On Complexity as Bounded Rationality

Yue & Mengrui
CS598HS

Recall Nash Equilibrium

A Nash Equilibrium is a set of strategies, one for each player, such that no player has incentive to change his or her strategy given what the other players are doing.

Our old friend: Prisoner's dilemma

	C	D
C	3,3	0,4
D	4,0	1,1

← **Nash Equilibrium**

Prisoner’s Dilemma contradicts social experience

“In real life we do not always behave in a selfishly antisocial way, and we often give up an advantage in order to behave in a cooperative manner.” (Papadimitriou & Yanakakis, 1994)

Real World Prisoner’s Dilemma Experiment

- Organizers: Menusch Khadjavi and Andreas Lange (Khadjavi & Lange, 2012)
- Participants:
  - female university Students
  - prisoners in Lower Saxony's primary women's prison
- Games Types:
  - simultaneous
  - sequential
- Payoff:
  - euros for students
  - coffee or cigarettes for prisoners
- Any Guess on the results?

Surprising Results on the Experiment

Collaboration Rate:

Group treatment	Students	Prisoners
Simultaneous PD	36.97%	55.56%
Sequential PD	63.04%	46.30%

Collective cooperation rates (both player cooperate):

Group treatment	Students (%)	Prisoners (%)
Simultaneous PD	13.14	30.16
Sequential PD	39.08	27.32

- Students are less collaborative than prisoners in Simultaneoius PD
- Prisoners's collaboration rate remains about the same

N-round Prisoner's Dilemma

- "In real life we do not play one-shot games; instead, we play **repeated** games. " (Papadimitriou & Yanakakis, 1994)
- "Unfortunately the only equilibrium is (D^n, D^n); that is, the only rational behavior is to defect all the time! "
- Proof: backwards induction
- Can we avoid the (D^n,D^n) equilibrium by **limiting strategic complexity**?
  - "Perhaps in a more realistic setting, in which the players can employ strategies that are in some sense simple, collaborative behavior is not ruled out."

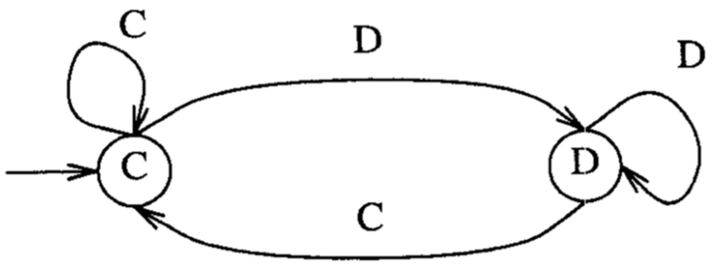
Bounded rationality

- Bounded rationality is the idea that we make decisions that are rational, but within the **limits** of the information available to us and our mental capabilities
- Rational agents will behave in complex and devious ways in order to extract a little more payoff. Such behaviors are ruled out if one assumes that agents **do not invest** inordinate amounts of computational resources and reasoning power to achieve small payoffs.
- For Prisoner's Dilemma, limiting the **size of states or memory** foster collaboration.

Axelrod’s Tournament

- In 1980, Robert Axelrod, professor of political science at the University of Michigan, held a tournament of various strategies for the prisoner's dilemma.
- 14 Strategies submitted:
  - All defect
  - Random
  - Unforgiving
- Winner: Tit-for-Tat (Anatol Rapoport)

Tit for Tat



- Two States:
  - Collaborate & defects
- Transition is labeled by the opponent’s strategy
- Begin by collaborating
- For all future round, copy opponent’s strategy from the previous round.

Simulation

https://repl.it/@MengruiLuo/SkyblueElatedMalware-1

Tit-for-Tat vs. Tit-for-Tat

Round 1	Round 2	Round 3	Payoff
C	C	C	3
C	C	C	3

Total Rational vs. Total Rational

Round 1	Round 2	Round 3	Payoff
D	D	D	1
D	D	D	1

Author: Papadimitriou

From Greece, (1949 ~), Currently working at UCB;

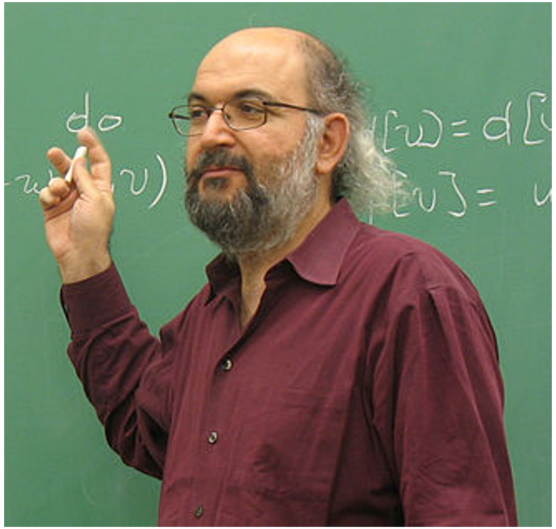
Won Donald E. Knuth Prize for his contribution in Complexity Theory;

Proved that solving for a Nash equilibria can be intractable. (PPAD-class, polynomial parity arguments on directed graphs)

Author: Yanakakis

From Greece, (1953~), currently working at Columbia University;

Won Donald E. Knuth Prize for theoretical Computer Science.



