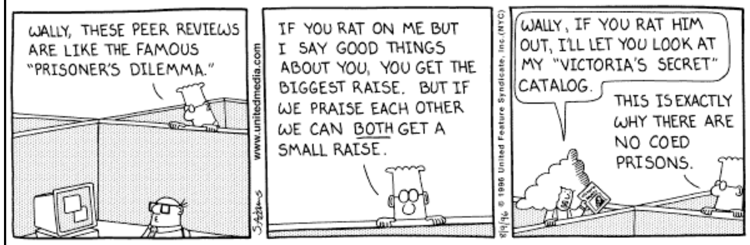
A Glimpse of Game Theory



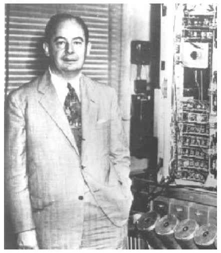
Basic Ideas of Game Theory

- Game theory studies the ways in which strategic interactions among rational players produce outcomes with respect to the players' preferences (or utilities)
  - The outcomes might not have been intended by any of them.
- Game theory offers a general theory of strategic behavior
- Generally depicted in mathematical form
- Plays an important role in modern economics as well as in decision theory and in multi-agent systems

Games and Game Theory

- Much effort has been put into getting computer programs to play artificial games like chess or poker that we commonly play for entertainment
- A larger issue is accounting for, modeling, and predicting how an agent (human or artificial) can or should interact with other agents
- Game theory can account for or explain a mixture of cooperative and competitive behavior
- It's applies to zero-sum games as well as non zero-sum games.

Game Theory



- Modern game theory was defined by von Neumann and Morgenstern

  von Neumann, J., and Morgenstern, O., (1947). The Theory of Games and Economic Behavior. Princeton: Princeton University Press, 2nd edition.
- It covers a wide range of situations, including both cooperative and non-cooperative situations
- Traditionally been developed and used in economics and in the past 15 years been used to model artificial agents.
- It provides a powerful model, with various theoretical and practical tools, to think about interactions among a set of autonomous agents.
- And is often used to model strategic policies (e.g., arms race)

Zero Sum Games

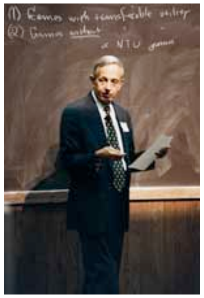
- **Zero-sum** describes a situation in which a participant's gain (or loss) is exactly balanced by the losses (or gains) of the other participant(s)
- The total gains of the participants minus the total losses always equals 0
- Poker is a zero sum game
  - The money won = the money lost
- Trade is not a zero sum game
  - If a country with an excess of bananas trades with another for their excess of apples, both may benefit from the transaction
- Non-zero sum games are more complex to analyze
- We find more non-zero sum games as the world becomes more complex, specialized, and interdependent

Rules, Strategies, Payoffs, and Equilibrium

Situations are treated as games.

- The **rules** of the game state who can do what, and when they can do it
- A player's **strategy** is a plan for actions in each possible situation in the game
- A player's **payoff** is the amount that the player wins or loses in a particular situation in a game
- A players has a **dominant strategy** if his best strategy doesn't depend on what other players do

Nash Equilibrium



- Occurs when each player's strategy is optimal, given the strategies of the other players
- That is, a strategy profile where no player can strictly benefit from unilaterally changing its strategy, while all other players stay fixed
- Every finite game has at least one Nash equilibrium in either pure or mixed strategies, a result proved by John Nash in 1950
  - J. F. Nash. 1950. Equilibrium Points in n-person Games. Proc. National Academy of Science, 36, pages 48-49.
  - Nash won the 1994 Nobel Prize in economics for this work
  - Read "A Beautiful Mind" by Sylvia Nasar or see the film.

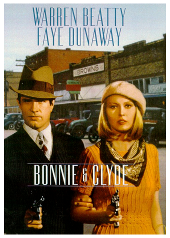
Prisoner's Dilemma


Will the two prisoners cooperate to minimize total loss of liberty or will one of them, trusting the other to cooperate, betray him so as to go free?

