Computer modeling of physical phenomena

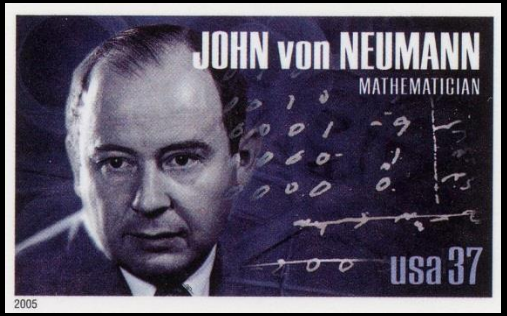
05-06.03.2024

Lecture II: Game theory

Game theory

Game theory studies the ways in which strategic interactions among rational players produce outcomes with respect to the players' preferences (the outcomes might not have been intended by any of them).

Modern game theory

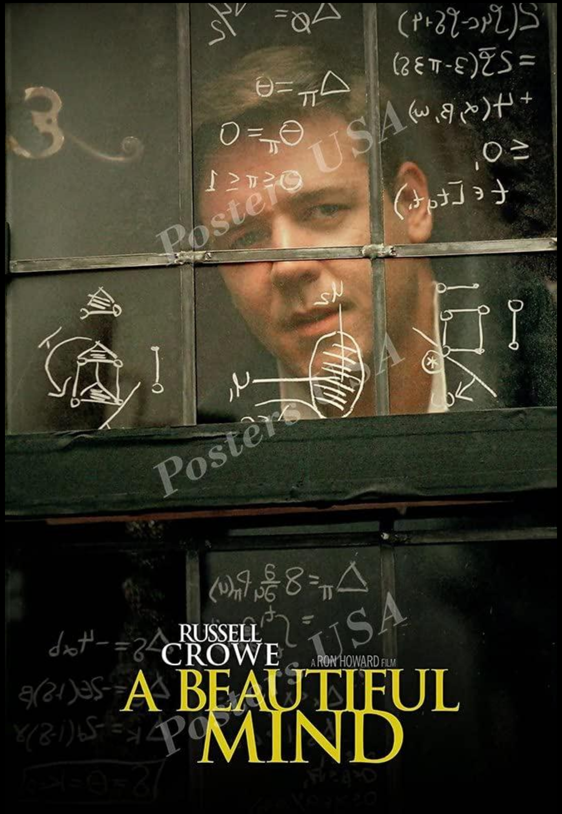


Neumann, J., and Morgenstern, O., (1947). The Theory of Games and Economic Behavior.

Nash, J., (1950). Equilibrium Points in n-Person Games, PNAS, 36, 48-49.

Nash, J., (1950). Non-Cooperative Games, PhD Thesis, Princeton.

John and Alicia Nash



A Beautiful Mind (2001)

Rules, Strategies, Payoffs, and Equilibrium

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

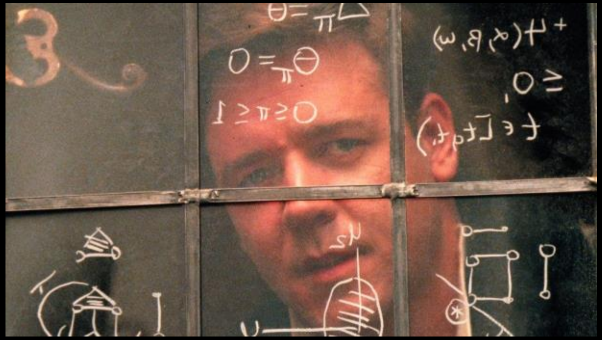
Zero-sum or not

Zero-sum describes a situation in which a participant's gain (or loss) is exactly balanced by the losses (or gains) of the other participant(s).

Non-zero-sum games are more complex to analyze and more interesting!

Nash Equilibrium

A situation when each player's strategy is optimal, given the strategies of the other players (no player can benefit from unilaterally changing his strategy, while all other players stay fixed).



Every finite game has at least one Nash equilibrium.