Papadimitriou's assumption

Agents who have a lot of computing resources (memories or states) at hand might use them to increase their utility,

which means, they might use their resources to come up with a devious strategy.

Theorem 1

	C	D
C	3,3	0,4
D	4,0	1,1

For every $\epsilon > 0$ in the n-round prisoner's dilemma played by two automata where at least one of them has a subexponential number of states, there is a mixed equilibrium with an average payoff of at least $(3 - \epsilon)$

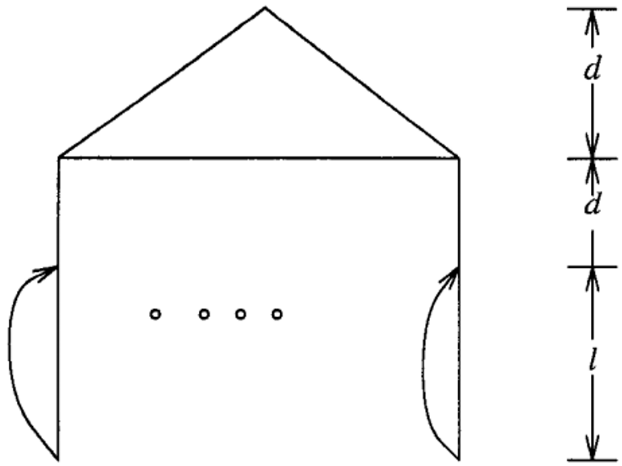
Proof: Intuition

1. Force the other player to memorize something at the start to fill up memory they might use to be devious otherwise.

2. Cooperate for a period of time and then prove to each other that you memorized what you were supposed to.

3. Punish any deviation by always defecting.

Sketch Proof: Business Card Game



$$d2^d2(1 + \frac{1}{\epsilon}) \geq s.$$

$$l \geq \lfloor d(1 + \frac{2}{\epsilon}) \rfloor$$

Random -> Exchange -> Loop

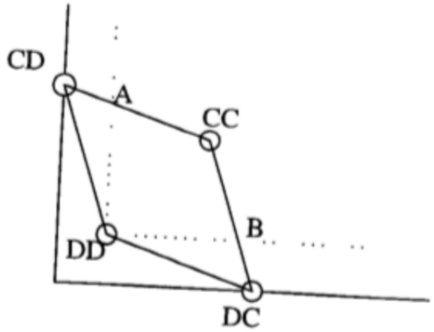
Comments on Theorem 1

1. If a young, simple and naive guy is playing the Prisoner's Dilemma game with his experienced 'Mind-reader' friend, they can end up their game with some collaborations.

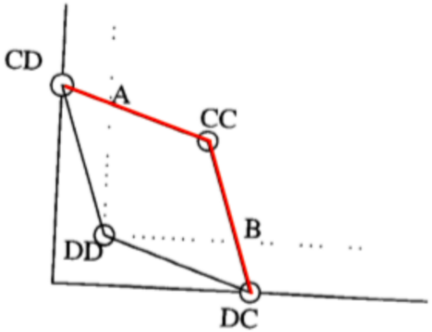
2. However, the Theorem 1 is kind of 'fragile'. Higher payoff can be achieved with a larger automata.
3. The theorem is still restricted within the PD game and sometimes only one person's rationality is limited, what if we generalize it into larger case and closer to life?

Introduction: Pareto's Region

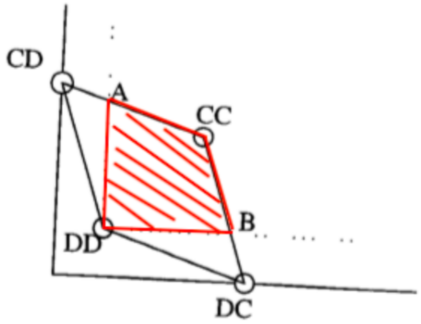
Payoff Geometry



Payoff Geometry: Pareto

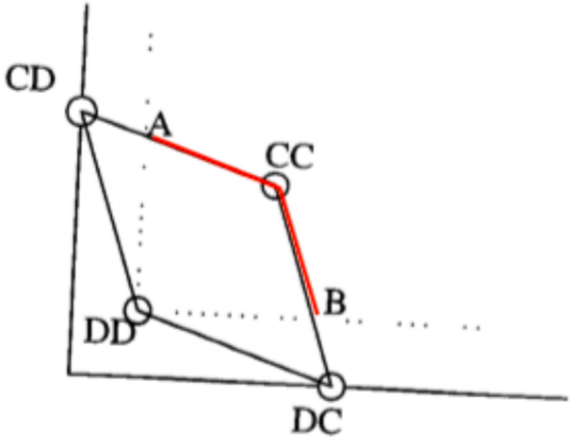


Payoff Geometry: IRR



Introduction: Pareto's Region (Cont.)