- Famous example of game theory
- Strategies must be undertaken without the full knowledge of what other players will do
- Players adopt dominant strategies, but they don't necessarily lead to the best outcome
- Rational behavior leads to a situation where everyone is worse off

---

Bonnie and Clyde



- Bonnie and Clyde are arrested by the police and charged with various crimes. They are questioned in separate cells, unable to communicate with each other. They know how it works:
  - If both resist interrogation (cooperating with each other) and proclaim mutual innocence, they will get a three year sentence for robbery
  - If one confesses (defecting) to all the robberies and the other doesn’t (cooperating), the confesser is rewarded with a light, 1-year sentence and the other will get a severe 8-year sentence
  - If they both confess (defecting), then the judge will sentence both to a moderate four years in prison
- What should Bonnie do? What should Clyde do?

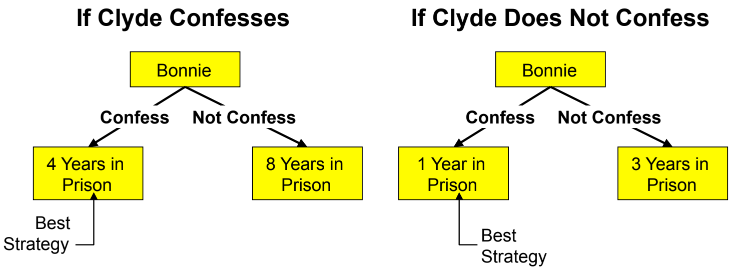
The payoff matrix

		CLYDE	
		Confess	**Not Confess**
BONNIE	**Confess**	4 years each	1 year for Bonnie and 8 years for Clyde
	Not Confess	8 years for Bonnie and 1 year for Clyde	3 years each

---

Bonnie’s Decision Tree

There are two cases to consider


If Clyde Confesses

Bonnie

Confess → 4 Years in Prison (Best Strategy)

Not Confess → 8 Years in Prison

If Clyde Does Not Confess

Bonnie

Confess → 1 Year in Prison (Best Strategy)

Not Confess → 3 Years in Prison

The dominant strategy for Bonnie is to confess (defect) because no matter what Clyde does she is better off confessing.

So what?

- It seems we should always defect and never cooperate
- No wonder Economics is called the dismal science

Some PD examples

There are lots of examples of the Prisoner's Dilemma in the world

- Cheating on a cartel
- Trade wars between countries
- Arms races
- Advertising
- Communal coffee pot
- Class team project

Prisoner's dilemma examples

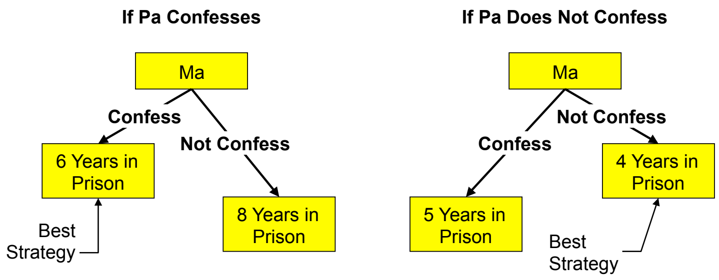
- **Cheating on a Cartel**
  - Cartel members' possible strategies range from abiding by their agreement to cheating.
  - Cartel members can charge the monopoly price or a lower price.
  - Cheating firms can increase profits
  - The best strategy is charging the low price
- **Trade Wars Between Countries**
  - Free trade benefits both trading countries
  - Tariffs can benefit one trading country
  - Imposing tariffs can be a dominant strategy and establish a Nash equilibrium even though it may be inefficient
- **Advertising**
  - The prisoner's dilemma applies to advertising
  - All firms advertising tends to equalize the effects
  - Everyone would gain if no one advertised

Games Without Dominant Strategies

- In many games the players have no dominant strategy.
- Often a player's strategy depends on the strategies of others.
- If a player's best strategy depends on another player's strategy, he has no dominant strategy.

		Pa	
		Confess	**Not Confess**
Ma	**Confess**	6 years for Ma 1 year for Pa	5 years for Ma 3 years for Pa
	Not Confess	8 years for Ma 0 years for Pa	4 years for Ma 2 years for Pa

Ma's Decision Tree



Ma has no explicit dominant strategy, but there is an *implicit* one since Pa does have a dominant strategy.

