Class Notes, Fall 2002 for: Agents, Games & Evolution OPIM 325 (Simulation)

http://opim-sun.wharton.upenn.edu/∼sok/teaching/age/f02/

Steven O. Kimbrough
University of Pennsylvania
The Wharton School
Jon M. Huntsman Hall (JMHH)
Suite 500, Room 565
3730 Walnut Street
Philadelphia, PA 19104-6340
215-898-5133
kimbrough@wharton.upenn.edu

5 December 2002. Final version for fall 2002.

$Id: age-classnotes.tex,v 1.15 2002/12/05 16:11:32 sok Exp $

Contents

Chapter 1

(R:09/05) First class. Introduction to the Course

1.1 Introduction to Agents, Games & Evolution

- Welcome. History and concept of the class.
- Emphasize: openness and flexibility; the AGEBook is a work in progress, only notes, really.
- Simulation $\rightsquigarrow$ computational modeling.
- Style: *not* foundational, didactic; instead, discursive, exploratory. Will plunge into actual experiments, discussing them and inviting your interpretations. Necessary background provided as we go.

1.2 Concept & Structure of the Class

- <u>Emphasize</u>: About new and exciting ideas, currently unfolding.
- <u>Computational modeling</u>. Procedural, constructive, representational (more later on this). Contrast with equational and axiomatic (foundational) modeling. Why do it?

- Other ideas include: emergence; natural computation and constructivism; selectionist or evolutionary or Darwinian processes and theories; complex adaptive systems; self-organizing systems; chaos; computation and complexity; evolution programs; reinforcement learning; genetic programming; zero-intelligence agents; distributed problem-solving; particle swarm optimization; learning classifier systems; and on and on.

1.3 Take Away? Who Is This For?

- For: insight into, and understanding of, strategy and strategic contexts (games).

- For: exposure to recent, surprising, and quickly-developing ideas. Conceptually, generally important; leading to applications (e.g., artificial agents doing business for us; investment strategy). Design? Management? Opportunities and threats?

- Note: computational $\not\rightarrow$ you do programming. (But you are welcome to do so.)

- Note: computational $\not\rightarrow$ formal. (Informal characterization; formal specification.)

1.4 Games or Strategic Contexts

- Two or more players; choices; rewards; interests. Contrast with: "game against Nature"

- Strong assumptions in classical game theory. Typically: common knowledge of game; rewards measured on an interval scale; cognitively unbound agents.

- Behavioral game theory. Diverges with predictions or prescriptions of classical game theory. Why? Are people stupid?

1.5 Computational Game Theory

- Artificial agents: strategy-centric vs identity-centric. (In that order.)
- "Copernican" reversal of sorts: How do raw strategies behave? What happens? A bottom up approach.
- Surprising results (which we'll discuss).

1.6 Axelrod's IPD Experiments: Canonical Prisoner's Dilemma

Strategic or normal game form:

	C	D
C	3 3	5 0
D	0 5	1 1

T = temptation = 5. R = reward = 3. P = punishment = 1. S = sucker = 0. T > R > P > S. Also R > (T + S)/2. Why?

1.7 How to Play PD?

- Defect is the dominant strategy for each player; DD is the (one and only) Nash equilibrium. But CC is a Pareto improvement (both players are better off). Hence the dilemma.
- Without communication and enforceable contracts, game theory predicts DD.
- What if the game is played repeatedly? Say 100 times. Say a random number of times?
- Axelrod's two tournaments: contestants submitted programs which played each other.

1.8 Synopsis

1. Games (or strategic contexts): more than one player, each of whom has interests. Ubiquitous.
2. Approaches to understanding: classical game theory, behavioral game theory, computational game theory. Focus here is on computational game theory.
3. Main methodology: computational modeling (to explain, predict, intervene). Computational models: procedural, constructive, representational. Surprising results.
4. This leads us to explore a collection of exciting ideas, currently very much under development.

1.9 And now. . .

- Handouts: syllabus, agebook1-3.pdf, Gigerenzer and Selten reading.
- Discuss syllabus.
- Read for next time: Axelrod, chapters 1–3
- About me. . . .
- Write a few words about you. . . . Interests, background, technical knowledge, etc.

Chapter 2

(T:09/10) Axelrod's IPD Tournaments: 1

2.1 Readings

As background for today, read/skim chapters 1 and 2 of the AGEbook (notes): "A Society of Ideas," and "Computational Explanation" which I am posting.

Do read chapters 3 and 4 of the AGEbook: "Introducing Games" and "Prisoner's Dilemma: Statics."

Also, and especially, read chapters 1–3, and appendices A and B of Axelrod's *The Evolution of Cooperation.*

For next time, for Thursday, read Axelrod, chapters 4–5. Handout: Maynard Smith, pp. 1–39. Read that, too.

I've decided to add a third lecture pertaining to the Axelrod book. (Actually, I've decided not to compress our discussion to just 2 lectures.) So next time we'll focus on dynamics and geography in IPD. Then on September 17 we'll have a discussion of what the policy recommendations are, focusing on the later chapters of Axelrod. Also, we'll use that time to cover any material not covered in the first two classes (9/10 and 9/12). I'll revise the syllabus suitably.

2.2 Today

1. Basic terminology in games. (It will be of continuing use.)

2. Introduce the problem of cooperation (of accounting for it) and set Hobbes as a foil.

3. Review Axelrod's tournament results for IPD (statics, next time is dynamics)

2.3 Introducing Games

- Read the chapter in the AGEbook.
- Strategic situations versus games
- The 2 × 2 game in *strategic form*

	C_1	C_2
R_1	r_1, c_1	r_2, c_2
R_2	r_3, c_3	r_4, c_4

2.4 Concepts

- player
- play of a game
- game in strategic form
- iterated or repeated game
- reward
- outcome
- equilibrium
- Pareto efficiency, deficiency

- strategy
- dominant strategy, super-dominant strategy
- constant-sum game, mixed-motive game, no-conflict game
- ordinal, interval, and ratio scales

2.5 Optimal Strategies

Proposition 1 (Absence of Absolutely Optimal Strategies) *For other than trivial games, played only once, there is no optimal strategy for a player, ignoring the strategies actually employed by the counterplayers.*

Here's the argument. In all nontrivial games player i can be prevented from achieving its highest reward, call it r_i^*, by the choice made by at least one other player. Without loss of generality, let $S_{i,1}S_{j,1}$ yield r_i^* for player i, all other outcomes ($S_{i,2}S_{j,2}$ etc.) yield a strictly lower reward for i.

	C_1	C_2
R_1	1 r_1^*	2 $r_{1,2}$
R_2	3 $r_{1,3}$	4 $r_{1,4}$

Assume merely that $r_1^* > r_{1,n} \quad n = 2, 3, 4$.
(See AGEbook for more info.)

2.6 Hobbes and the Problem of Coöperation

NB: Why be moral? See the Gyges ring story and Plato's *Republic.*

But the issue raised most forcefully in "modern" times (and with us today) by Thomas Hobbes (1588–1679). Here is the famous passage from his *Leviation* [1651].

> Whatsoever there is consequent to a time of war, where every man is enemy to every man; the same is consequent to the time, wherein men live without other security, than what their own strength, and their own invention shall furnish them withal. In such condition, there is no place for industry; because the fruit thereof is uncertain: and consequently no culture of the earth; no navigation, or use of the commodities that may be imported by sea; no commodious building; no instruments of moving, and removing such things as require much force; no knowledge of the face of the earth; no account of time; no arts; no letters; no society; and which is worst of all, continual fear, and danger of violent death; and the life of man, solitary, poor, nasty, brutish, and short.

2.7 Prisoner's Dilemma

The story/setup…(no communication, why?) Seems to capture essential aspects of many real-world strategic contexts. Examples?

- Arms races
- Firms' entries into markets
- Price-setting by firms
- Tipping—coöperative behavior generally

2.8 Canonical Prisoner's Dilemma in Strategic Form

	C	D
C	3 3	5 0
D	0 5	1 1

T = temptation = 5. R = reward = 3. P = punishment = 1. S = sucker = 0.

T > R > P > S. Also R > (T + S)/2. Why?

2.9 Note on PD

1. DD is the single Nash equilibrium outcome.
2. The reasoning for DD is...
3. DD is the single Pareto-deficient outcome
4. All other outcomes are Pareto-efficient
5. CC is intuitively the "social optimum"

   But remember: sometimes we want to prevent coöperation.
6. One-shot case vs definitely-repeated (a known number of times) case vs indefinitely-repeated (an unknown, e.g. random, number of times) case.

   DD is the single Nash equilibrium for the one-shot case and for the definitely-repeated case. There are many, many Nash equilibria for the indefinitely-repeated case.

2.10 Further Points

- Invented about 1950 at the RAND Corporation.
- Since then surely the most discussed and studied game.
- Human experiments, beginning at RAND in 1950, have consistently produced results other than all mutual defection.
- Poundstone's *Prisoner's Dilemma* is a nice, popular-science treatment
- Axelrod's tournaments (early 1980s) demonstrated how much remained to be learned about the game, and games of mixed-motive in general.

$Id: prisoners-dilemma-statics-foils.tex,v 1.2 2002/09/11 16:26:39 sok Exp $

2.11

Axelrod 1:[1]

The Evolution of Cooperation

Chapters 1–3, Appendices A & B

Steven O. Kimbrough

2.12 Chapter 1

- "Under what conditions will cooperation emerge in a world of egoists without central authority?"
- Why be moral? Altruistic?
- Higher authority? How else could this arise, naturally (constructively)?
- Hobbes: you need a dictator; else war of all against all
- Plato[?]: self-realization; Social contract

2.13

Chapter 1: Examples

- Tipping
- International affairs. Nuclear/biological/etc. war. MAD.

  Note: Soviet Invasion of Afghanistan, 1979.
- In legislatures; U.S. Sentate. Civil discourse today? Presidential nominations? "Outing" opponents?

Cooperation: "the cement of society"

[1] File: axelrod-1-slides.tex.

2.14

Chapter 1

- "The approach of this book is to investigate how individuals pursuing their own interests will act, followed by an analysis of what effects this will have for the system as a whole." [page 6]
- "The object of this enterprise is to develop a theory of cooperation that can be used to discover what is necessary for cooperation to emerge." [page 6]

2.15

Comment: Copernican revolution (of sorts)

- Not: How can we be rational?

  Rationality ≈ efficient means to a given end
- But: How do strategies fare in games?

  Angels and insects. Even bacteria.
- Why the change in view? What's wrong with rationality? Isn't it better to be rational?

2.16

Problems with classical game theory

1. Non-cooperative game theoretic techniques require completely rigorous specification of the 'rules of the game,' thereby limiting the scope of applicability of the techniques.

2. The focus on equilibria has amounted to a focus on the statics, at the expense of the dynamics, of games. While classical game theory may predict and explain outcomes (static equilibria), it has little to say about the constructive, or algorithmic, procedures that yield the favored equilibria. Consequently, the theory has little to say about the design of intelligent agents in strategic situations. Moreover, equilibria can easily exist but be effectively unreachable in finite time.

2.17

Problems with classical game theory

3. Many games have multiple equilibria. While theoretical progress has been made in choosing among multiple equilibria, the progress has been quite limited. Moreover, the 'folk theorem' for repeated games lurks: very roughly, in repeated games just about everything can be an equilibrium.

2.18

Problems with classical game theory

4. Competing solution concepts impose a further layer of indeterminacy on games, as nicely summarized by Shubik:

> Except for the two-person constant-sum game, it has long been recognized that there is no unique solution concept that can be regarded as *the* natural extension of individual rational behavior to multiperson decision making. This is one of the central problems of game theory.

2.19

Problems with classical game theory

5. Poor success at prediction.

> Ironically, [classical] game theory is often hoisted on its own pétard: many of its most fundamental predictions—predictions that would have been too vague to test with any confidence in the pre-game-theoretic era—are *decisively and repeatedly disconfirmed,* in laboratory settings, with substantial agreement among experimenters, regardless of their theoretical priors.

2.20

Evolutionary...

> ... *game theory is about the emergence, transformation, diffusion, and stabilization of forms of behavior.* Traditionally, game theory has been seen as a theory of how "rational agents" *do* behave, and/or how the rest of us *should* behave. Ironically, game theory which for so long was predicated upon agent rationality, has shown us, by example, the shakiness of the concept. For one thing, the centipede game and others like it show that there is nothing substantively "rational" about even so simple a thing as eliminating dominated strategies .... Moreover, the solution to some games (even when unique) is often so sophisticated that it is implausible that ordinary people would be willing to spend the resources to discover it. This supports the evolutionary notion that good strategies diffuse across populations of players rather than being learned by "rational optimizers." Finally, experimental studies of dictator, ultimatum, and public goods games indicate that if people are "rational," it must be in a sense far more sophisticated than the simple, self-interested, maximization of expected utility.
>
> It is better to drop the term "rational" altogether, which is what we do in this book ....