Cake Dilemma

How to split a cake between two bratty children so that none of them can complain they were slighted with the smaller piece?

The payoff matrix

		CHOOSER	
		Choose bigger	**Choose smaller**
SLICER	**Slice equally**	Equal minus ε for S Equal plus ε for C	Equal plus ε for S Equal minus ε for C
	Slice unequally	Smaller for S Bigger for C	Bigger for S Smaller for C

Prisoner’s Dilemma

Merrill Flood and Melvin Dresher (RAND, 1950)

Intro: Buick sale

In June 1949, Flood wanted to buy a used Buick from a RAND employee who was moving back East. Buyer and seller were friends. They weren't looking to cheat each other, just to agree on a fair price for the car. How should they set a price?

Dealer: buy for 500$, sell for 800$, should they split an extra $300 profit?

What if one defects (I will sell for 750$!) - the other is still better off to accept it than to go the dealer.

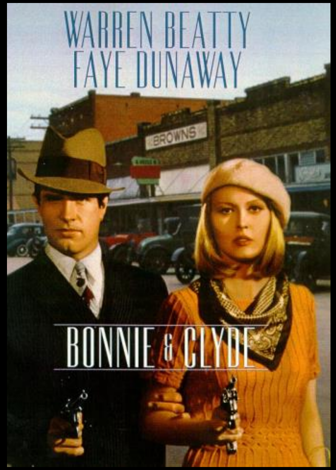
The party who is more unreasonable is apt to get the better of the deal!

based on: William Poundstone, *Prisoner's Dilemma*, 1992

Prisoner’s dilemma

Bonnie and Clyde are arrested by the police and charged with various crimes. They are questioned in separate cells, unable to communicate with each other. They know how it works:

- if both resist interrogation (cooperating with each other) and proclaim mutual innocence, they will get a 2-year sentence for robbery,
- if one confesses (defecting) and the other doesn’t (cooperating), the confesser goes free and the other will get a severe 8-year sentence,
- if they both confess (defecting), then the judge will sentence both to moderate 4 years in prison.

