be a much greater amount of happiness than your “1”. My “1” just means the maximum I could get from the bargaining, and your “1” is the maximum you could get.
  - Intermediate values determined by people’s preferences over lotteries:
    Utility of “x” goes to something such that I would be indifferent between getting that thing for sure, and having a probability of x of getting 1 unit of utility.

- The Nash Solution:
  The agreement point will be that point that maximizes the product of the two parties’ utilities.

Some motivations:

1. Nash argues: It is the unique solution satisfying some formal constraints: (a) Pareto optimal, (b) dominates the non-agreement point, (c) not dependent on the units of utility used, (d) symmetry, (e) independence of irrelevant alternatives.
   - Objection: Not clear how satisfaction of these constraints shows that the solution represents the maximization of each party’s self-interest.

2. Harsanyi says: See the solution as the outcome of a series of offers and counter-offers. The party with more to lose at any given point should be the one to make the next concession. This turns out to lead to the Nash solution.
   - At any point in the negotiations, you can accept the other party’s last offer (and get some utility from that), or hold out for more.
   - Holding out for more carries a risk of conflict (failure to agree).
   - So you should only hold out if the amount you stand to gain outweighs the risk.
   - The person most willing to make a concession is the person for whom the ratio of potential gain to risk is lowest. Harsanyi says this is (u’ - u“)/u’, where u’ is the utility you get if your current offer is accepted and u” is the utility you get if the other player’s current offer is accepted.

- What if you meet a person who sincerely insists on more than the Nash solution?
  - If they can convince you of this, then you should accept their offer.
  - But if you know they are rational, then they could not convince you.

An Example:

- Dividing $1 million between P and R. P has borrowed $100,000 from the Mafia and had no hope of otherwise repaying it.
  - Intuitively, P is in a weak bargaining position. R can hold out for $900k.
    - Aside: But note that P is in a “strong” bargaining position in this slightly different case: Suppose P borrowed $900k from the Mafia.
    - What puts P in a weak position is that he has quickly diminishing marginal utility of money.
    - Note that this has little or nothing to do with P’s being “poor” or R’s being “rich”.

Arbitration:

Parties bargaining about something very important might hire an arbitrator, to reduce the risk of nonagreement.

- Note: the arbitrator must be acceptable to both parties. Hence, each party must expect the arbitrator to make a decision about as good as the party could get by direct bargaining.
- Arbitrator has more leeway when the nonagreement point is worse (or: the stakes are high).
I. The Ultimatum Game

- Two people are in a room with $10. They must divide the money between them. Amounts must be multiples of $1.
- Player 1, the Proposer, proposes a way of dividing the money. (An ultimatum.)
- Player 2, the responder, accepts or rejects the offer.
  - If he accepts, they get that division.
  - If he rejects, they both get nothing.
- What is the game theoretic solution? The subgame perfect equilibrium:
  - Player 1: "$9 for me, $1 for you."
  - Player 2: "Accept."

- Reality:
  - Modal response: "$5 for me, $5 for you." "Accept."
  - If Proposer demands much more than $5, Responder usually rejects.
  - Aggressiveness of Proposer demands increased by:
    i) Use of market terminology ("buying" and "selling").
    ii) Proposer's "earning" his position.

- Why offer more than $1?
  a. Concern for welfare of other player?
  b. Ideal of fairness?
  c. Fear of rejection by responder?
  d. Concern about others' opinion of you?

- These hypotheses can be tested ...

II. The Dictator Game

- Like the Ultimatum, except that Player 2 has no option to reject. Player 1 unilaterally determines division.
  - Result: Dictators are much more aggressive. Some take all $10. But some still offer $3 or $5.
- A variation: As above, except that the offers are "double blind": neither the experimenter nor the other players know which Dictators made which offers.
  - Result: Dictators are extremely aggressive. Most take all $10.

- Conclusions:
  - Explanations (a) and (b) above are not correct. (c) and (d) are correct.
  - But (b) has something to do with it, because offers change when one party "earns" the right to be the proposer.

III. Bilateral Monopoly

- Player 1 and Player 2 must agree on a division of $10. Either party can veto. Non-agreement both get nothing.
- No unique solution in traditional game theory.
- In reality, almost everyone agrees on $5, $5.
- Brinksmanship strategy:
  - Try to commit yourself to a certain offer (favorable to you). Prevent yourself from changing. E.g., make promises.
  - This converts the Bilateral Monopoly into an Ultimatum Game with you as Proposer.
- Theoretically, the other party then should give in.
  - In reality, this will very likely fail. (Recall results of Ultimatum Game above.)

