Lecture 26

11/26/07 1

Matching Pennies

- Al and Barb each independently picks either heads or tails
- If they are both heads or both tails, Al wins
- If they are different, Barb wins

11/26/07 2

Payoff Matrix

Minimum of each pure strategy is the same		Barb	
		head	tail
Al	head	+1, −1	−1, +1
	tail	−1, +1	+1, −1

11/26/07 3

Mixed Strategy

- Although we cannot use maximin to select a pure strategy, we can use it to select a mixed strategy
- Take the maximum of the minimum payoffs over all assignments of probabilities
- von Neumann proved you can always find an equilibrium if mixed strategies are permitted

11/26/07 4

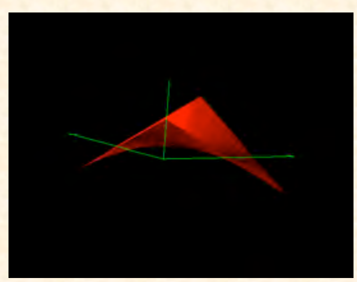
Analysis

- Let P_A = probability Al picks head
- and P_B = probability Barb picks head
- Al's expected payoff:

$$E\{A\} = P_A P_B - P_A (1 - P_B) - (1 - P_A) P_B + (1 - P_A)(1 - P_B)$$
$$= (2 P_A - 1)(2 P_B - 1)$$

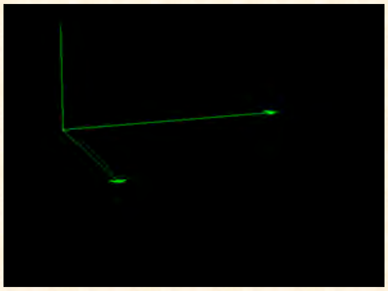
11/26/07 5

Al's Expected Payoff from Penny Game

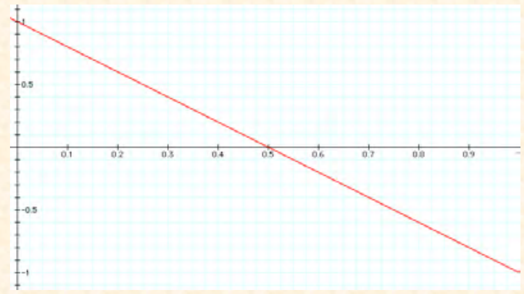


11/26/07 6

How Barb's Behavior Affects Al's Expected Payoff



How Barb's Behavior Affects Al's Expected Payoff



More General Analysis (Differing Payoffs)

- Let A's payoffs be:
  H = HH, h = HT, t = TH, T = TT
- E{A} = $P_A P_B H + P_A(1 - P_B)h + (1 - P_A)P_B t + (1 - P_A)(1 - P_B)T$
  $= (H + T - h - t)P_A P_B + (h - T)P_A + (t - T)P_B + T$
- To find saddle point set ∂E{A}/$\partial P_A = 0$ and ∂E{A}/$\partial P_B = 0$ to get:

$$P_A = \frac{T - t}{H + T - h - t}, \quad P_B = \frac{T - h}{H + T - h - t}$$

Random Rationality

"It seems difficult, at first, to accept the idea that 'rationality' — which appears to demand a clear, definite plan, a deterministic resolution — should be achieved by the use of probabilistic devices. Yet precisely such is the case."

—Morgenstern

Probability in Games of Chance and Strategy

- "In games of chance the task is to determine and then to evaluate probabilities inherent in the game;
- in games of strategy we *introduce* probability in order to obtain the optimal choice of strategy."

— Morgenstern

Review of von Neumann's Solution

- Every two-person zero-sum game has a maximin solution, provided we allow mixed strategies
- But— it applies only to two-person zero-sum games
- Arguably, few "games" in real life are zero-sum, except literal games (i.e., invented games for amusement)

Nonconstant Sum Games

- There is no agreed upon definition of rationality for nonconstant sum games
- Two common criteria:
  - dominant strategy equilibrium
  - Nash equilibrium

Dominant Strategy Equilibrium

- Dominant strategy:
  - consider each of opponents' strategies, and what your best strategy is in each situation
  - if the same strategy is best in all situations, it is the dominant strategy
- Dominant strategy equilibrium: occurs if each player has a dominant strategy and plays it