2.21

Continuing:

> In the same vein, we do not follow classical game theory in asking how agents "learn" to play optimal strategies, because the cognitive processes involved in "learning" are probably, under most conditions, much less important than the forms of imitation underlying the replicator dynamic. . . and cultural transmission. . . . In short, evolutionary game theory replaces the idea that games have "solutions" that agents "learn," with the idea that games are embedded in antural and social processes that product agents who play effectively.
>
> Dispensing with the rationality postulate does not imply that people are *irrational* (whatever that means). the point is that the concept of "rationality" does not help us understand the world.

Comment: Going too far. But first things first.

2.22

Canonical Prisoner's Dilemma

Strategic or normal game form:

	C	D
C	3 3	5 0
D	0 5	1 1

T = temptation = 5. R = reward = 3. P = punishment = 1. S = sucker = 0.

T > R > P > S. Also R > (T + S)/2. Why?

2.23

Chapter 1

- Conditions of play: no way to make meaningful threats or promises.

  Think: space alien. (Or should you?)

- "What makes it possible for cooperation to emerge is the fact that the players might meet again."

- Finite, known number of meetings: still a D-D situation by lights of classical game theory. Backwards induction argument.

  Note: "surprise exam paradox."

2.24

Useful fact: $\sum_{i=0}^{\infty} w^i = (1-w)^{-1}$ for $0 < w < 1$

- But with (indefinite) repetition a form of meaningful communication (not cheap talk) can arise. Actions speak for themselves.

- But: future gains are worth less than gains now. What's the discount rate? What's the interest rate?

PV factor =

$$\sum_{i=0}^{\infty} \frac{1}{(1+r)^i}$$

r = interest rate. Axelrod's discount rate, $w = (1+r)^{-1}$

2.25

Comment

- r and w are inversely related
- When the interest (hurdle) rate is small, the discount rate is large.
- When the interest rate is small, we are more interested in looking to the future.

2.26

Strategy (for a game): complete directions on what to do

- In these tournaments, how A does relative to B depends on what strategies A and B have, but also on what strategies the Cs (the other players) have!
- Proposition 1 [page 15]: "If the discount parameter, w, is sufficiently high, there is no best strategy independent of the strategy used by the other player."

Proof?

2.27

Proof of Proposition 1

Mutual AllD yields $\sum P \cdot w^i = P \cdot (1-w)^{-1}$.
Mutual AllC yields $\sum R \cdot w^i = R \cdot (1-w)^{-1}$.
AllD in the face of SuperRetaliate yields $T + (\sum P \cdot w^i - P)$.
So if $R \cdot (1-w)^{-1} > T - P + P \sum w^i$ or
$(R-P) \cdot (1-w)^{-1} > T - P$ or $w > \frac{(T-R)}{(T-P)}$
Note: In the canonical case: $\frac{(5-3)}{(5-1)} = \frac{1}{2}$. $w = (1+r)^{-1} \Rightarrow r < 50\%$.

2.28

Comments [pages 17–19]

- Payoffs need not be comparable.
- Payoffs need not be symmatric.
- Payoffs to a player need only be measured on an interval scale. (Utility theory)
- Cooperation may or may not be desirable.
- Players need not be rational in any sense. Strategies may simply describe what they do and the game how they get rewarded. Bacteria.

2.29

Chapter 2: Computer Tournaments. Axelrod:

p. 29 : Existing literature doesn't tell you how to play IPD (iterated prisoner's dilemma).

NB: comments above on classical game theory.

p. 30 : "the proposition of the previous chapter demonstrates that what is effective depends not only upon the characteristics of a parfticular strategy, but also upon the nature of the other strategies with which it must interact."

p. 30 : "An effective strategy must be able at any point to take into account the history of the interaction as it has developed so far." Why?

2.30

Computer Tournament #1

- 14 entries (computer programs) + RANDOM. All 15 played each other and themselves. 200 rounds/moves in each game.
- "In a preliminary tournament, TIT FOR TAT scored second place"
- "The striking fact is that *none* of the more cmplex programs submitted was able to perform as well as the original, simple TIT FOR TAT."

  Contrast with computer chess tournaments.

2.31

Results and Comments

- Average scores: page 193.
- 200·3 (=R) = 600. 200·1 (=P) = 200. T4T got 504.
- "Surprisingly, there is a single property which distinguishes the relatively high-scoring entries from the relatively low-scoring entries. This is the property of being *nice,* which is to say never being the first to defect." [p. 33]
- Nice guys do well when they play each other (600).

2.32

Results & Comments (con't.)

- "What is not obvious is that the relative ranking of the eight top rules was largely determined by just two of the other seven rules. These two rules are *kingmakers* because they do not do very well for themselves, but they largely determine the rankings among the top contenders." [p. 34] Why?
- "*Forgiveness* of a rule can be informally described as its propensity to cooperate in the moves after the other player has defected. Of all the nice rules, the one that scored lowest was also the one that was least forgiving." [p. 36]
- JOSS, sneaky TIT FOR TAT. [p. 36] Probabilistic defecting exploitation. #12, 304 points on average.

2.33

"A major lesson of this tournament is…

…the importance of minimizing echo effects in an environment of mutual power. …a single defection may be successful when analyzed for its direct effects, and perhaps even when its secondary effects are taken into account. but the real costs may be in the teriary effects when one's own isolated defections turn into unending mutual recriminations." [p. 38]

We all know people like that.

2.34

More lessons

- "These results from supplementary rules [that would have won the tournament] reinforce a theme from the analysis of the tournament entries themsleves: the entries were too competitive for their own good." [p. 40]
- In the second tournament, the entrants "were aware not only of the outcome of the first round, but also of the concepts used to analyze success, and the strategic pitfalls that were discovered." [p. 41]

2.35

The second tournament

62 entries. "As announced in the rules, the length of the games was determined probabilistically with a 0.00346 chance of ending with each given move. this is equivalent to setting w=0.99654. Since no one knew exactly when the last move would come, end-game effects were successfully avoided in the second round." [pp. 42-3]

Submitted by, and only by, Anatol Rapoport, "TIT FOR TAT was the simplest program submitted in the first round, and it won the first round. It was the simplest submission in the second round, and it wn the second round. Even though all the entrants to the seocnd round knew that TIT FOR TAT had won the first round, no one was able to design an entry that did any better." [p. 42]

2.36

The second tournament

"There were a number of rules in the second round of the tournament that deliberately used controlled numbers of defections to see what the could ge away with. To a large extent, what determined the actual ranking of the nice rules was how well they were able to cope with these challengers." [p. 44]

"So while it pays to be nice, it also pays to be retaliatory. TIT FOR TAT combines these desirable properties. It is nice, forgiving, and retaliatory." [p. 46]

2.37

Is TIT FOR TAT robust?

- See Appendix A, especially beginning page 197, and 203. 5 rules, regression equation. Best in 5 of 6; second in 1 of 6 hypothetical tournaments.
- Replicator dynamics test; cf., page 51. "Ecological perspective." Are the sucessful rules successful because they beat up on the losers?
- "On the other hand, TIT FOR TAT foregoes the possibility of exploiting other rules."

Chapter 3

(R:09/12) Prisoner's Dilemma: Dynamics

3.1 Announcements

Handout: chapter 2 of Maynard Smith's *Evolution and the Theory of Games.* Read it for next time. Also, read Axelrod, chapters 6, 7, 8, and 9.

Today, focus on dynamics of IPD. An introduction to dynamics for games, especially from our computational (procedural, constructive, representational) perspective. But this is just the beginning; much more later as we progress. Handouts: classnotes for today; notes for AGEbook chapter 5, "PD: Dynamics." Both of these are now on the Web site, too:

http://opim-sun.wharton.upenn.edu/∼sok/teaching/age/f02/.

Will provide revised syllabus next time, too. Discussion then, too, of applications. Then on to Skyrms.

3.2 Why the Interest in Dynamics?

- Axelrod's tournaments yielded surprising and important results, yet they were merely particular tournaments. Is there something to be learned more generally about the behavior of strategies? Here: IPD. Later, but very soon: other games.

- Statics question: Given a fixed pool of players and a regime of play, etc., who wins?

Interesting, but recall proposition 1.

- Dynamics question types: (a) Population trajectories (*population questions*), (b) Properties of individual strategies (or types of strategies) (*individual questions*)

  Related questions. Note: later replace individual strategies with (identify-centric) individuals.

- Why are we interested?

  Explain, predict, intervene?

3.3 Dynamics Questions

- Dynamics questions: Given a pool of players and a regime of play, etc., how does the pool change over time? Does it converge to a particular constitution? If so, what is it? Short of predicting the paths of change for all possible strategies, what can we say about what happens when particular (types of) strategies win (or lose) out? &c.

  Note: The very large number of possible strategies (proposition 2).

- Keep in mind: modeling assumptions made in obtaining results.

3.4 Axelrod Chapter 3: The Chronology of Cooperation

- "Suppose that everyone came to be using the same strategy. Would there be any reason for someone to use a different strategy, or would the popular strategy remain the choice of all?" [p. 55]

- See Maynard Smith and ESS (next time).

  Biological underpinnings and interests. (Can this stuff evolve and stay evolved?)

  Note: Could provide prior information about new situations. Compare with "The doctors know how to cure cancer, but they are keeping it secrete in order to make more money."

3.5 "Ecological" Simulation

- Axelrod, p. 51. Aka: replicator dynamics. But terminology is not standard. Population questions.

- The setup: Pool of n strategies, $s_1, s_2, \ldots, s_n$ each with its own frequency, $w_1, w_2, \ldots, w_n$, $w_i \geq 0$, $\sum w_i = 1$. Draw pairs of strategies randomly in proportion to their weights (or fitnesses), play the game (with iterations) and accumulate points obtained in this generation. Repeat some large number of times. Refix the w_is based on points accumulated and reset the accumulators. Repeat numGenerations times.

  Note: This is a discrete-generation model. Can also do a continuous-generation model.

- More successful strategies will become more prominent in the pool/population, and vice versa.

3.6 Results of Ecological/Replicator Dynamics?

- Axelrod: Did this for for the tournament strategies and Tit for Tat won out (p. 51)

- (sok): Compare AllD with AllC, starting out equal in numbers.

  For $R = 3, T = 4$:

  999,0.9573873123549033,0.042612687645096785

  1000,0.9574091789730933,0.04259082102690666

  For $R = 100, T = 101$:

  999,0.542669398474112,0.457330601525888

  1000,0.542636303984413,0.45736369601558696

3.7 Other Sorts of Experiments?

- Forthcoming, as we proceed.

- 8 "one deep" or M-1 IPD strategies, e.g., 1=C, 0=D

Initial play	Response to 1	Response to 0
1	1	0

- 64 M-2 IPD strategies, e.g.,

Initial play	Response to 11	Response to 10	to 01	to 00
11	1	0	1	0

3.8 Nondeterministic Strategies (or effects)

- Suppose noise or miscommunication, etc., or even deliberate policy, e.g., Generous Tit for Tat:

Initial play	Response to 1	Response to 0
1	1	0.333

- See "Chaos and the Evolution of Cooperation," by Martin Nowak and Karl Sigmund, *Proc. Nat. Acad. Sci.* vol. 90, pp. 5091–5094, 1993. Neglecting first move, strategies as 4-tuples:

Receive:	R	S	T	P
Play:	$1-\epsilon$	ϵ	$1-\epsilon$	ϵ

3.9 Discount rate and all that

See §§5.2–3 of the AGEbook.

3.10 Collective stability. (Similar to ESS.)

- Look for stability in the population by asking whether a uniform population can be "invaded" by a single "mutant" strategy.
- "Therefore a new strategy is said to *invade* a native strategy if the newcomer ges a higher score with a native than a native gets with another native."

- "A strategy is *collectively stable* if no strategy can invade it."

- "...collectively stable strategies are important because they are the only ones that an entire population can maintain in the long run in the face of any possible mutant." NB: Mutations change strategies, not outcomes.

3.11 The stability of TIT FOR TAT

- "In fact, no rule can invade TIT FOR TAT if the discount parameter, w, is sufficiently large." [p. 58] (Proposition 2, page 59.)

- See §5.4 of the AGEbook.

- Comment: careful. What does the proof assume?

- Proposition 3, page 61: "Any strategy which may be the first to cooperate can be collectively stable only when w is sufficiently large."

- Intuition?

3.12 Nonrandom enounters; clustering

- Proposition 5, page 63: AllD is always collectively stable

- Why?

- But, clusters of new arrivals could defeat AllD. "thus, even a small cluster of TIT FOR TAT layers can get a higher average score than the large population of meanies they enter." [p. 64] Why?

- Is clustering, of any sort, ever realistic? NB: How small the clustering numbers are. Also: looking ahead to cellular automata and spatialization of games.