IV. Application: Nuclear Deterrence & Doomsday
- If the Russians attacked the U.S., it would be irrational to retaliate.
  - Game theoretically, either side should have rushed to attack first.
  - They did not. Why?
    a) Perhaps they didn’t hate the other side enough.
    b) They were irrational.
    c) They thought the other side was irrational.
    d) The larger the stakes are, the more risk-averse the players will be. Ex.: the Cuban Missile Crisis. Why did Kruschev back down, instead of running the blockade?
- A solution: Build a doomsday machine. It will trigger automatically if the Russians attack, killing everyone.
  - Even better if we could rig it to go off if the Russians did anything we didn’t like.
  - The other party must know that this is the case.
- Should one really do this, if we can? See point (d) above.
Introduction
Q: Why do people recognize ‘rights’?
   • Moral concept of ‘rights’: “A has a right to do B”. “It would be wrong (or prima facie wrong) to prevent A from doing B.”
   • Legal concept of ‘rights’: “A has a right to do B”. “There’s no law against A doing B.”
   • These don’t explain all rights-respecting behavior:
     - Both fail to explain Britain’s ‘right’ to control Hong Kong or ‘ownership’ of the Falkland Islands.
     - For the moral concept: some aspects of accepted property rights seem arbitrary.
     - For the legal concept: there is some circularity, since the behavior of government people is partly explained by the rights that they have. (Better argument: legal rights depend upon the laws themselves having been made ‘legitimately’, i.e., by people who had the right to make those laws; and how is that to be understood?)
   • Alternative: a positive account of rights: A theory of why people would engage in rights-recognizing behavior, independent of moral beliefs or laws. “Rights ... are a consequence of strategic behavior and may exist with no moral or legal support.” (p. 2)

I. Schelling points, self-enforcing contracts, and the paradox of order
Q: What is the difference between civil society and the ‘Hobbesian state of nature’?
   • [Note: the ‘Hobbesian state of nature’ is a state of social chaos; no rule of law.]
   • Not because of the physical objects present. (Courtrooms, law books, etc.)
   • Not because of the people present. (Police officers, politicians, etc.) For what makes those people act in the socially orderly way?
   • Answer: People in a civil society face a different strategic situation than people in a state of nature. (They have different incentive structures.)

A. Schelling points
   • A type of problem: When there is an advantage in coordinating, but people can’t communicate, how can they coordinate?
   • A Schelling point is a solution that people will tend to converge on in the absence of communication, because it seems natural or ‘special’ to them.
     - Ex.: You and a friend have to try to pick the same number from the following sequence:
       2, 5, 9, 25, 69, 73, 82, 96, 100, 126, 150
       Which number do you pick?
   • The game of bilateral monopoly:
     Two people are in a room with $100. They may keep the money if they can both agree on how to divide it between them. If they can’t agree, then neither gets anything.
     - In terms of pure game theory, any division (other than 0-100) benefits both parties, so both parties should be willing to accept it. Should you insist on 99-1 in your favor?
     - Note how this is similar to the above game. What division should you propose/accept?
   • Schelling points also provide alternatives to continued bargaining.

B. Up from Hobbes
   • Imagine 2 people in a state of nature.
- Both want to avoid conflict. (Similar to above ‘bilateral monopoly’ game.)
- Best way is to agree on a system of rights, esp. property rights.

• The Schelling point:
  - May use some natural boundary to divide the land. Neither party pays tribute to the other.
  - Any previous agreement is thereafter itself a Schelling point.

• The establishment of the agreement does not alter our physical situation or physical power. But it alters the strategic situation. Neither party violates the agreement, because that would return them to the ‘Hobbesian jungle.’ If one party violates the agreement, the other party fights, because allowing the violation “implies unlimited demands.” If A allows B to steal from him, A loses not only that property, but also the advantage of having agreed-upon property rights.

• Important: The contract enforces itself, without either moral beliefs or legal sanctions.

II. Two routes from Hobbes to here

• Main idea:
  - There is a process of evolution of norms, whereby more efficient rules win out.
  - It produces locally efficient but not necessarily globally efficient norms. I.e., a norm will not be adopted if its benefits depend upon almost everyone adopting it. It will be adopted if it benefits small groups who adopt it.

III. Law, justice, and efficiency

• The 3-way coincidence: the following tend to be about the same:
  1. The (intuitively) morally correct rules
  2. The economically efficient rules
  3. The rules that are actually in effect

   Why this coincidence?

• (II) above explains why 2 & 3 tend to coincide.