Another Example

Price Competition		Beta		
		$p = 1$	$p = 2$	$p = 3$
Alpha	$p = 1$	0, 0	50, –10	40, –20
	$p = 2$	–10, 50	20, 20	90, 10
	$p = 3$	–20, 40	10, 90	50, 50

There is no dominant strategy

Nash Equilibrium

- Developed by John Nash in 1950
- His 27-page PhD dissertation: *Non-Cooperative Games*
- Received Nobel Prize in Economics for it in 1994
- Subject of *A Beautiful Mind*

Definition of Nash Equilibrium

- A set of strategies with the property: No player can benefit by changing actions while others keep strategies unchanged
- Players are in equilibrium if any change of strategy would lead to lower reward for that player
- For mixed strategies, we consider expected reward

Another Example (Reconsidered)

Price Competition		Beta		
		$p = 1$	$p = 2$	$p = 3$
Alpha	$p = 1$	0, 0	50, –10	40, –20
	$p = 2$	–10, 50	20, 20	90, 10
	$p = 3$	–20, 40	10, 90	50, 50

better for Beta better for Alpha

Not a Nash equilibrium

The Nash Equilibrium

Price Competition		Beta		
		p = 1	*p* = 2	*p* = 3
Alpha	*p* = 1	0, 0	50, –10	40, –20
	p = 2	–10, 50	20, 20	90, 10
	p = 3	–20, 40	10, 90	50, 50

Nash equilibrium

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

Price Competition		Beta		
		p = 1	*p* = 2	*p* = 3
Alpha	*p* = 1	0, 0	50, –10	40, –20
	p = 2	–10, 50	20, 20	90, 10
	p = 3	–20, 40	10, 90	50, 50

Cooperation

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

		Bob	
		cooperate	defect
Ann	cooperate	–1, –1	–20, 0
	defect	0, –20	–10, –10

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

11/26/07 25

Ann's "Rational" Analysis (Dominant Strategy)

		Bob	
		cooperate	defect
Ann	cooperate	–1, –1	–20, 0
	defect	0, –20	–10, –10

- if cooperates, may get 20 years
- if defects, may get 10 years
- ∴, best to defect

11/26/07 26

Bob's "Rational" Analysis (Dominant Strategy)

		Bob	
		cooperate	defect
Ann	cooperate	–1, –1	–20, 0
	defect	0, –20	–10, –10

- if he cooperates, may get 20 years
- if he defects, may get 10 years
- ∴, best to defect

11/26/07 27

Suboptimal Result of "Rational" Analysis

		Bob	
		cooperate	defect
Ann	cooperate	–1, –1	–20, 0
	defect	0, –20	–10, –10

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

11/26/07 28

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

11/26/07 29

Classification of Dilemmas

11/26/07 30

General Payoff Matrix

		Bob	
		cooperate	defect
Ann	cooperate	CC (*R*) Reward	CD (*S*) Sucker
	defect	DC (*T*) Temptation	DD (*P*) Punishment

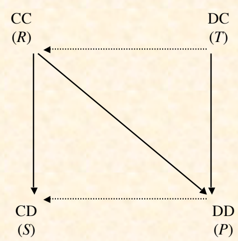
11/26/07 31

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

11/26/07 32

Three Possible Orders



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

11/26/07 33

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

11/26/07 34

The Iterated Prisoners' Dilemma

and Robert Axelrod's Experiments

11/26/07 35

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

11/26/07 36

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

		B	
		cooperate	defect
A	cooperate	3, 3	0, 5
	defect	5, 0	1, 1

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

⇓ playing ⇒	ALL-C	RAND	ALL-D	Average
ALL-C	3.0	1.5	0.0	1.5
RAND	4.0	2.25	0.5	2.25
ALL-D	5.0	3.0	1.0	3.0

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

Demonstration of Iterated Prisoners' Dilemma

Run NetLogo demonstration
PD N-Person Iterated.nlogo

11/26/07 43

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

11/26/07 44

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 **TFT**s into endless cycle of retaliation
- May be exploited by feigning accidental defection