- See §5.5 of the AGEbook.

3.13

Axelrod 2:[1]

The Evolution of Cooperation

Chapters 4–5

Steven O. Kimbrough

3.14

Chapter 4: Live and Let Live

> "The live-and-let-live system was endemic in trench warfare. It flourished despite the best efforts of senior officers to stop it, despite the passions aroused by combat, despite the military logic of kill or be killed, and despite the ease with which the high command was able to repress any local efforts to arrange a direct truce.
>
> "The is a cost of cooperation emerging despite great antagonism between the players." [page 74]

4. Why was it characteristic of trench warfare in World War I, but of few other wars?

[1]File: axelrod-2-slides.tex.

3.15

Chapter 4: Live and Let Live

- "Thus the situation meets the conditions for a Prisoner's Dilemma between small units facing each other in a given immobile sector." [page 75]
- "For our purposes, two key factors make the battalion the most typical player. On the one hand, it was large enough to occupy a sufficient sector of the front to be 'held responsible' for aggressive actions which came from its territory. On the other hand, it was small enough to be able to control the individual behavior of its men, through a variety of means, both formal and informal." [page 76]

3.16

Chapter 4: Live and Let Live

- "Locally, the dilemma persisted: at any given moment it was prudent to shoot to kill, whether the other side did so or not. What made trench warfare so different from most othe combat was that the same small units faced each other in immobile sectors for extended periods of time." [page 77]
- "But direct truces were easily suppressed." [page 78] Why?
- "Another way in which mutual restraint got started was during a spell of miserable weather."

3.17

Chapter 4: Live and Let Live

- "More effective in the long run were various methods which allowed the two sides to coordinate their actions without having to resort to words." [page 78]
- "After all, if you prevent your enemy from drawinghis rations, his remedy is simple: he will prevent you frm drawing yours." [page 79]
- "The strategies that could sustain mutual cooperation were the ones which were provocable." [page 79]
- "What finally destroyed the live-and-let-live system was . . . "

3.18

Chapter 4: Live and Let Live

- "Thus, the evolution of strategies was based n deliberate rather than blind adaptation." [page 84]
- Two additions to the theory [page 84]: ethics and ritual
- "Under suitable circumstances, cooperaton can develop even between antagonists." [page 87]

`/* $Header: /home/sok/cvsrepos/agebook/axelrod-2-slides.tex,v 1.2 2002/09/17 1`

Chapter 4

(T:09/17) ESS and Policy

4.1 Announcements

Today: (a) Collective stability, ESS and related matters; (b) Discussion of policy implications of Axelrod's (and related) findings

For Thursday, 9/19/02, read chapters 1–2 of Skyrms's *Evolution of the Social Contract.* (The promised revised syllabus is on the way.) Remember our site is:

http://opim-sun.wharton.upenn.edu/~sok/teaching/age/f02/

Today will use notes in AGEbook a lot in lecture. For (b), policy matters, see notes in AGEbook, chapter 6.

4.2 Why dynamics? redux

- Because off-equilibrium states matter to us; because we simply want to understand (explain, predict, intervene) the evolution of strategic interaction; because we want to explain certain phenomena (altruism, coöperation, reciprocity, symbiosis, all non-Hobbesian conditions)
- Because we want to eliminate equilibria.

  Explain: too many equilibria in repeated (finite or indefinite or infinite) games; folk theorem; which equilibria do plausible dynamics lead to? Recall: hidden cure for cancer; argument from the market
- Now, first, recall Axelrod's notion of collective stability: AGEbook, §5.4.

4.3 Axelrod's chapter 5 (with Hamilton): 1

- "...foresight is not necessary for the evolution of cooperation." [page 88]
- The locus of selection: group or individual? "The original individualistic emphasis of Darwin's theory is more valid." [page 89]
- For "popular" writings, see Dawkins.
- But then a puzzle: How does cooperation happen? Even symbiosis.

4.4 Axelrod's chapter 5 (with Hamilton): 2

- "...broadly, genetical kinship theory and reciprocity theory." [page 89]
- Emphasis here on individuals re-encountering each other
- "Many of the benefits sought by living things are disproportionally available to cooperating groups." [page 92] Symbiotic relationships included. Total lack of mentality included.
- "...if the payoffs are in terms of fitness, and the interactions between pairs of individuals are random *and not repeated,* then any population with a mixture of heritable strategies evolves to a state where all individuals are defectors." [page 92]

4.5 Axelrod's chapter 5 (with Hamilton): 3

- "For any value of w, the strategy of unconditional defection (ALL D) is always stable; if everyone is using this strategy, no mutant strategy can successfully invade the population." [page 93]
- "Stated formally, a strategy is evolutionarily stable if a population of individuals using that strategy cannot be invacted by a rare mutant adoption a different strategy." [page 93]

  Note: Compare with collective stability, Axelrod.

4.6 Chapter 5: Biological Systems: 3 separate questions, page 95: 4

- *Robustness.* What type of strategy can thrive in a variegated enviornment…?
- *Stability.* Resist invasion?
- *Initial viability.* Get a foothold?

Emphasize: high enough w needed.

4.7 Axelrod's chapter 5 (with Hamilton): 5

Now, focus on foothold

- Kinship
- Reciprocity

"story that emerges" [page 99]

Key: "an individual must not be able to get away with defecting without the other individuals being able to retaliate effectively." [page 100]

4.8 Now, (Maynard Smith, John) and ESS

A collectively stable strategy (Axelrod) is a special case of an ESS, which is more general.

See §§5.7f in the AGEbook.

$Id: ess-policy-foils.tex,v 1.1 2002/09/17 15:37:46 sok Exp $

Chapter 5

(R:09/19) Skyrms, Chapter 1

See the notes in the AGEbook.

Chapter 6

(T:09/24) Skyrms, Chapters 2 and 3, maybe 4

6.1 Administration

Today: Skyrms, chapters 2–3 and maybe 4.
For next time, read the rest of Skyrms.

6.2 Skyrms, Chapter 2: Commitment

- Perhaps the most fundamental assumption, or maxim, of rational choice theory is: *Never pick a dominated option.*
- Discuss: subgame perfection/modular rationality.
- Rational choice theory and prescriptions against violations of modular rationality prevent certain kinds of credible threats.

  Examples: *Dr. Strangelove,* MAD, *Gianni Schicchi, Ulysses and the Sirens,* Schelling's smoker and hiding the cookies. Dracula?

  Other examples?

6.3 Further points

- Threats and promises: empty if they violate modular rationality?

- Other examples[?]: threaten price cuts to deter market entry. Saddam?
- Now let's look at the notes in the AGEbook. Emphasize: the main point is that the evolutionary dynamic does not respect modular rationality. [Be sure to be clear on what this means.]

6.4 Skyrms, Chapter 3: Mutual Aid

Again, see the AGEbook notes.

Chapter 7

(R:09/26) Finishing up Skyrms

7.1 Announcements

- Elster talk.

  FYI: Professor Elster's talk on 'Emotions and Rationality' will be held in Logan Hall, G17 (Ground Level) at 4:00 P.M. Wednesday, October 9, 2002.

- Kearns talk.

  "The Algorithmic Revolution in Game Theory: The Internet to Artificial Intelligence," A lecture by Michael Kearns, Professor of Computer and Information Science Co-Director, Institute for Research in Cognitive Science.

  Tuesday, October 15, 2002 4:00–5:15 p.m. Amado Recital Hall, Irvine Auditorium

  Reception immediately following. For more information or to request an e-ticket please contact:

  provlec@pobox.upenn.edu or phone 215-898-7227

- Handout: "The Stag Hunt" by Brian Skyrms. (also at: http://opim-sun.wharton.upenn.edu/~sok/papers/). Read this for next time (10/1). Also next time: other games and a presentation of experiments on the stag hunt game.

- Will hand out a revised syllabus, but we're only one day behind and I'll skip "Evolution and Equilibria" for now. So, next Thursday we'll be back on schedule with "Surprises and the Surprise Exam." See:

  `http://opim-sun.wharton.upenn.edu/~sok/-`

  `comprats/surprise-exam-slides-poland.pdf`

  which you should download and glance through. (Do I need to print it out for you?)

7.2 Chapters in Skyrms

Last time we were still discussing chapter 3. So that's where we'll start today. See AGEbook for notes on chapters 3, 4, and 5.

7.3 Skyrms: Postscript

Bargaining with more than 2 players: an open question. Basically, much territory remains open for exploration.

7.4 Taking Stock

Where are we?

- Have moved focus to repeated games (definitely, indefinitely). Actually, generalize to repeated encounters by common players; game may change. Do this for greater realism.

- Often: no optimal strategy, independent of the strategies played by counter-players.

  Even when restricted to "reasonable" strategies. ("rational" strategies may lead to optimal counter-strategies, but are typically implausible)

- Find that the replicator dynamics (ecological simulation), based on random pairing, will often produce more sensible, intuitively reasonable outcomes. E.g., emergence of cooperation in Prisoner's Dilemma.

  Concept: ESS. Refinement of Nash equilibrium.

- Finding: modular rationality (subgame perfection) is not "respected" by the replicator dynamics.

- Finding: correlation matters. With the right kind of correlation, often quite small, strictly dominated strategies can take over a population.

- Finding: coordinating on 'external' events may produce Pareto-improved results/equilibria (by affording the players new strategies or mixtures thereof).

- Finding: meaning$_{nn}$ can plausibly arise entirely through an evolutionary process.

Lots of questions remain, but these are important results and fundamental. NB The (extended) evolutionary dynamic has been quite fruitful in providing a constructivist, naturalistic mode of explaining strategic behavior in populations. What next?

7.5 Discussion of an Example: Fashion

Useful as background (only): "Design Innovation and Fashion Cycles," by Wolfgang Pesendorfer, *The American Economic Review,* vol. 85, no. 4, 1995, pp. 771–92.

Classically, fashion:

- Certain things—clothes, hairstyle, etc.—are "in" temporarily, then out.

- Change is regular

- Attributes largely non-functional

- Motivation?

> [Fashion] satisfies the need of differentiation because fashions differ for different classes—the fashions of the upper stratum of society are never identical with those of the lower; in fact, they are abandoned by the former as soon as the latter prepares to appropriate them.

We might think, then, of 'fashion' as a set of tags that agents take on in order to signal certain information. With Skyrms's essay on meaning in mind, can we sketch out an account of 'fashion'?

Note:

- Yogi Berra: "Some places are so popular that no one goes there anymore."
- Irony; Captain Renault to Richard Blain in the presence of Major Strasse (in the movie *Casablanca*): "Rick, Major Strasse is one of the reasons the German Reich enjoys the reputation it has today."

Chapter 8

(T:10/01) Beyond PD

8.1 Administration

First today, Annapurna on stag hunt. Then, discussion of Skyrms's paper "The Stag Hunt." Then, time permitting, discussion of some interesting 2×2 games.

Next: Continuing from today, and sok's "Surprising Ramifications of the Surprise Exam Paradox."

First, this, then Annapurna.

8.2 *Tour d'Horizon*

- Recall: games, strategic interactions are everywhere; key concept for understanding all social phenomena, including economics and commerce.

- Recall: classical game theory limited in its ability to explain, predict, and support intervention. Failure of prediction; often not applicable in design.

- Recall: focus on iterated or repeated games. Up to now: mainly PD; also ultimatum; split the cake; mention of stag hunt and chicken, (battle of the sexes?)

- Many interesting, promising findings (summarized last time) based on looking at iterated games and study of the dynamic involved.

- Dyanmics so far: exhaustive (Axelrod's tournaments), "ecological" or replicator dynamics (*mean-field* dynamics); some hints on other dynamics (e.g., correlation, spatialization).

- Looking forward, big themes (variables):

  games × dynamics × individual intelligence and learning

  Note: nothing yet on individual intelligence and learning

Now, Annapurna.

8.3 Stag Hunt, aka: Assurance Game

Read Skyrms, "The Stag Hunt." Also very useful: "A Dynamic Model of Social Network Formation," Brian Skyrms and Robin Pemantle, *Proceedings of the National Academy of Sciences,* vol. 97, no 16, August 1, 2000, pp. 9340–9346. (Quite technical)

	S	H
S	2,2	0,1
H	1,0	1,1

Table 8.1: Stag Hunt, version 1. H = hunt hare; S = hunt stag

	S	H
S	9,9	0,8
H	8,0	7,7

Table 8.2: Stag Hunt, version 2 (assurance game)

In version 2, if you hunt hare and the counter-player hunts stag, you do better than if both hunt hare.

Both versions have two equilibria: HH and SS. While SS is Pareto-efficient (and HH Pareto-deficient), SS is *risk-dominant.* Intuitively, playing H is safer than playing S; maxmin grounds. In a symmetric 2×2 game, a strategy is risk dominant if both players prefer it on the assumption that the other player is playing his strategies 50:50. So, we have risk dominance for HH in version 2.