• Good question: why does (1) correspond with (2) and (3)? What, if anything, does this show about the nature of morality? Think about this.
Phil. 4800  
Notes #14: Commons, Public Goods, & Externalities

I. The Tragedy of the Commons
    • There is a common grazing area. Many ranchers have free access.
      - Ranchers profit according to # of cattle they have.
      - If land is overgrazed, it becomes useless to everybody.
    • What will happen?
      - Each rancher gets 100% of the benefit of each cow he adds to the land.
      - He gets 1/n of the cost, where n is the # of ranchers.
      - He “should” then put as many cows on the land as possible.
      - Problem: The land becomes overgrazed & worthless. Maybe he should restrain himself? Problem:
        - If one other rancher overgrazes, that rancher gets all the benefit, and everyone else is a “sucker”.
        - The larger the group, the more certain that this happens.
        - If one rancher thinks someone else will overgraze, then he “should” do so first.
        - No natural rule for how much each can graze (Schelling point).
    • Analogies:
      - Population & overbreeding.
      - Pollution.
      - Use of National Parks.

II. Public Goods
    • Like the tragedy of the commons, except with something good.
      - Non-excludable: The good must be provided either to all or to none of the members of some
        pre-existing good ± Provider cannot collect money proportional to the value of the good. ±
        Provider bears the costs, other people get the benefits.
      - Example: A dam could protect a valley from flooding.
        - Anyone who builds it pays the cost.
        - The benefit goes to everyone in the valley.
        - Problem: no one will individually choose to build it.

III. Externalities
    • Positive externalities: Benefits your action has that you can’t charge people for. (Like public goods.)
    • Negative externalities: Harms your action causes that you don’t have to pay for.
    • Problem:
      - Positive externalities are underproduced.
      - Negative externalities are overproduced.

IV. Solutions
    A. Appeal to conscience?
      • This won’t work because people will not want to be suckers.
      • Also, it causes anxiety.
      • And conscience will be selected out by cultural/biological evolution.
    B. Private property
      • For the ranchers, this would internalize the externalities.
      • Problems: What about population? Air pollution? Oceans?
    C. “Mutual coercion, mutually agreed upon.”
• We appoint an authority to hurt people who produce negative externalities.
  - We could prohibit harmful actions.
  - Or we could tax them. This is better, when the optimal amount is nonzero.
• Problem: Who will watch the watchers?
  - We must “invent corrective feedbacks to keep custodians honest”. (Doesn't say what these would be.)

V. The Problem Hardin Doesn’t Notice
“Mutually agreed upon coercion” creates another tragedy of the commons.
• Custodian may:
  - Exploit the resource for his own benefit (directly or indirectly): Benefits go to custodian. Harms go to society.
  - Protect resource responsibly: Benefits go to society. Costs go to custodian.
  - What will he do?
• Response: The public will watch over the authorities.
• Problem: Each citizen may:
  - Carefully watch over the authorities: Benefits go to society. Costs go to citizen.
  - Ignore authorities, let other people do the watching: Benefits go to citizen, costs to society.
  - What will citizen do?

VI. Schmidtz on Jamestown & Customs
• Jamestown: First permanent settlement in America, 1607 on.
• Most colonists starved. Why?
  - Food was a commons: Any food was to be shared equally. ± Food is a positive externality.
  - Human “skeletons” were standing in the streets bowling, waiting for someone else to plant crops.
• Governor Dale in 1614 divides land into private plots.
  - Production rises 7-fold.
  - Why? Internalized positive externalities.
• An alternative to conscience, chaos, and government: Custom. Works best when
  - Group is very interdependent.
  - They can easily monitor.
  - They get frequent feedback.
• Difference between Hutterites and Jamestown?
  - In Jamestown, you get food no matter what.
    I think Schmidtz’ point is: The consequences of shirking must be more than disapproval. Like, starvation. The rewards for good behavior should be tangible.
  - Note: Punishments might be enacted only rarely. But they must be there.
Phil. 4800
Unit 2 Review

At the end of unit 2, students should know:

These games, their payoffs, & their solutions:
Prisoner’s Dilemma
   & iterated version
   & indefinitely iterated
Ultimatum Game
Dictator Game
Bilateral Monopoly

These solution concepts:
Dominance
IEDS
Nash Equilibrium
Nash bargaining solution

Other concepts:
Schelling points
Pareto improvements/ optimality/ efficiency
Public goods
Negative/ positive externalities
Tit-for-tat strategy
   Nice strategies
   Forgiving strategies
   Provocable strategies

These examples & what they illustrate:
British & German WWI soldiers
Collusion in a market
Doomsday device
Tragedy of the commons
   & pollution
   & population
   & Jamestown

These people & their main views:
Dawkins on PD
Nash on bargaining
Friedman on property rights
   & why we observe them
Hardin
   On commons problem
   On appeal to conscience
   Solution to commons problem
Schmidtz on Jamestown
   & problem w/ it

Note: Working w/ Payoff Matrices:
You will have to do it. E.g., correlate them with types of games, work out Nash equilibria, etc.