11/26/07 45

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

11/26/07 46

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

11/26/07 47

Kant's Categorical Imperative

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

11/26/07 48

Reading

- CS 420/594: Flake, ch. 18 (Natural & Analog Computation)

Ecological & Spatial Models

Ecological Model

- What if more successful strategies spread in population at expense of less successful?
- Models success of programs as fraction of total population
- Fraction of strategy = probability random program obeys this strategy

Variables

- $P_i(t)$ = probability = proportional population of strategy i at time t
- $S_i(t)$ = score achieved by strategy i
- $R_{ij}(t)$ = relative score achieved by strategy i playing against strategy j over many rounds
  - fixed (not time-varying) for now

Computing Score of a Strategy

- Let n = number of strategies in ecosystem
- Compute score achieved by strategy i:

$$S_i(t) = \sum_{k=1}^{n} R_{ik}(t) P_k(t)$$

$$\mathbf{S}(t) = \mathbf{R}(t)\mathbf{P}(t)$$

Updating Proportional Population

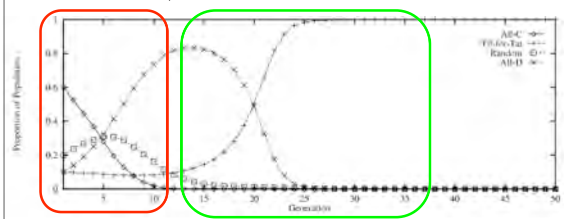
$$P_i(t+1) = \frac{P_i(t) S_i(t)}{\sum_{j=1}^{n} P_j(t) S_j(t)}$$

Some Simulations

- Usual Axelrod payoff matrix
- 200 rounds per step

Demonstration Simulation

- 60% ALL-C
- 20% RAND
- 10% ALL-D, TFT



NetLogo Demonstration of Ecological IPD

Run EIPD.nlogo

Collectively Stable Strategy

- Let *w* = probability of future interactions
- Suppose cooperation based on reciprocity has been established
- Then no one can do better than **TFT** provided:

$$w \geq \max\left(\frac{T-R}{R-S}, \frac{T-R}{T-P}\right)$$

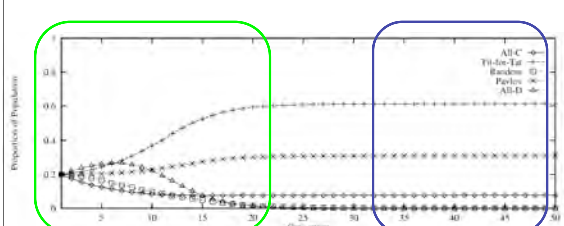
- The **TFT** users are in a Nash equilibrium

"Win-Stay, Lose-Shift" Strategy

- Win-stay, lose-shift strategy:
  - begin cooperating
  - if other cooperates, continue current behavior
  - if other defects, switch to opposite behavior
- Called **PAV** (because suggests Pavlovian learning)

Simulation without Noise

- 20% each
- no noise



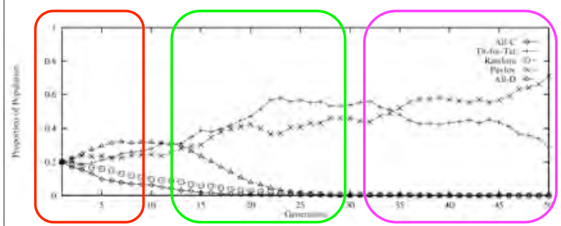
Effects of Noise

- Consider effects of noise or other sources of error in response
- **TFT**:
  - cycle of alternating defections (CD, DC)
  - broken only by another error
- **PAV**:
  - eventually self-corrects (CD, DC, DD, CC)
  - can exploit **ALL-C** in noisy environment
- Noise added into computation of $R_{ij}(t)$

11/26/07 61

Simulation with Noise

- 20% each
- 0.5% noise



11/26/07 62

Spatial Effects

- Previous simulation assumes that each agent is equally likely to interact with each other
- So strategy interactions are proportional to fractions in population
- More realistically, interactions with "neighbors" are more likely
  - "Neighbor" can be defined in many ways
- Neighbors are more likely to use the same strategy