8.4 Hobbes, Rousseau and Hume

Rousseau. Hume's meadow-draining problem; rowing.

p. 5: Hobbes: The Foole's mistake is to ignore the future.

8.5 Transforming IPD to Stag Hunt

Skyrms, page 6. P. 7:

> The problem of instituting, or improving, the social contract can be thought of as the problem of moving from riskless Hunt Hare equilibrium to the risky but rewarding Stag Hunt equilibrium.

Discuss: range of application? (Is it just social theory?)

8.6 Dynamics and the Stag Hunt

Is HH evolutionarily stable wrt SS? vice versa? Well S,H < H,H and H,S < S,S. Both H,H and S,S are collectively stable (Axelrod). Neither can invade the other. (Look at small numbers, too; all this with random pairing, mean-field pairing.

Recall on ESS (AGEbook):

Page 14: "Since I is stable, $W(I) > W(J)$." [Discuss; assuming I is stable, what follows?]. "Since $p \ll 1$, this requires, for all $J \neq I$,"

$$\textit{either} \quad E(I,I) > E(J,I) \tag{8.1}$$

$$\textit{or} \quad E(I,I) = E(J,I) \textit{ and } E(I,J) > E(J,J) \tag{8.2}$$

[Discuss! There's actually an error here in Maynard Smith. p is the rather high frequency of incumbent Is. So, instead, $1 - p \ll 1$] Essential here are the fitness equations at the bottom of page 13 and on 14:

$$W(I) = W_0 + pE(I,I) + (1-p)E(I,J) \tag{8.3}$$

$$W(J) = W_0 + pE(J,I) + (1-p)E(J,J) \tag{8.4}$$

and

$$p' = pW(I)/\bar{W} \tag{8.5}$$

$$\bar{W} = pW(I) + (1-p)W(J) \tag{8.6}$$

OK, suppose I is the incumbent. Can it be invaded by J having (small) frequency $1-p$? Neglect W_0. [Why? discuss] Now it's easy:

1. If $E(I,I) > E(J,I)$, then since $(1-p)$ is small we may neglect it and obviously, $W(I) > W(J)$

   Comment: size matters.

2. If $E(I,I) = E(J,I)$, then we can't neglect the $(1-p)$ and require that $E(I,J) > E(J,J)$

   Comment: again, size matters. Why?

8.7 Noise?

p.10, "we know how to analyze this system but the news is not good."

8.8 Local Interaction

P. 11, no good news if arrayed in a circle.

Good news if there's a leader (page 11). "the structure of local interaction can make a difference"

8.9 Dynamic Interaction Structures

From the PNAS article:

> Individual agents begin to interact at random. The interactions are modeled as games. The game payoffs determine which interactions are reinforced, and the social network structure emerges as a consequence of the dynamics of the agents' learning behavior.

"It can be shown, both by simulation and analytically, that stag hunters learn to interact with other stag hunters. This, perhaps, should come as no surprise. It is not quite so obvious that hare hunters will end up interacting with other hare hunters."

"Here, we finally have a model that can explain the institution of a modest social contract."

8.10 from the PNAS paper

Each morning, each agent goes out to visit some other agent. The choice of whom to visit is made by chance, with the chances being determined by the relative *weights* each agent has assigned to the others. For this purpose, agent number i has a vector of weights $\langle w_{i1}, \ldots, w_{in} \rangle$ that she assigns to other players (assume $w_{ii} = 0$). Then she visits agent j with probability

$$\text{Prob(agent } i \text{ visits } j) = \frac{w_{ij}}{\sum_k w_{ik}}\,. \tag{8.7}$$

Here we are interested in a symmetric baseline model, so we will assume that all initial weights are 1. Initially, for all agents, all possible visits are equiprobable.

The weights w_{ij} are initialized to 1 for $i \neq j$, and are then updated according to

$$w_{ij}(t+1) = w_{ij}(t) + u(i, j; t), \tag{8.8}$$

where $w_{ij}(t)$ is the weight agent i gives to agent j at time t and $u(i, j; t)$ is the payoff of the game played at time t between visitor i and host j (and zero if this visit did not occur at time t). This, together with specification of the visitation probabilities in Eq. **8.7**, defines the model. Changing the initial weights does not affect the qualitative behavior of any model, so there is no need to vary the initialization.

Chapter 9

(R:10/03) The Surprise Exam and Rationality

9.1 The Surprise Examination Paradox

- Aka: Surprise Hanging Problem
- "There will be a surprise exam given in one of the next 6 meetings of the class."
- Reasoning by backwards induction…

9.2 From Grim et al., *The Philosophical Computer*, page 163

The similarity of this reasoning to that of the argument for dominant defection throughout a series of known finite length is worth noting because of course the Surprise Examination is treated standardly in the philosophical literature as a *paradox,* thought to hide some fallacious piece of logical legerdemain. That the same form of reasoning is thought of as valid in the theoretical economics literature, though perhaps inapplicable in some practical sense, indicates that important work remains to be done in bridging the two bodies of work.

9.3 First question on the exam

1. Explain the fallacy in the reasoning that led you to believe it impossible for me to give you a surprise exam as announced.

Elementary confusion by the students: Speech alone (in this context) does not have the power to prevent an exam from being given, even a surprise exam.

The real question: Can the teacher give a surprise exam *and* speak truly in saying that there will be a surprise exam?

Will try to reconstruct and present a proper way of reasoning about this.

Begin with the one-shot problem: "There will be a surprise exam tomorrow."

9.4 Version 1 Applied to the One-Shot Surprise Exam Problem

1. $E \vee \neg E$

   There will be an exam tomorrow or not, a tautology.

2. $l = \frac{2}{3}$

   Arbitrary threshold (<u>l</u>ine or <u>l</u>evel); may be changed without loss of generality.

3. $P(E) \geq l \rightarrow \neg S$

   If the probability of an exam tomorrow is greater than or equal to the stipulated threshold, then no surprise.

4. $V(a) \rightarrow (E \wedge S)$

   There will be an exam tomorrow and it will be a surprise, asserted by the teacher. $a =$ the assertion by the teacher that there will be an exam on the next class (E) and that it will be a surprise (S). If the assertion is truthful or veridical, $V(a)$, then $(E \wedge S)$.

5. $E \rightarrow P(E) = 1$

   If there is an exam tomorrow, then the probability of that there is an exam tomorrow is 1.

6. $P(E) = 1 \rightarrow P(E) \geq l$

   A simple mathematical truth.

$\vdash$ $(\neg S \wedge \neg V(a)) \vee \neg V(a)$

Comment: Valid, but unsound. Premise (5) is the problem.

9.5 Version 2 of the Students' Reasoning Applied to the One-Shot Surprise Exam Problem

1. $E \vee \neg E$
2. $l = \frac{2}{3}$
3. $P(E) \geq l \rightarrow \neg S$
4. $V(a) \rightarrow \Box(E \wedge S)$
5. $\Box E \rightarrow P(E) = 1$
6. $P(E) = 1 \rightarrow P(E) \geq l$

$\vdash$ $\neg V(a)$

Comment: Valid, but now premise (4) is problematic.

9.6 Version 3 of the Students' Reasoning Applied to the One-Shot Surprise Exam Problem

1. $E \vee \neg E$
2. $l = \frac{2}{3}$
3. $P(E) \geq l \rightarrow \neg S$
4. $V(a) \rightarrow (P(E) = 1 \wedge S)$
5. $P(E) = 1 \rightarrow E$
6. $P(E) = 1 \rightarrow P(E) \geq l$

$\vdash \neg V(a)$

Comment: Now premise (4) is problematic. If this *is* what our teacher meant, then we'll simply get another teacher, who will mean something else. The real question is whether when we find such a teacher she can speak truly and give the surprise exam.

9.7 Version 4 of the Students' Reasoning Applied to the One-Shot Surprise Exam Problem

1. $E \vee \neg E$
2. $l = \frac{2}{3}$
3. $P(E) \geq l \rightarrow \neg S$
4. $V(a) \rightarrow (E \wedge S)$
5. $P(E) = 1$
6. $P(E) = 1 \rightarrow P(E) \geq l$

$\vdash \neg V(a)$

Comment: Again, the argument is valid and the key premise is (5). Its justification is that there's no where else to put the probability mass. There will be an exam and there is only one day available for it, so all the probability has to be on that day. But this is wrong. $P(E)$ can in principle be anything at all.

9.8 Version 5 of the Students' Reasoning Applied to the One-Shot Surprise Exam Problem

1. $E \vee \neg E$
2. $l = \frac{2}{3}$
3. $(P(E) < l \wedge E) \rightarrow S$
4. $V(a) \leftrightarrow (E \wedge S)$
5. $P(E) < l$

$\vdash (E \wedge S \wedge V(a)) \vee (\neg E \wedge \neg V(a))$

9.9 Bingo!

Notice that assumption (4) has been strengthened to a biconditional. This is harmless and could have been done for the earlier versions. The strengthening amounts to accepting a rule that credits the teacher with speaking truthfully, $V(a)$, if what she said—that $(E \wedge S)$—is in fact true.
The upshot: In the one-shot surprise exam problem, the teacher must either speak falsely (e.g., by making a self-contradictory statement) or speak truly but with a probability no larger than l. Only by putting herself at risk of falsehood is it possible for her to speak truly in this case. By taking a risk (of speaking falsely) the teacher expands her scope of action.

The teacher can trade risk (of falsehood) for reward (being able to give a surprise exam).

9.10 Now the 6-shot (n-shot) Surprise Exam Problem

If the teacher is willing to run a risk, $r > 1 - l$ of speaking falsely, then it is not true that surprise could not lurk on the last day. This suffices to block the backwards induction and to undue the students' reasoning in the n-shot case.

But...The soundness of version 5 relies on the teacher being willing to accept a risk of at least $1 - l$ of speaking falsely. Given this, many would choose not to utter the one-shot surprise exam assertion. Honesty, integrity, prudence, or whatever may well prevent a reasonable person from saying something they know to have a chance higher then $\frac{1}{3}$ of being false. Better to keep silent.

What if the students know this?

9.11 Students' Reasoning Applied to the Augmented One-Shot Surprise Exam Problem

1. $E \vee \neg E$
2. $l = \frac{2}{3}$

3. $(P(E) \geq l \wedge E) \rightarrow \neg S$

4. $V(a) \leftrightarrow (E \wedge S)$

5. $P(E) \geq (1 - r)$

6. $r < (1 - l)$.

7. $(P(E) \geq (1 - r) \wedge r < (1 - l)) \rightarrow P(E) \geq l$

$\vdash$ $(\neg V(a) \wedge E \wedge \neg S) \vee (\neg E \wedge \neg V(a))$

Teacher's falsehood, validly deduced.

9.12 A deeper lesson lurks

More shots attenuate the teacher's verisimilitude scruples. Sufficient is:

1. In each period, initially the probability of the exam is less than l, and

2. The total probability of having the exam is greater than or equal to $(1 - r)$

This is trivially arranged for any l, and for any $(1 - r) < 1$, provided enough periods are available. Simply decide to give the exam with equal probability to every period. Again concretely with the l and $(1 - r)$ values given above, give a probability of $\frac{1}{5}$ of holding the exam on each of 5 days. The exam is held. If the exam is held on the fifth day, that morning the students will know that the exam will be held that day (assuming they are certain the exam will be held at all). There is only a 1 in 5 chance this will happen, which is above $(1 - r)$ and acceptable to the teacher. If the exam occurs on day 4, the students have a 50% certainty that morning, which is below l. And earlier is even better for the teacher's veracity. Further, with enough periods the teacher can set the probability of giving the exam to 1 and still have a probability $< r$ of not surprising the students, as we have just seen in the example.

9.13 In sum on the Surprise Exam Paradox

The n-period case generalizes the 1-shot case. The teacher can speak truly in this form provided the teacher is willing to undertake some risk, $r > 0$, of speaking falsely and providing n is large enough (given r and l). The teacher cannot be certain of speaking truly, but in this respect the case is like most. Usually, when we assert we take some chance of speaking falsely, even with the best of intentions. What is odd is to interpret a speaker otherwise. The only way the teacher could have spoken truthfully and given the surprise exam was to have spoken with some chance of speaking falsely.

The students erred in presuming incorrectly regarding the teacher's toleration of risk.

9.14 Question 2 on the Exam

2. In a 100-shot Repeated Prisoner's Dilemma game, played between the teacher and an unknown, but fully competent human subject, the teacher announces that she will gain the reward from mutual coöperation at least 2 times, net. That is, if P is the penalty for mutual defection and R is the reward for mutual coöperation, the teacher is asserting that she will get at least $98 \cdot P + 2 \cdot R$ points from the 100 trials. Can this assertion be plausibly justified? Why or why not?