Some games have no simple solution

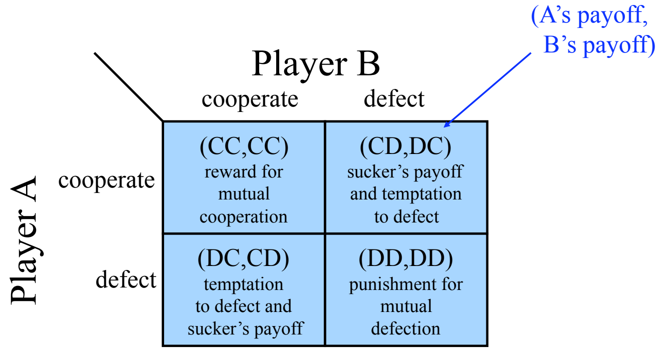
In the following payoff matrix, neither player has a dominant strategy. There is no non-cooperative solution

		Player B	
		1	2
Player A	1	1, -1	-1, 1
	2	-1, 1	1, -1

Repeated Games

- A repeated game is a game that the same players play more than once
- Repeated games differ form one-shot games because people's current actions can depend on the past behavior of other players
- Cooperation is encouraged

Payoff matrix for the generic two person dilemma game



(A's payoff, B's payoff)

		Player B	
		cooperate	defect
Player A	cooperate	(CC,CC) reward for mutual cooperation	(CD,DC) sucker's payoff and temptation to defect
	defect	(DC,CD) temptation to defect and sucker's payoff	(DD,DD) punishment for mutual defection

Payoffs

- There are four payoffs involved
  - CC: Both players cooperate
  - CD: You cooperate but other defects (aka “sucker’s payoff”)
  - DC: You defect and other cooperates (aka “temptation to defect”)
  - DD: Both players defect
- Assigning values to these induces an ordering, with 24 possibilities (4!); three lead to “dilemma” games

  Prisoner’s dilemma: DC > CC > DD > CD
  Chicken: DC > CC > CD > DD
  Stag Hunt: CC > DC > DD > CD

Chicken

- DC > CC > CD > DD
- Rebel without a cause scenario
  - Cooperation: swerving
  - Defecting: not swerving
- The optimal move is to do exactly the opposite of the other player

Stag Hunt

- CC > DC > DD > CD
- Two players on a stag hunt
  - Cooperating: keep after the stag
  - Defecting: switch to chasing the hare
- Optimal play: do exactly what the other player(s) do

Prisoner’s dilemma

- DC > CC > DD > CD
- Optimal play: always defect
- Two rational players will always defect.
- Thus, (naïve) individual rationality subverts their common good

More examples of the PD in real life

- **Communal coffeepot**
  - – Cooperate by making a new pot of coffee if you take the last cup.
  - – Defect by taking the last cup and not making a new pot, depending on the next coffee seeker to do it.
  - – DC > CC > DD > CD
- **Class team project**
  - – Cooperate by doing your part well and on time.
  - – Defect by slacking, hoping the other team members will come through and sharing the benefit of a good grade.
  - – (Arguable) DC > CC > DD > CD

Iterated Prisoner’s Dilemma

- Game theory shows that a rational player should always defect when engaged in a prisoner’s dilemma situation
- We know that in real situations, people don’t always do this
- Why not? Possible explanations:
  - – People aren’t rational
  - – Morality
  - – Social pressure
  - – Fear of consequences
  - – Evolution of species-favoring genes
- Which of these make sense? How can we formalize these?

Iterated Prisoner’s Dilemma

- **Key idea:** In many situations, we play more than one “game” with a given player.
- Players have complete knowledge of the past games, including their choices and the other player’s choices.
- Your choice in future games when playing against a given player can be partially based on whether he has been cooperative in the past.
- A simulation was first done by Robert Axelrod (Michigan) in which computer programs played in a round-robin tournament (DC=5,CC=3,DD=1,CD=0)
- The simplest program won!

Some possible strategies

- Always defect
- Always cooperate
- Randomly choose
- Pavlovian

  Start by always cooperating, switch to always defecting when “punished” by the other’s defection, switch back and forth at every such punishment.
- Tit-for-tat

  “Be nice, but punish any defections”. Starts by cooperating and, after that always does what the other player did on the previous round
- Joss

  A sneaky TFT that defects 10% of the time
- In an idealized (noise free) environment, TFT is both a very simple and a very good strategy

Characteristics of Robust Strategies

Axelrod analyzed the various entries and identified these characteristics

- **nice** - never defects first
- **provocable** - responds to defection by promptly defecting. Promptly responding defections is important. "being slow to anger" isn't a good strategy; some programs tried even harder to take advantage.
- **forgiving** – programs responding to single defections by defecting forever thereafter weren't very successful. It's better to respond to a TIT with 0.9 TAT; might dampen some echoes and prevent feuds.
- **clear** - Clarity seemed to be an important feature. With TFT you know exactly what to expect and what would/wouldn't work. Too many random number generators or bizarre strategies in a program, and the competing programs just sort of said the hell with it and began to all Defect.