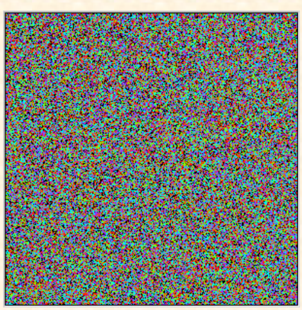
11/26/07 63

Spatial Simulation

- Toroidal grid
- Agent interacts only with eight neighbors
- Agent adopts strategy of most successful neighbor
- Ties favor current strategy

11/26/07 64

Typical Simulation ($t = 1$)

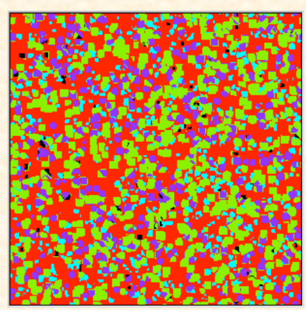


Colors:

ALL-C
TFT
RAND
PAV
ALL-D

11/26/07 65

Typical Simulation ($t = 5$)

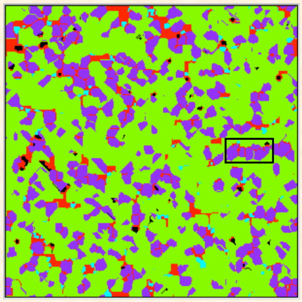


Colors:

ALL-C
TFT
RAND
PAV
ALL-D

11/26/07 66

Typical Simulation (t = 10)

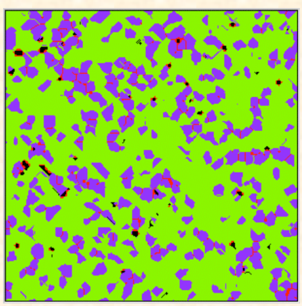


Colors:

ALL-C
TFT
RAND
PAV
ALL-D

Typical Simulation (t = 10) Zooming In

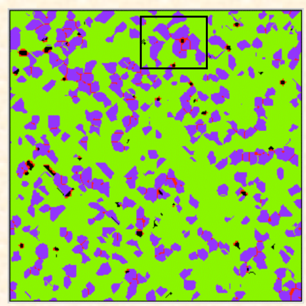
Typical Simulation (t = 20)



Colors:

ALL-C
TFT
RAND
PAV
ALL-D

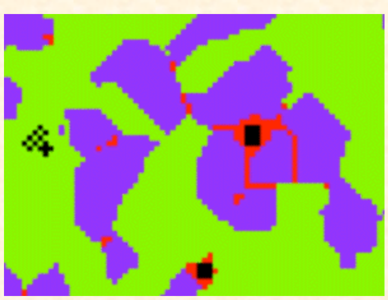
Typical Simulation (t = 50)



Colors:

ALL-C
TFT
RAND
PAV
ALL-D

Typical Simulation (t = 50) Zoom In



Simulation of Spatial Iterated Prisoners Dilemma

Run sipd.nlogo

SIPD Without Noise

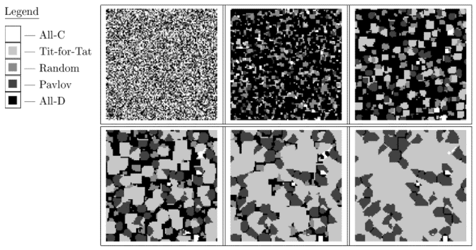


Figure 17.4 Competition in the spatial iterated Prisoner's Dilemma without noise

Conclusions: Spatial IPD

- Small clusters of cooperators can exist in hostile environment
- Parasitic agents can exist only in limited numbers
- Stability of cooperation depends on expectation of future interaction
- Adaptive cooperation/defection beats unilateral cooperation or defection

Additional Bibliography

1. von Neumann, J., & Morgenstern, O. *Theory of Games and Economic Behavior*, Princeton, 1944.
2. Morgenstern, O. “Game Theory,” in *Dictionary of the History of Ideas*, Charles Scribners, 1973, vol. 2, pp. 263-75.
3. Axelrod, R. *The Evolution of Cooperation*. Basic Books, 1984.
4. Axelrod, R., & Dion, D. “The Further Evolution of Cooperation,” *Science* **242** (1988): 1385-90.
5. Poundstone, W. *Prisoner’s Dilemma*. Doubleday, 1992.