9.15 The one-shot Prisoner's Dilemma game

The (one-shot) Prisoner's Dilemma game involves two players each with two strategies: C (coöperate) and D (defect). In strategic form the game is:

	C	D
C	R R	T S
D	S T	P P

with the requirement that $T > R > P > S$ and that $2 \cdot R > T + S$. Typically, even usually, in experiments $T = 5, R = 3, P = 1$, and $S = 0$. Since $T > R$

and $P > S$, there is only one equilibrium point (EP): both players play D. The dilemma, of course, is that if both players could play C, both would be better off, since $R > P$.

9.16 Game theory says the teacher is wrong

John Nash proved that, allowing pure and mixed strategies, every finite n-person game has at least one equilibrium point (EP).

The *Nash equilibrium solution concept* holds that the "solution" or predicted outcome of any game (among rational players) will be an EP. Since the one-shot prisoner's dilemma has only one EP, the Nash equilibrium solution concept predicts that both players will defect.

Further, for any fixed number, n, of iterations of Prisoner's Dilemma, a backwards induction argument suffices to prove that the n-period Iterated Prisoner's Dilemma (IPD or RPD) game still has exactly on EP: both players play D in each game.

9.17 Experiments say the teacher is right

It is interesting, and significant, that in the first human experiment with repeated prisoner's dilemma the human subjects were asked to record their thoughts as the game was being played. Comments such as

- "Perverse!"
- "Oh ho! Guess I'll have to give him another chance."
- "In time he could learn, but not in ten moves so:"
- "What's he doing?!!"
- "I'm completely confused. Is he trying to convey information to me?" and
- "This is like toilet training a child—you have to be very patient."

appear throughout the 100 iterations of the game. Even so, the two subjects jointly cooperated in 60 of the 100 iterations. By the lights of classical game theory this was a remarkably rewarding triumph of irrational behavior.[1]

[1] These results are not inconsistent with subsequent empirical findings.

9.18 Consider the one-shot PD

By defecting, the teacher can guarantee that she will receive a payoff greater than S. Similarly, by not announcing a surprise exam as she did, the teacher can avoid uttering a falsehood (in that case). If the teacher/player is willing to undertake some risk, however, there is also a chance that the teacher can do better than to receive P or to do without a surprise exam. Suppose the teacher has the following policy: play C with probability r and play D with probability $(1-r)$. If the other player plays the same strategy, then the expected return for each player, $E(r)$, is:

$$R \cdot r^2 + T \cdot (r - r^2) + P \cdot (1-r)^2 \tag{9.1}$$

on the harmless assumption that $S = 0$. Rearranging, the gamble pays off (in expectation) if

$$\frac{R \cdot r^2 + T \cdot (r - r^2)}{(1 - (1-r)^2)} > P \tag{9.2}$$

or

$$\frac{R \cdot r + T \cdot (1-r)}{(2-r)} > P \tag{9.3}$$

Obviously, setting $r = 1$ (for both players) yields an expected value of R, and for fixed T, R, and P (with $S = 0$) this maximizes the expected value.

More interestingly, note that on the left-hand side both the numerator and the denominator are positive, so fixing r and P, it is always possible to increase R and T sufficiently to make the inequality hold. Of course, on the standard assumptions of game theory, this should not matter.

9.19 This is but a crude model

of how a player might reasonably deal with risk in the PD game. Still it tells us something. Let $d = T - R$. Prisoner's Dilemma requires that $0 < d < R$. Fix d at some small value, say $d = 2$ in:

$$\frac{R \cdot r + (R+d) \cdot (1-r)}{(2-r)} > P \tag{9.4}$$

In the usual PD problem, $R = 3, T = R+d = 3+2 = 5, P = 1, S = 0$. Using the $R = R+d$ formulation, fix d, P, S. We see from expression (9.4) that $E(r)$

(the expected return for each player, assuming an independent probability of r of coöperating) increases as R increases, for any fixed value of r. Suppose the outcome values are dollar amounts. Let $R = 100$ ($R = \$100, T = \$102$ etc.).

9.20 Consider now our rational human teacher

and her rational human counter-player. Each sees the situation; each understands that the goal is to maximize dollars captured individually, not to get more dollars than the other player. It remains true that each player can with certainty avoid the sucker's payoff S by defecting. ... Surely, many players would reason that jointly they have much to lose by not coöperating and little extra to gain by individually defecting. Given that the other player has a coinciding interest in mutual coöperation, why not take a chance an play C? Surely as well, the strength of such sentiments increases with R. Suppose $R = \$1,000,000$. Suppose it is much larger than that.

9.21 Believe it? OK. Don't? OK.

Consider now the teacher's 100-play IPD game. Even if the teacher and her counter-player both find the one-shot PD reasoning above unconvincing, they surely would give pause when faced with 100 plays, each with a reward of \$1,000,000 for mutual coöperation. Do they both really want to follow a policy of defecting every time, no matter what? Must they conclude that their actions cannot have an effect—positive or negative—on the other player? Surely that is an awfully strong and unduely pessimistic assumption about a supposedly rational player. Why not try coöperating and if it is reciprocated continue to do so?

If these numbers aren't convincing for these arguments, increase the number of plays to a million and R to a billion. Increase them all you want. If the teacher's r is large enough given T, R, P, S, the chance she will get her points is, I think we should agree, an excellent one.

9.22 The point may be summarized in the following manner

Suppose each player reasons that there is some chance, r, that the counter-player can be induced during the 100 iterations to play C fairly often. Given the counter-player's interest in maximizing returns (as opposed to gaining relative points only), is it reasonable to assume, without probing, that $r = 0$? Let there be exactly n iterations of the game, known to players. Let $P = \$\frac{1}{n}$. Let $d = T - R$ be small, as above. Are there no large values of n and R for which it would not be folly not to probe the counter-player for mutual coöperation? This obviously rhetorical question answers itself in the negative.

9.23 Comments

- "Birds do it, bees do it..."
- People do it.
- Horses do it.

The Nash equilibrium is a solution *concept* for n-person games. The concept is that the games are solved by finding an EP (equilibrium point). That is where the players will end up. Long-established experience has shown that humans in (definitely) Iterated Prisoner's Dilemma games consistently do better than mutual all defect. Assuming no pervasive flaws in experiments conducted over a 50-year period, either the human subjects are consistently, egregiously irrational, or the Nash equilibrium concept is flawed (or both).

9.24 Contra Nash

The fundamental error on the part of the students was to assume that the teacher's r is, or even must be, 0. The students assumed that in making the tradeoff between risk (of speaking falsely) and reward (being able to offer a surprise exam), the teacher would place no value (or no sufficiently large value) on the reward, at the expense of taking some risk. Similarly, I have argued, in the definitely IPD both players face a tradeoff between risk (of getting the sucker's payoff, S) and reward (achieving P very often during the iterations). The flaw in the Nash equilibrium solution concept (at least for IPD) is to impose the assumption that both players are unwilling to trade any risk at all (however small) for any reward at all (however large).

9.25 On Beyond Nash (Something positive)

Suppose we have a game with n players (in IPD, this n is 2). At the conclusion of play, each player i has played some strategy, s_i. This constitutes the *strategic configuration* or *SC.*

$$SC = (s_1, s_2, \ldots, s_n) \tag{9.5}$$

Let $H_i(s_1, s_2, \ldots, s_n) = H_i(SC)$ be the payoff to player i given the strategies, including i's, played in *SC.* An *equilibrium point* or *EP* is any SC^* such that for each $i = 1, 2, \ldots, n$

$$H_i(SC^*) = \max_{s_i} H_i(s_1^*, s_2^*, \ldots, s_i, \ldots, s_n^*) \tag{9.6}$$

In other words, an *EP* (equilibrium point) is an *SC* (strategic configuration) such that no individual player can (or could) unilaterally do better by picking a different strategy than the one the player has in *SC.* The Nash equilibrium solution concept for games is that among rational players every game will conclude at an *EP.*

9.26 Generalizing the Nash solution concept

Given s, a particular *SC,* define the *improvement vector* for s or IV_s, as

$$IV_s = (c_1, c_2, \ldots, c_n) \tag{9.7}$$

where c_i is the *count* or number of ways i distinct players can jointly alter their strategies in such a way that each of the i players does equally well or better. More carefully, if $c_3 = 2$ then there are two distinct groups of 3 players who collectively have strategies that if taken, while all the strategies outside the group remain as in s the relevant *SC,* would make everyone in the group no worse off. Now a single group of 3 (or whatever) might have many ways to do this. IV_s ignores this and only counts the number of groups of a given size with 1 or more opportunities. (We can also define a *strict improvement vector* or *SIV,* as an *IV* in which everyone in every group is strictly better off.)

9.27 Points arising

1. The *IV* concept is a generalization of the *EP* concept. $IV_{SC} = (0, c_2, \ldots, c_n)$ if and only if *SC* is an *EP*.

2. The payoff vector, H_s, of a strategic configuration s is denoted:

$$H_s = (H_1(s_1), H_2(s_2), \ldots, H_n(s_n))$$

3. In the one-shot PD, the *table of space of improve vectors* or the *improvement space table* or *IS* is:

SC	IV	H_{SC}	target Hs
(D, D)	$(0, 1)$	(P, P)	(R, R)
(D, C)	$(1, 0)$	(T, S)	(P, P)
(C, D)	$(1, 0)$	(S, T)	(P, P)
(C, C)	$(2, 0)$	(R, R)	(T, S)
			(S, T)

where the target Hs are the payoffs if the associated IVs are realized.

4. We might similarly define how the players might act to make things worse. This would result in a disimprovement space table.

5. We say that SC_x *strictly dominates* SC_y iff for all $i = 1, 2, \ldots, n$,

$$H_i(s_i^x) > H_i(s_i^y)$$

or equivalently

$$H_x > H_y$$

Note that (C, C) strictly dominates (D, D) in the one-shot PD. (Hence the dilemma for the Nash solution concept.)

6. In the n-shot IPD, there is a correspondingly larger improvement space, generated by the Cartesian product of the one-shot spaces.

7. Any well-defined game will have well-defined improvement and disimprovement spaces. In learning to play the game, or in rational deliberation about it, players should be thought of as exploring these spaces.

   At risk of being wrong, but with potential for reward, players can form hypotheses about their counterparts, as play unfolds (actually or prospectively).

8. Given a well-defined game, with its improvement/disimprovement spaces and attendant payoff vectors, how play actually unfolds will depend upon the strategy formation and selection algorithms employed by the players.

9. In general and under realistic assumptions regarding strategy formation and selection algorithms, ending at an EP will be unusual in games with IVs having non-zero terms past the first position with attendant large payoffs.

   Note: This is what the now large literature on computational dynamics in games has yielded—well something consistent with this.

10. We can look for and expect robustness results for classes of learning algorithms in games. Think: replicator dynamics, reinforcement learning. Convergence (and speed of it) to SCs with particular properties. Example: in Q-learning, with potential reward R and risk r, the players will achieve the R with such-and-such characteristics.

    And that is how we can improve predictions and better understand decision processes (human or not).

9.28 Note vis à vis Evolutionary Game Theory

- Views here accord in many ways with evolutionary game theory, e.g., Gintis in *Game Theory Evolving*
- And also differ on the underlying theory of rationality. EGT wishes to save the classical theory (and the Nash equilibrium) by finding refinements (via evolutionary processes). How? We do what we do because of the way we're built.
- Yes, but... Why are we built the way we are?

  Why has evolution produced the traits we have? Why are they adaptive (and rational in an extended sense)?
- Suggestion here: Rationality, properly understood, prompts us to be willing to take risks in strategic contexts. It's true that we are cognitively bound, but that's not—at least superficially—why we do better than AllD in repeated prisoner's dilemma.

9.29 What Next?

Conjecture 1 (SAGE: Smart Agents Go Efficient) *Under reasonable discount rates, reasonable learning regimes, and indefinite (or definite but sufficiently long) repeated play of a stage game, agents can stably get to Pareto optimal outcomes.*

This is a conjecture and cannot be proved, but it could be disproved (or narrowed) in certain cases and demonstrated in other particular cases. Interesting to see when it obtains and when not. Let's take a preliminary look.

NB smart ≈ probes the counter-player, collects information, revises accordingly

9.30 IPD Again

Suppose you are playing IPD and are confident of having a reasonably long run of it. Consider the following learning strategy.

1. Pick a run length, L_r over which you will explore the strategy space. Here, let $L_r = 3$, so you have 8 possible (nonreactive strategies).

2. Pick an iteration length, $L_i =$ the number of iterations you will play each tested strategy before deciding how to exploit the information you get from your exploratory experiment. Here again, let $L_i = 3$.