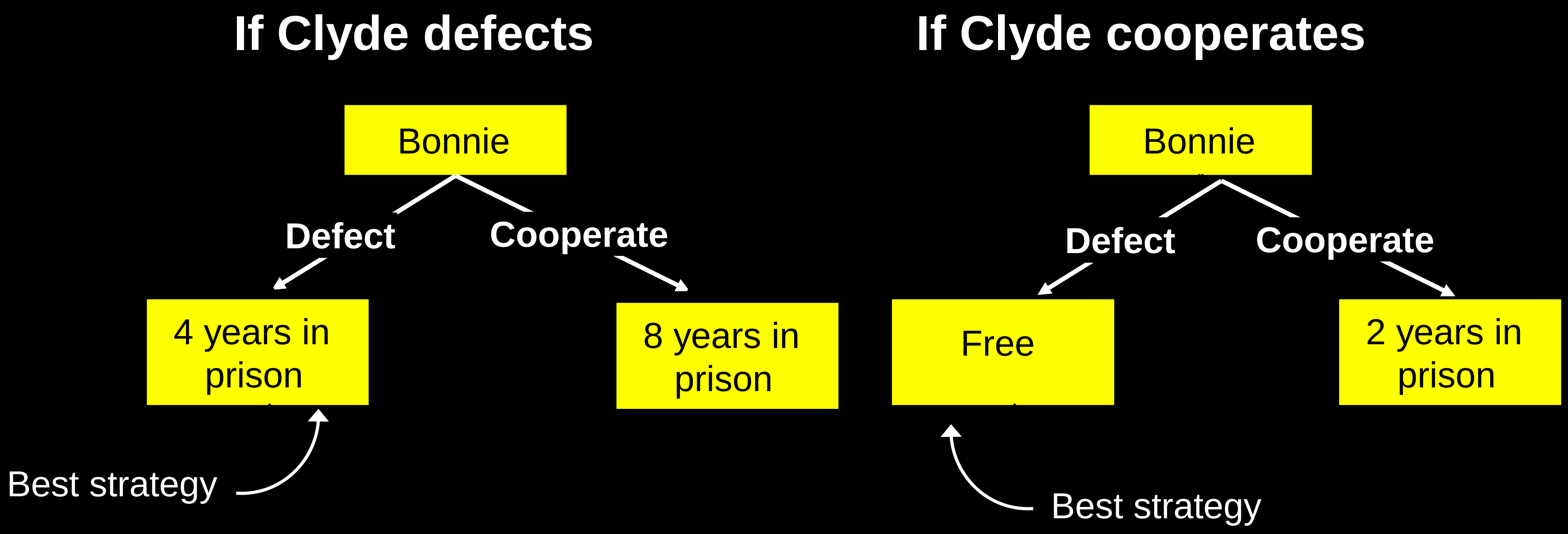


What should Bonnie do? What should Clyde do?

The payoff matrix

		CLYDE: Cooperate	CLYDE: Defect
BONNIE	Cooperate	2 years for B 2 years for C	8 years for B 0 years for C
BONNIE	Defect	0 years for B 8 years for C	4 years for B 4 year for C

Bonnie’s Decision Tree



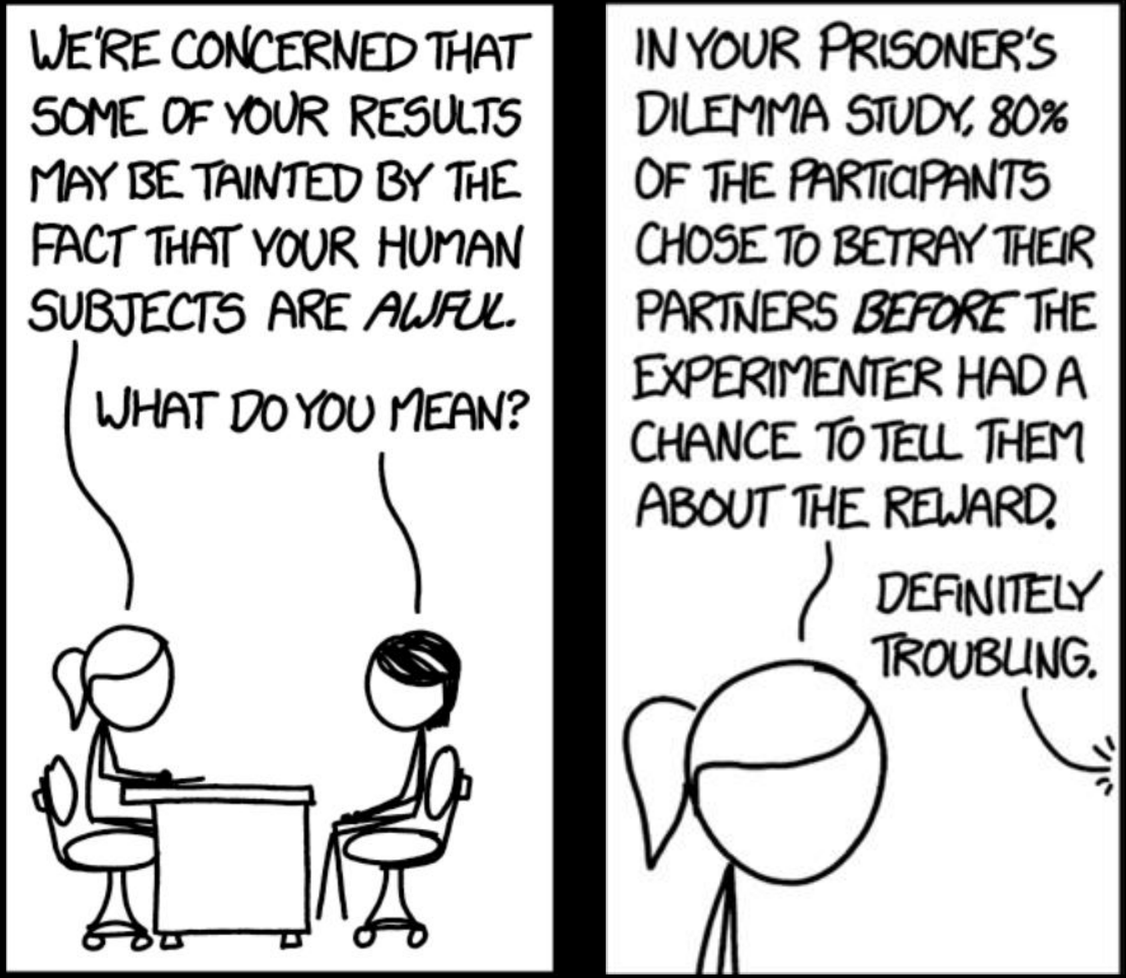
The dominant strategy for Bonnie is to confess (defect) because no matter what Clyde does, she is better off confessing.