Implications of Robust Strategies

- You do well, not by "beating" others, but by allowing both of you to do well. TFT never "wins" a single encounter! It can't. It can never do better than tie (all C).
- You do well by motivating cooperative behavior from others - the provocability part.
- Envy is counterproductive. It does not pay to get upset if someone does a few points better than you do in any single encounter. Moreover, for you to do well, then the other person must do well. Example of business and its suppliers.
- You don't have to be very smart to do well. You don't even have to be conscious! TFT models cooperative relations with bacteria and hosts.
- Cosmic threats and promises aren't necessary, although they may be helpful.
- Central authority is not necessary, although it may be helpful.
- The optimum strategy depends on environment. TFT is not necessarily the best program in all cases. It may be too unforgiving of JOSS and too lenient with RANDOM.

Required for emergent cooperation

- A non-zero sum situation.
- Players with equal power and no discrimination or status differences.
- Repeated encounters with another player you can recognize. Car garages that depend on repeat business versus those on busy highways. Gypsies. If you're unlikely to ever see someone again, you're back into a non-iterated dilemma.
- A temptation payoff that isn't too great. If, by defecting, you can really make out like a bandit, then you're likely to do it. "Every man has his price."

Ecological model

- Assume an ecological system that can support N players
- On each round, players accumulate or loose points
- After each round, the poorest players die and the richest multiply.
- *Noise* in the environment can model the likelihood that an agent makes errors in following a strategy or that an agent might misinterpret another's choice.
- There are simple mathematical ways of modeling this, as described in Flake's book.

Evolutionary stable strategies

- Strategies do better or worse against other strategies
- Successful strategies should be able to work well in a variety of environments
  - E.g., ALL-C works well in an mono-culture of ALL-C's but not in a mixed environment
- Successful strategies should be able to "fight off mutations"
  - E.g., an ALL-D mono-culture is very resistant to invasions by any cooperating strategies
  - E.g., TFT can be "invaded" by ALL-C

Population simulation

(a) TFT wins

(b) A noise free version with TFT winning

(c) 0.5% noise lets Pavlov win

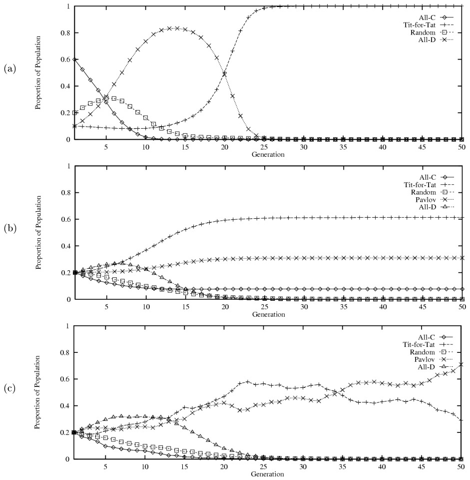


Figure 17.3 Population simulations of the ecological version of the iterated Prisoner's Dilemma: (a) an idealized version that illustrates the rise of **TFT**; (b) a noise-free simulation with **TFT** winning; (c) with 0.5 percent noise **PAV** wins

For more information

- Prisoner's Dilemma: John von Neumann, Game Theory, and the Puzzle of the Bomb, William Poundstone, Anchor Books, Doubleday, 1993.
- The Origins of Virtue: Human Instincts and the Evolution of Cooperation, Matt Ridley, Penguin, 1998.
- Games of Life : Explorations in Ecology, Evolution and Behaviour, Karl Sigmund, 1995.
- Nowak, M.A., R.M. May and K. Sigmund (1995). The Arithmetic of Mutual Help. Scientific American, 272(6).
- Robert Axelrod, The Evolution of Cooperation, Basic Books, 1984.
- The Computational Beauty of Nature: Computer Explorations of Fractals, Chaos, Complex Systems, and Adaptation, Gary William Flake, MIT Press, 2000.
- New Tack Wins Prisoner's Dilemma, By Wendy M. Grossman, Wired News, October 2004.

Hawk and Dove