3. In random order, play each strategy L_i times in succession, recording what happens.

4. Evaluate the results and choose a revised policy of play.

9.31 What if Your Counter-Player Plays Tit-for-Tat?

No.	Strategy	Reward Sequence	Reward
1	(000)(000)(000)	(TPP)(PPP)(PPP)	T+8P
2	(100)(100)(100)	(RTP)(STP)(STP)	3T+R+3P+2S
3	(010)(010)(010)	(TST)(PST)(PST)	4T+2P+3S
4	(001)(001)(001)	(TPS)(TPS)(TPS)	3T+3P+3S
5	(110)(110)(110)	(RRT)(SRT)(SRT)	3T+4R+2S
6	(101)(101)(101)	(RTS)(STS)(RTS)	3T+2R+4S
7	(011)(011)(011)	(TSR)(TSR)(TSR)	3T+3R +3S
8	(111)(111)(111)	(RRR)(RRR)(RRR)	9R

9.32 Evaluation

Recall: $T > R > P > S$ and $2R > T + S$.

Well, fixing R, P, S, it's always possible to increase T so that $v(8) < v(i), i \neq 8$. OK, but $v(5) > v(2), v(4), v(6), v(7)$ What about $v(5) = v(110)$?

When is it that $v(8) = 9R > v(5) = 3T + 4R + 2S$? Assume $S = 0$. Then:

$9R > 3T + 4R$ or $R > \frac{3}{5}T$

Note: Let $L_i = n$, then $v(8) = n3R$ and $v(5) = nT + nR + R$ (assuming $S = 0$)

$n3R > nT + nR + R$ when $R > \frac{n}{(2n-1)}T$

What about $v(3) = v(010)$? $R > \frac{T(n+1)+P(n-1)}{3n}$

$\lim_{n\to\infty} \frac{T(n+1)+P(n-1)}{3n} = \frac{T+P}{3}$

9.33 Discussion

So, if R is at all close to T, then the learner with this very simple regime will quickly learn to cooperate with the Tit-for-Tat player.

Also, for any values of R and T, the lesson can be learned with sufficiently high L_i.

What does it cost the Tit-for-Tat player to teach the lesson? When will it be worth it?

This is a very limited test and confirmation of the SAGE conjecture. What could we do to broaden and deepen our testing of it? What happens if we, say, try all m-1 strategies and both players are probing and learning?

Recall: smart $\approx$ probes the counter-player, collects information, revises accordingly. What can we say about optimality and smartness?

Chapter 10

(T:10/08) Finishing the Surprise Exam and Starting GAs

10.1 Announcements, etc.

I would like to invite you to a colloquium by Jon Elster of Columbia University on Wednesday October 9th, at 4pm, in Logan Hall in Room G-17. The colloquium is sponsored by the PPE Program.

Last time: the surprise exam, but only question 1. Today, on to question 2, including a few new foils.

Next time: I'll hand out an assignment, due later, after midterm quiz. Also next time: chapter 17 of the AGEbook, "Notes on Computability and Complexity."

What about chapters 13-15 of the AGEbook? Examples. Will discuss as the opportunity arises. Considering an extra class session before chapter 17.

Chapter 11

(R:10/10) Starting GAs

Assignment handed out:

***

Assignment #1 for OPIM 325, Fall 2002

11 October 2002

Read the review article, Michael W. Macy and Robert Willer, "From Factors to Act ors: Computational Sociology and Agent-Based Modeling," Annual Review of Sociolo gy, 2002, vol. 28, pp. 143-66. (Handed out; see also: `http://opim.wharton.upenn.edu/~sok/papers/m/macy-willer-2002-ann-rev-sociology.pdf`.)

Pick one of the recent articles mentioned in the review paper, read it, write up a short description and critical assessment of it (5 pages max), and prepare a short presentation (5-10 minutes max) on it.

The write up and the presentation are due in class on October 29, 2002. We'll de vote that class period to your presentations and their discussion.

Feel free to ask questions along the way, but note that I will be out of town an d mostly unavailable from the afternoon of October 24 through the morning of Oct ober 29.

***

Then, in the AGEbook, chapter 11, "Notes on Evolution: 1," followed by chapter 12, "Notes for Introducing Genetic Algorithms."

I asked you to read, and I handed out, chapter 1 of Melanie Mitchell's book on GAs.

[Did not finish or even start the schema theorem. Will do this Thursday 10/17]

Chapter 12

(T:10/15) Coding, Computation, and Complexity

BIG topics; will do a speed drill. First, coding-computation-slides.pdf, which is on our site.

Then chapter 17 of the AGEbook.

If time permits: GAs and knapsack problems; two-market GA.

For next time:

1. Will start out with the schema theorem for GAs.
2. Read "What Is Life?" handed out today. (skimming it is ok)
3. Read chapter 19 in the AGEbook, "Cellular Automata and the Game of Life"
4. Star Logo demo from Annapurna

Remember: we have a quiz in class on 10/22. On readings/discussion to date. Closed book. OK to bring one sheet (two sides) of notes.

Chapter 13

(R:10/17) Misc.

Remember: we have a quiz in class on 10/22. On readings/discussion to date. Closed book. OK to bring one sheet (two sides) of notes. Today's material *not* on exam. Format 3–5 questions with short answers.

Today:

1. Finish up from last time, chapter 17 of the AGEbook on computation and complexity. Two topics, briefly: (a) P, NP, etc.; complexity (b) what all this means for game theory and the theory of rationality.
2. Very briefly, discussion of the two-market GA (see reading in AGEbook)
3. Demo/discussion by Annapurna on Star Logo
4. Game of Life

Note: still dangling: Holland's schema theorem for GAs. Important, especially in light of our expanded view of rationality theory; we'll get to it.

Game of Life Web sites:

1. http://www.math.com/students/wonders/life/life.html
2. http://www.bitstorm.org/gameoflife/

Chapter 14

(T:10/22) Midterm Exam

Chapter 15

(R:10/24) Genetic Programming

Movie of Koza III, plus discussion.

Chapter 16

(T:10/29) In-class presentations

Assignment #1.

Chapter 17

(R:10/31) GP Experiment

Finish up student presentations of Assignment #1. Discussion of Dworman, Kimbrough, Laing, PG96 paper (handed out). Very brief discussion of Kimbrough, Lu, Wood, and Wu GECCO-2002 paper (handed out).

Read for next time (handed out):

Grim et. al, chapters 4 and 5, and some of 6. Sigmund, 1993, *Games of Life* chapter 8.

Chapter 18

(T:11/5) Spatialization

18.1 Administration

We're close to the syllabus and will remain so, with minor deviations. Handouts for today:

- "Zones of Cooperation in Demographic Prisoner's Dilemma" by Joshua Epstein

- "Why Agents? On the Varied Motivations for Agent Computing in the Social Sciences" by Robert Axtell

Today: focus on Sigmund, chapter 8 (in part) and Grim et al., chapter 4. First, discuss final projects. Also, one more assignment, similar to the first. Next time (a) Grim chapter 5 and the beginning of 6, (b) today's handouts.

18.2 Grim et al. chapter 4

- Begins with a standard, familiar, but powerful and clear reference to Hobbes, who perhaps should be credited with first recognizing the Prisoner's Dilemma

- There follows introductory material, which should be a review for us.

- §4.3: Generosity in an Imperfect World

Nowak and Sigmund article, "Tit for Tat in Heterogenous Populatins," *Nature* 355 (1992), pp. 250-2. Focus on *stochastic reactive strategies.* Tit for Tat is a deterministic reactive strategy.

18.3 Nowak & Sigmund (1992)

- Instead of, e.g., $\langle 1, 1, 0\rangle$ (deterministic Tit for Tat) we have, e.g., $\langle 0.99, 0.99, 0.01\rangle$ (stochastic Tit for Tat).
- Simplify to $\langle 0.99, 0.01\rangle$, since assuming an infinite game the first move doesn't matter. (See formulas page 169 of Grim et al.)
- Results? See also Sigmund (1993), pp. 195f. 100 random strategies in the $(0.01, \ldots 0.99)^2$ grid. Usually, AllD (or something close) takes over. Sometimes, however, something close to Tit for Tat is present. Dyanmic response! See Figure 1 of Grim et al. In the end GTFT takes over, $\langle 0.99, 0.3\rangle$. Discuss. Quite remarkable. See Figure 1 in Grim et al., chapter 4.

18.4 Sigmund (1992), pp. 192f

- Tit for Tat, as we know is vulnerable to errors. Endless cycle of cooperate and defect.
- See page 193, Figure 8.2. Zone of cooperation.
- The zone has surprising dynamics. Although $\langle 0.4, 0.1\rangle$ is an ESS, it is not stable; the dynamics drives the system away.
- Page 195: "Strangely enough, the situation with q is just the opposite. If all members of the population used $p = 90$ per cent, say, the the value $q = 35$ percent is evolutionarily stable, and also attainable, in a sense unavoidable even."

18.5 Sigmund (1992), pp. 203f

- Enhancing cooperation: keep in touch. Why? "Preferential assortment helps cooperation to spread." As we saw from Axelrod, too. I would add: With small steps, the absolute value of the increment from defection is reduced. (Recall: utility theory is only on an interval scale. Fitness is on a ratio scale.)

- Tragedy of the commons; free riders. When cooperation fails. Why?

18.6 Grim et al., §4.4 Spatialization

- On a grid. Review of the Game of Life.
- Page 174. Strategies play Moore (8) neighbors, with repetition (200 times). Convert to neighbor's strategy if it gains more points.
- 64×64 array; torus. Each cell occupied by a randomly-chosen m-1 IPD strategy.
- See Figure 5 in Grim et al. Note initial success of 'vicious' strategies.
- But (page 177) it doesn't always work out this way. Discuss.

18.7 Grim et al., §4.5 Deeper Strategies

- That is, m-2 strategies, all 64: $\langle i1, i2, cc, cd, dc, dd\rangle$. Note: Hard to interpret. Figure 6, a typical run. Takeover by STFT 2 and TFT.
- Page 181: "Total conquest by any single strategy in a 'two-deep' competition turns out to be the exception rather than the rule, however. A much more common pattern of evolution is one toward a stable equilibrium among several strategies." See Figure 7.
- "...in the move from one generation to the next cells adopt the strategy of the highest-scoring neighbor, and thus the fact that a cell stayw with a particular strategy may indicate not that *that* cell is doing particularly well with that strategy but that a neighboring cell is. This allow for the possibility of a 'buffer situation', which appears to be what is happening here."

18.8 Grim et al., §4.6 Greater Generosity in an Imperfect Spatial World

- "What happens when we turn to a *spatialized* form of the stochastic Prisoner's Dilemma?"
- Repeated the Nowak and Sigmund experiments, but on a 100×100 grid.
- "With a full 121 strategies represented, such an array is highly sensitive to initial configuration..." See Figure 9. Often, as here, conquest by $\langle 0.99, 0.10 \rangle$.
- "It is clear that a primary factor in these first resutls is the limit of our spatial array..."

18.9 Grim et al., §4.6 Greater Generosity...

- "None of Nowak and Sigmund's stochastic strategies has a true score of zero against any other. In the evolution plotted in figure 1, the proportion of stochastically imperfect TFT falls at one point before its rise to somewhat more than half of its original representation. But the more generous $\langle 0.99, 0.2 \rangle$ falls to less than a millionth, and $\langle 0.99, 0.3 \rangle$ falls toless than a trillionth." On a small grid, rounding error eliminates such strategies.
- Also, Grim et al. find that often "there is a direct victory by $\langle 0.99, 0.1 \rangle$ without the necessity of prior conquest by a statistically imperfect TFT $\langle 0.99, 0.01 \rangle$."
- What to do?

18.10 Two more sophisticated studies

2. 8 strategies (from 121) randomly dropped on a 100×100 grid. When a strategy dies out, another is randomly drawn from the 121 and randomly seeded in $\frac{1}{8}$ of the cells. See Figure 10 for two typical end results, dominance by $\langle 0.99, 0.4 \rangle$ and $\langle 0.99, 0.6 \rangle$.

"Although it may take longer to develop, the basic result is the same if we start with 16 strategies rather than 8, if we introduce only 1% of an alternative strategy when one dies, or both." Etc.

". . . it became clear that it was the possibility of small clusters that was crucial to ecological dynamics."

18.11 Two more sophisticated studies

3. "Rather than replace a full 12.5% of the display only on the death of a strategy, we replaced a clustered 0.06% of the display with a randomly chosen stragy at each generation."

   See Figure 11. "In this third model an even clearer dominance by $\langle 0.99, 0.5\rangle$ and $\langle 0.99, 0.6\rangle$ was evident.

But, page 189, "No strategy is impervious to invasion by small clusters of all other strategies."

18.12 Invasion Dynamics

"[I]t becomes important to distinguish between different notions of invasion." See Figure 12.

1. growth (for a while)
2. sustained growth
3. invasion to conquest

"Although GTFT is still impervious to invasion by a single unit of any other strategy, a GTFT of $\langle 0.9999999, 0.3333333\rangle$ is invadable to conquest by clusters of as small as four units of, for example, $\langle 0.9999999, 0.5\rangle$."