Pareto-IRR



Theorem 2

For an arbitrary game G if $p = (p_1, p_2)$ is an individually rational Pareto optimal point, then for every ϵ, there is a subexponential bound on automata size such that an automata smaller than the bound exists for both players corresponding to a mixed equilibria with average payoff at least $p_i - \epsilon$ for each player in the n repeated game of G

Comment on Theorem 2

1. It is a more general case, and Pareto-IRR is introduced into this theorem, which is more accurate and elegant;
2. Again, it proves the figurative example we use in the last example is true.

Reflections on Previous Part

Tit-for-Tat vs. Tit-for-Tat

Round 1	Round 2	Round 3	Payoff
C	C	C	3
C	C	C	3

At least 3-epsilon, with 3 being pi

Total Rational vs. Total Rational

Round 1	Round 2	Round 3	Payoff
D	D	D	1
D	D	D	1

Think about epsilon and d in the business card game?

$$d2^d2(1 + \frac{1}{\epsilon}) \geq s.$$

Complexity Theory of Games

Suppose there is a game:

A game scheme *g* is a polynomially computable function from 3 strings to 2 integers g(z,x,y) = (a,b)

z encodes the game, x player 1's strategy, y player 2's -> a is player 1's payoff, b is player 2's

Complexity Theory of Games (Cont.)

Decision problem	Equals complexity class
There exists a strategy y, which given x and z has a payoff at least b	NP
There exists an equilibrium which pays player 1 at least b for the zero-sum game z	EXP
There exists an equilibrium in game z which pays both player 1 and player 2 at least b	NEXP

Related Work - Complexity

Papadimitriou’s landmark paper

Papadimitriou first proves that for N players (N>=4), the complexity class of solving Nash Equilibrium is PPAD-Complete. (PPAD-class, polynomial parityarguments on directed graphs)

Chen and Deng proves that this theorem also applies to N=2.

Related Work - Machine Learning

Rubinstein(1993), Cho and Libgober(2019)

One seller, infinite buyers of two groups (high cost and low cost);

Asymmetric information; (Seller has more knowledge about good’s value)

Seller tends to exclude high cost buyers, and he use arbitrary pricing strategy;

Conclusion 1: If buyers are totally rational, the reasonable decision is to compare the price with the expected value of posterior distribution after Beyasian updating.

Buyers decision function of time may not be monotonic.

Related work - Machine Learning (Cont.)

Rubinstein further proposes:

For high cost buyers, they can only take ‘Threshold Strategy’, which means they can only decide based on the threshold value.

Conclusion 2: Sellers can approach maximum value of profit as close as possible, which means high cost buyers are almostly all excluded.

Related Work - Machine Learning (Cont.)

What if we view this problem from a machine learning perspective?

AdaBoosting (Adaptive Boosting): by increasing number of ‘weak learners’ (Binary learner in this case) and giving weights, an accurate classification can be reached.

Economically speaking, AdaBoosting reflects whether given infinite bounded rational individuals can make a rational decision by adjusting their weights.

Related Work - Machine Learning (Cont.)

To evaluate AdaBoosting, Uniformly PAC Learnability is introduced (Probably Approximate Correct).

Conclusion:

If seller can choose arbitrary pricing strategy and publicize it after choosing, AdaBoosting is PAC learnable;

If seller can choose arbitrary pricing strategy but not publicize it after choosing, AdaBoosting is not PAC learnable.

Some strategies cheat clever ML algorithms! And what are they?

Related Work - Machine Learning (Cont.)

Conclusion:

If seller only choose the best response pricing strategy against AdaBoosting, even seller does not publicize his strategy, AdaBoosting can also recognize it.

Which means,

If seller is being rational, and try to cheat buyers, exclude high cost buyers and maximize his profit, AdaBoosting can read his mind, then he will suffer profit loss; but if not being rational, he can cheat AdaBoosting.

Connected!

Related Works and Reflection

Bounded rationality is also commonly seen in Evolutionary Dynamics, Biology and so on...

Very indiscreet and personal thought:

Such a simple set model can provoke great interdisciplinary ideas, and sometimes there is an analogy between one conclusion and the other.

Reference

Cho, I. & Libgober, J. (2019). Machine Learning for Strategic Inference in a Simple Dynamic Game. Working Paper.

KAZNATCHEEV, ARTEM. “Short History of Iterated Prisoner's Dilemma Tournaments.” Theory, Evolution, and Games Group, 3 Mar. 2015, egtheory.wordpress.com/2015/03/02/ipd/.

Khadjavi, Menusch, and Andreas Lange. “Prisoners and Their Dilemma.” Journal of Economic Behavior & Organization, North-Holland, 12 June 2013, www.sciencedirect.com/science/article/pii/S0167268113001522.

Rubinstein, A. (1993). On price recognition and computational complexity in a monopolistic model.Journal of Political Economy,101(3), 473-484.