18.13 Further studies

Fascinating patterns. See Figure 13.

- GTFT thus (fig. 13) is invadable to conquest by more generous strategies.

- "Strategies more generous than GTFT, up to an including $\langle 0.9999999, 0.6666666\rangle$, prove invadable by other strategies but seem invadable *to conquest* by no others in at least standard patterns of small clusters. Here our work took the form of an empirical survey..." (page 192)
- "Spatialization alone seems to favor an important increase in the level of generosity one can expect to evolve in a Hobbesian model."

18.14 Let's Talk

- What's going on here? Why does spatialization have such an effect? (And how might you go about discovering the answer to this question?)
- Arguably, spatialization yields a more accurate model. Examples?
- How might we improve or add detail to the stochastic, m-1, spatialized model?
- What do you think will happen when other games are tried? When PD is tried with different outcome values?

Chapter 19

(R:11/7) Finishing up Grim

19.1 Administration

Discussion today:

1. Finish up on chapter 4, spatialization, of Grim et al.
2. Briefly discuss chapters 5 and first pages of 6 in Grim et al.
3. "Why Agents? On the Varied Motivations for Agent Computing in the Social Sciences" by Robert Axtell
4. "Zones of Cooperation in Demographic Prisoner's Dilemma" by Joshua Epstein [will do next time]

Next week: Tuesday, Epstein and Axtell, chapters I–III; Thursday, chapters IV–VI.

19.2 Administration

General comments. This is about (nearly) the last of the very theoretical, foundational classes. I want to move more towards discussing particular models and applications. We move in that direction with Epstein and Axtell next week. Still, there is some material on fundamental methods coming up (e.g., reinforcement learning, learning classifier systems, swarm optimization). I'm going to try to compress this and give us more time to discuss particular models and applications. Thoughts and suggestions are welcome.

19.3 Grim et al., chapter 5

- Read through (and discuss here) for appreciation of more general models, e.g., Real Life. Skip §5.5 on infinite-value logics (although it's quite fascinating).
- §5.1, Real Life. "Why not construct a variation of Conway's Game of Life that countenances eral-valued degrees of life and death? We call this game 'Real Life.' It turns out that Real Life, unlike Conway's Game of Life, exhibits a sensitive dependence on initial conditions that is characteristic of chaotic systems." [Discuss: rule translation to reals; chaos.]

19.4 §5.3 Real-Valued Prinsoner's Dilemmas

- "No matter how convincingly one spins the standard Prisoner's Dilemma story, however, it remains artificial in a number of respects. 'Confession' is treated as an all-or-nothing affair, for example, with distinct punishmens allotted in terms of it. But surely the normal case is one in which there are *degrees* of cooperation..." Model the payoffs as in Figure 10.
- Experiments. "In general, we took our results to confirm a Nowak and Sigmund-like result for the spatial and continuous-valued Prisoner's Dilemma."
- But... see §5.4.

19.5 Note: Error on page 211

The formula given is nonsense, probably a serious typo. This is better:

$$f(x,y) = f(x|y) = \begin{cases} P + (R-P)x & x = y \\ P - (P-S)x + (R-S)y & x > y \\ P - (T-R)x + (T-P)y & x < y \end{cases} \tag{19.1}$$

(and $f(y|x) = f(x|y)$.)

And on page 217: elementary error in the diagram.

19.6 §5.4 Pavlov and Other Two-Dimensional Strategies

- "A reactive or *one-dimensional* strategy is one for which each move, after the first, is strictly determined by the opponent's previous move. A *two-dimensional* strategy is one for which each move, after the first, is strictly determined not by [sic] merely by its opponent's previous move but by the previous *pair* of moves by both players."
- PAVLOV or "win-stay lose-shift". "PAVLOV cooperates only after mutual cooperation or mutual defection."

19.7 §5.4 Pavlov and Other Two-Dimensional Strategies

"Nowak and Sigmund show PAVLOV to be superior to GTFT in a nonspatial competition. Is PAVLOV still superior to GTFT once we spatialize the Prisoner's Dilemma? Is there another a [sic] two-dimensional strategy that is superior even to PAVLOV in a spatial context? The answer to both questions is 'yes'. PAVLOV is consisently beaten in spatial competition by [$\langle 1, 1, 0, 0, 0\rangle$] the continuous-valued [sic: two-dimensional?] counterpart to a GRIM strategy." [This last phrase is confused.]

Then they try clustering and TFT does better than GRIM. All this is for *bivalent* two-dimensional strategies, in a spatial regime.

19.8 §5.6 Discrimination

- Moral ≈ inevitable? [Recall Skyrms and fairness; same thing?]
- P. 229: "We call a strategy *discriminatory* it it adopts on strategy against players of another color and a different strategy against strategies of its own color."
- In a first tournament we pitted the eight possible bivalent one-dimensional or reactive strategies against two forms of DTFT. The two DTFTs quickly eliminated all the first-order strategies with the possible exception of clusters of TFT." Figure 18.

19.9 §5.6 Discrimination, #2

- “What if we include all discriminatory strategies in such a tournament—all reactive strategies that may treat others differently than they treat themselves?” Examples:

DTFT	$\langle 110, 000\rangle$
DQ	$\langle 111, 000\rangle$
TFT	$\langle 110, 110\rangle$

- Pp. 230–1: Evolved under a genetic algorithm regime. Crossing over with maximum-valued neighbor (one of eight). Figure 19.

- In a field of discriminatory strategies, TFT is no longer the optimal strategy: discrimination is more successful than impartial fairness.”

19.10 §5.6 Discrimination, #2

- “It is clear from even these simple examples that discriminatory strategies play an important and domineering role in the simple game theoretic environments at issue. This comes as close as a formal argument can, we think, to making the philosophical point that socially dominant strategies need not in any way be genuinely ethical cases.”

19.11 §5.7 Continuity and Forgiveness

- “What happens if we once again open our modeling to encompass a range of degrees of cooperation and defection?”

- “We modeled a *veil of ignorance* by introducing a stochastic parameter v as a measure of the extent to which a strategy is able to correctly identify the color of its opponent.” Made a little difference.

- Figure 20, simple experiment. Figure 21, some optimism, given the specially encoded representation, “suggests that a high enough veil of ignorance may allow even Forgiving TFT to prevail over discrimination.”

19.12 Grim et al., Chapter 6, pp. 237–245

- Later, this is heavy-going stuff, although clear and nicely done. We just need to note the results.

- Page 240 "Average scores within the repeated unit of play alone can thus be taken as a limit toward which average scores will tend over finite games of increasing length. What we take as the score of Strategy 1 versus Strategy 2 in an infinitely iterated game is simply that limit."

  "Though scores fo infinitely iterated games are used throughout this chapter, the basic results will also hold for finite games of sufficient length."

- Note and compare Figures 1 and 2.

19.13 Grim et al., Chapter 6, pp. 237–245

- Pages 242–3, "A question often arises. . .

  Given a computer big enough or fast enough, a mind without limits at attention span or memory or attention to dtail, is there some systematic computation that will tell us in each case whether an embedded finite array will result in progressive conquest or not?

  The work outlined below answers this question firmly in the negative. There is no effective procedure that will in each case tell us whether or not a given configuration of Prisoner's Dilemma strategies embedded in a uniform background results in progressive conquest."

19.14 Grim et al., Chapter 6, pp. 237–245

- Pages 242–3, footnote 12:

  "The question used to illustrate undecidabiltiy throughout the chapter is whether a single strategy will dominate to conquest. This is only an example, however. Given the basic method of proof other properties of arrays can be shown to be undecidable in precisely the same way: whether TFT will ever be completely extinguished in an array, for example, or whether good in the guise of a certain level of generosity will eventually triumph."

19.15 Axtell, "Why Agents?" OOP+3 cases. First OOP.

- OOP metaphor. SmallTalk, C++, Java, even Microsoft. Will use Java nomenclature.
- Program: a collection of objects, sending messages to each other; objects respond to messages and do their things. Decomposition idea. World view.
- Classes, objects, variables. methods. Public, private. Encapsulation.
- Inheritance. Specialize; varieties of things.
- Classes: countries, cities. Objects: Sweden, Philadelphia.

19.16 OOP

- Define classes: `Agent`, `Population`, `InteractionRegime`, `DataCollection`. Add "controller" class/object that: initializes and instantiates the objects, then sets the world in motion. Population interacts (including updating), record data: until done.
- Conceptually simple and elegant. Facilitates maintenance and code reuse (yours and others).

19.17 Traditional Simulation

- Pseudo-random numbers
- Monte Carlo
- Discrete event
- Well-established, widely-used, and accepted in practice

- Often, low-status academically (compared to mathematical models)

19.18 First Use: Agent Models as Classical Simulation

- "*not* the most important use of agents, but it is the use that best fits the conventional meaning of 'simulation'"
- "a useful check on the numerical solution" but "not a common use of agent-based computational models."
- Another form/version of Monte Carlo simulation (really: discrete event simulation). Example: queuing system.
- In addition, useful for people with weak mathematical skills/intuitions

19.19 Second Use: Partially Soluble [sic: Solvable] Models

- When you can write down a mathematical model, but you can't solve it. Explore the solution space.

  Yes, and an amplification: make discoveries and find patterns in the solution space. John Miller and 'model busting.' SOK and candle-lighting analysis.
- Example: equilibria exist are are effectively uncomputable.
- Example: equilibrium not attend by boundedly rational agents.

  Santa Fe stock market
- Example: on the path(s) to equilibrium
- Etc.

19.20 Third Use: Models Ostensibly Intractable or Provably Insoluble [sic: Unsolvable]: Agent Computing as a *Substitute* for Analysis

- This happens more than is admitted and than you might think.
- Either: impossible or way too hard.
- Gives an example. Grim et al. example today. SOK example earlier: Diaphontine equations.

  In general, really almost all the interesting problems are like this.

Chapter 20

(T:11/12) General Discussion of GAS

Chapter 21

(R:11/14) Finishing Up GAS

21.1 Administration

- Plan: next 3 sessions, a paper a time, with a model; will discuss at beginning of class. We'll discuss these papers first in the hour, then I'll move on to a lecture, for which there may not be any required reading.

- Next time (11/19): Zero-intelligence traders

  You can find the paper on JSTOR. Here is the reference:

  "Allocative Efficiency of Markets with Zero-Intelligence Traders: Market as a Partial Substitute for Individual Rationality," Dhananjay K. Gode, Shyam Sunder, *The Journal of Political Economy,* Vol. 101, No. 1. (Feb., 1993), pp. 119-137.

- A week from today (11/21): Assignment #2. Read Hoffmann's paper, "Entrepreneurs and Norm Dynamics: An Agent-Based Model of the Norm Life Cycle," which is at

  `http://opim-sun.wharton.upenn.edu/~sok/papers/`

  Write a two-paper (2-3 pages), discussing his model and suggesting improvements/alternatives. (a) Briefly describe his model. What is it a model of? What are its strengths and weaknesses (what does it capture, does it fail to capture?) (b) How might the model be improved and experiments run to obtain additional results?

- (11/26): Beer game paper at:

`http://opim-sun.wharton.upenn.edu/~sok/-`

`papers/fmec-dss-si/kwz/beer-game-dss-final.doc`

Slides at

`http://opim-sun.wharton.upenn.edu/~sok/-`

`sokpapers/2000-1/beergameslides.ppt`

- Note: get to work on your term papers. Due: at the final exam date/time.

- Today, Epstein and Axtell, GAS, chapters IV, V, and VI. Opportunity to look at more realistic computational models.

21.2 *Growing Artificial Societies*, Chapter IV

- "Sugar and Spice: Trade Comes to the Sugarscape"

  Perhaps the main, most crucial chapter in the book. They take on General Equilibrium theory in economics. Page 136:

  > In many ways, the central question of economic theory is this: To what extent can economic markets efficiently allocate goods and services among agents?

  'Efficiently'? Two meanings? Which is primary here?

21.3 *GAS*, Chapter IV

Neoclassical General Equilibrium theory and the Walrasian auctioneer. Recall page 17:

> In the usual general equilibrium story it is assumed that every agent "takes" a price issued from the top down, by the so-called Walrasian auctioneer.

and page 101:

> The neoclassical theory of general equilibrium describes how a single centralized market run by a so-called auctioneer can arrive at an equilibrium price vector for the entire economy—a set of prices at which all markets clear. The image of an auctioneer announcing prices to the entire economy is quite unrealistic; no individual or institution could ever possess either complete knowledge or agent preferences and endowments or sufficient computational power to determine the appropriate prices. And even if market-clearing prices were somehow identified, why wouild all agents use them, why would all agents be price-takers?

And why would anyone use such a model for policy making?

21.4 *GAS*, Chapter IV: Another Kind of Model

> A more recognizable image is presented by Kreps...
>
> > ... we can imagine consumers wandering around a large market square, with all their possessions on their backs. They have chance meetings with each other, and when two consumers meet, they examine what each has to offer, to see if they can arrange a mutually agreeable trade. ... If an exchange is made, the two swap goods and wander around in search of more advantageous trades made at chance meetings.
>
> We implement trade in precisely this fashion, as welfare-improving (that is, mutually agreeable bilateral barter between agents. No use is made of an auctioneer or any similar artifice. [page 101]

21.5 *GAS*, Chapter IV: Another Kind of Model

In their distributed model here

> Individual agents do not use any nonlocal price information in their decisionmaking. Because price formation is local, this is a model of *completely decentralized exchange* between neoclassical agents. Later we relax some of the least realistic aspects of the

> neoclassical setup; for example, giving the agents finite lives and nonfixed preferences.
>
> The main issue we address is the extent to which interacting agents are capable of producing *socially optimal* outcomes, that is, allocations of resources having the property that no agent can be made better off through further trade.
>
> ...It seems natural to think of market processes as a form of social computation, with the agents operating as distributed processing "nodes" and the flow of commodities serving as inter-node communication. Each node (agent) executes a local optimization algorithm (purposive behavior), attempting to maximuze a local objective (utility) function through decentralized interactions with other nodes (agents). The market as a whole—the social computer—tends toward a globally optimal allocation of goods, as if it were "attempting" to *compute* such an allocation. In this chapter we study how the success of this social computation depends on agent specification.

21.6 Spice: A Second Commodity, pp. 96f.

> A each position [on the sugarscape] there is a sugar level and capacity, as before, as well as a spice level and capacity.
>
> Each agent now keeps two separate accumulations, one of sugar and one of spice, and has two distinct metabolisms, one for each good. These metabolic rates are heterogeneous over the agent population... Agents die if *either* their sugar or their spice accumulation falls to zero.

See Figure IV-1. Sugar mountains northeast and southwest. Spice mountains northwest and southeast.

21.7 *GAS*, Chapter IV: The Agent Model(s)

1. The Agent Welfare Function, p. 97.

$$W(w_1, w_2) = w_1^{m_1/m_T} \cdot w_2^{m_2/m_T} \tag{21.1}$$

OK: $m_T = m_1 + m_2$. w_i = accumulation. m_i = metabolism. The idea is to compute what you should seek (at the margin) in order to avoid starvation. Question: Could we get this more directly? What would be a simpler algorithm? What would be a more intelligent algorithm?

21.8 *GAS*, Chapter IV: The Agent Model(s)

2. The Agent Movement Rule in the Presence of Two Commodities, p. 98f.
   Multicommodity agent movement rule **M**:

   - Look out as far as vision permits in each of the four lattice directions, north, south, east, and west;
   - Considering only unoccupied lattice positions, find the nearest position producing maximum welfare;
   - Move to the new position;
   - Collect all the resources at that location.

   The effect is "profound." See animation IV-1.

21.9 *GAS*, Chapter IV: The Agent Model(s)

Recall:

> We implement trade…as welfare-improving (that is, mutually agreeable) bilateral barter between agents. No use is made of an auctioneer or any similar artifice.

1. Internal Valuations, pp. 102f.

   Basically, agents wander around until encountering one another. If one values sugar more than spice and the other values spice more than sugar, then there is a potential for mutually-beneficial trade. But at what price? Geometric mean of $[MRS_A, MRS_B]$. Why? Full disclosure and no bargaining. Future work (footnote 14, page 104).

2. The Trade Algorithm, pp. 104f.

   Summary, page 107:

> The Sugarscape interagent trade rule can be summarized as follows: If neighboring agents have different marginal rates of substitution then they attempt to arrange an exchange that makes them both better off. Bargaining proceeds and a trade price is "agreed" to. Quantities of sugar and spice in proportion ot the trade price are specified for exchange. If exchangeof the commodities will not cause the agents' *MRS*s to cross over then the transaction accurs, the agents recomput their *MRS*s, and bargaining begins anew. In this way nearly Pareto optimal allocations are produced locally.

21.10 Markets of Bilateral Traders, pp. 108f.

> The immediate question for us—having banished the auctioneer and all other types of nonlocal information—is *whether our population of spatially distributed neoclassical agents can produce anything like an equilibrium price through local interactions alone.* It turns out that there is a definite sense in which they can! [page 108]

And that is? See Figure IV-3.

> There is a sense in which this completely decentralized, distributed achievement of economic equilibrium is a *more* powerful result than is offered by general equilibrium theory, since *dynamics* of price formation are fully accounted for, and there is no recourse to a mythical auctioneer. [page 111]

21.11 *GAS*, Chapter IV: But...

> The decentralized economy is always far from general equilibrium in this sense.
>
> This result is of prime significance. For whenever the actual trade volumes are less than the general equilibrium ones, agent society is not extracting all the welfare from trade that it might. If the agents could coordinate their activities beyond their local neighborhoods they could all be made better off. Here we see that

> *even though* ***T*** *produces exchanges that are nearly Pareto-optimal locally, the resulting market has from from optimal welfare properties globally.* [page 116]

21.12 *GAS*, Chapter IV: Refinements and Discussion

They proceed to add distributions of agent vision, sex, finite lives, 'culture', pollution, foresight, credits (loans between agents) to the model. And what do they see?
Policy Implications [pp. 136–7]

> Do plausible departures from the axioms of general equilibrium theory produce markets that behave almost as well as ideal markets?

And if not, what are the policy implications? How might we use ABM (agent-based modeling) to help us formulate policy? What other kinds of investigations would be useful?

21.13 *GAS*, Chapter V: Disease Processes

Something of a *tour de force* for ABM

- Immune systems as binary strings
- Diseases as (shorter) binary strings
- Infection if no match, metabolism ups, agent changes to make a match
- Note genotype/phenotype distinction
- Interesting and plausible dynamics

See Annimations V-1 and -2.

21.14 *GAS*, Chapter V: Disease Processes

- Note genotype/phenotype distinction
- The mechanism isn't very realistic. Does this matter?
- How might we use a model of this sort for policy making? (Smallpox and bioterrorism?)
- How might we build and apply a model of this sort in other areas? (Viral marketing?)

21.15 *GAS*, Chapter VI: Conclusions

> The main point of the preceding chapters is simply this: *A wide range of important social, or collective, phenomena can be made to emerge from the spatio-temporal interaction of autonomous agents operating on landscapes under simple local rules.* [page 153]

- Summary beginning on page 153. Turning it all on, pages 154–9.
- "At the very least, artificial societies raise important policy questions. And they may help answer some of them." [page 160]. Discuss.
- Schelling's models, treated here. See Animations VI-4ff.

Chapter 22

(T:11/19) Dynamic Programming

22.1 Administration

Today: (1) discussion of zero-intelligence traders; (2) "Identity-Oriented Agents" chapter (just notes) in the AGE book; (3) "Dynamic Programming" chapter in the AGE book.

Thursday: (1) discussion of Hoffmann's paper (assignment due); (2) "Reinforcement Learning" chapter in the AGEbook. Recommended reading: Kaelbling, Littman, Moore, "Reinforcement Learning: A Survey" *Journal of Artificial Intelligence Research,* vol. 4, 1996, pp. 237-285. Download at `http://opim-sun.wharton.upenn.edu/~sok/papers/`

And (11/26): Beer game paper at:
`http://opim-sun.wharton.upenn.edu/~sok/–`
`papers/fmec-dss-si/kwz/beer-game-dss-final.doc`

Slides at
`http://opim-sun.wharton.upenn.edu/~sok/–`
`sokpapers/2000-1/beergameslides.ppt`

22.2 Zero-Intelligence Traders

"Allocative Efficiency of Markets with Zero-Intelligence Traders: Market as a Partial Substitute for Individual Rationality," Dhanajay K. Gode and Shyam Sunder, *Journal of Political Economy,* vol 101, no. 1, 1993, pp. 119–137.

- Three kinds of traders: ZI-U, ZI-C, human. Each given redemption value for amounts of goods to be purchased, and cost value for amounts of goods to be sold. Prices from 1 to 200 in increments of 1.

- ZI-C constrained to make random offers (bids) below redemption value, v_i, and asks above cost, c_i. Random within the region permitted by the constraints.

- All traders given (private) demand and supply functions.

- Double auction, both buys and sellers are permitted to buy and sell. Sellers because they are given goods at start.

22.3 Interpretation

- ZI-C vs ZI-U measures the "discipline of the market."

- Human vs ZI-C measures the value or effect of intelligence, added to the discipline of the market.

- Five experiments, characterized by 5 different supply and demand schedules.

22.4 Results

- ZI-U prices are totally random. ZI-C are noisy about the equilibrium price (Walrasian). Humans' prices are less noisy about the equilibrium.

- Humans show signs of learning; not so with ZI agents, of course.

- Efficiency (Is all the consumer and producer surplus extracted?). Table 2: ZI-C guys do well.

- Profit dispersion? Table 3. ZI-C generally higher than humans. "Paradoxically, profit maximization seems to be associated with lowering, not raising, profit dispersion across individuals."

22.5 Discussion

- Nice complement to Epstein and Axtell, chapter 3.
- Do you think the findings generalize? How might we use ABM to answer this question? Suggestion: efficiency, prices at the price equilibrium, etc., are the issues; study them from the bottom up.
- What if we introduced more intelligent agents into a population of ZI-C (or even ZI-U)? Is intelligence an arms race?
- Are there (likely) simple, robust rules for trading in such an environment? (Bid low, ask high, then back off in time?)
- What other properties are we interested in for such markets? What about dynamic changes in supply and demand schedules, what might happen then? What kinds of traders would adjust faster?

22.6 Identity-Centric Agents

- So far, mostly, agents have been naked strategies. Epstein and Axtell edge beyond this, yet their agents do very little learning; adaptation, yes, learning, not in any strong sense.
- The one learning regime we've discussed, the genetic algorithm, was (and normally is) presented as a way of evolving good solutions (naked strategies) from a population.
- How, then, are we to think about designing an identity-oriented agent? To begin, see the notes in the chapter "Identity-Oriented Agents" in the AGE book. . . .

22.7 Dynamic Programming

See the AGE book chapter (notes). DP is on the way to a general view of reinforcement learning, learning from feedback.

Chapter 23

(R:11/21) Reinforcement Learning

23.1 Administration

Today: (1) discussion Hoffmann's paper and his model; (2) Reinforcement learning, chapters from AGEbook (handed out).

Recommended reading: Kaelbling, Littman, Moore, "Reinforcement Learning: A Survey" *Journal of Artificial Intelligence Research,* vol. 4, 1996, pp. 237-285. Download at `http://opim-sun.wharton.upenn.edu/~sok/papers/`

And (11/26): Beer game paper at:
`http://opim-sun.wharton.upenn.edu/~sok/-`
`papers/fmec-dss-si/kwz/beer-game-dss-final.doc`

Slides at
`http://opim-sun.wharton.upenn.edu/~sok/-`
`sokpapers/2000-1/beergameslides.ppt`

23.2 Comment on IPD as a dynamic program

How would we formulate IPD as a dynamic program?

Well, each epoch/stage has just one state, s_t. At each s_t the player has decision, either C or D. The (one-stage) reward at t is determined by p_t, the probability that the counter-player plays C. After obtaining the reward, the player goes to state/stage s_{t+1} and things are exactly as before.

Suppose the game is finitely iterated? If p_t is unaffected by anything our player does, then clearly, no matter what p_t is it is preferable (for the player) to always choose D (defect) in IPD. Isn't this also true if it's an infinitely-repeated IPD?

In any event, the assumption that p_t is unaffected by anything our player does is crucial and highly problematic. If our player can induce the counter-player to play more cooperatively, both can gain. Perhaps we ought to think of Tit for Tat as a policy designed to reinforce the counter-player for good behavior.

Chapter 24

(T:11/26) Reinforcement Learning 2

Today, discussion of the Beer Game paper. Then chapter 26, Reinforcement Learning 2 in the AGEbook. Then Thanksgiving on Thursday.

Chapter 25

(T:12/3) LCS: Learning Classifier Systems

Chapter 23 of the AGEbook.

Chapter 26

(R:12/5) Last Class

Will email you all comments on your assignment #2 by Monday. Also, the final quiz. All in PDF. Please write down your emails for me now. Send me email if you don't hear from me.

Final projects and quiz due by 5 p.m. Friday 12/13/02. Email is OK, so is slipping the documents under my door (565 JMHH).

I'll have office hours next week and I invite you to come by (or make an appointment) and talk about your final projects. Friendly chats.

The final versions of the AGEbook and the Class Notes are now posted at the course Web site:
`http://opim-sun.wharton.upenn.edu/~sok/teaching/age/f02/`.

Today: VII Upwards and Onwards in the AGEbook. Mostly, chapter 28, Note: Rationality Redux. If time, chapter 29, Notes: Dual Interpretations.