The payoff matrix

		CLYDE	
		Cooperate	**Defect**
BONNIE	**Cooperate**	2 years for B 2 years for C	8 years for B 0 years for C
	Defect	0 years for B 8 years for C	4 years for B 4 year for C

Nash equilibrium

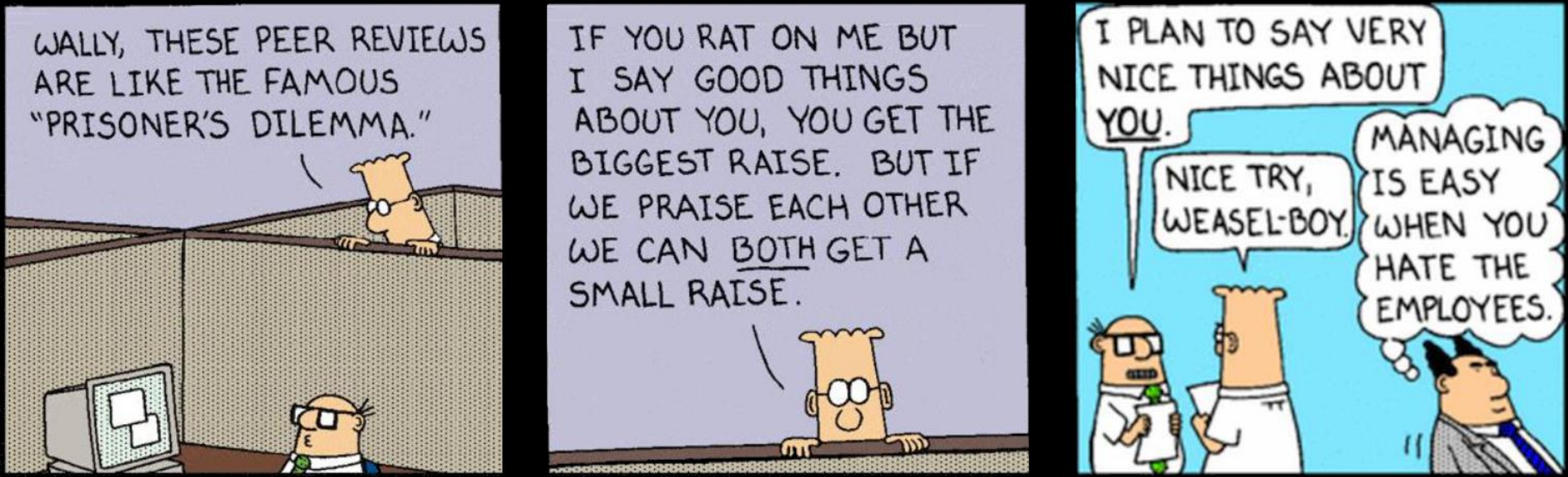
It seems we should always defect and never cooperate...



Examples...

- Cheating on a cartel:
  - Cartel members can charge the monopoly price or a lower price.
  - The best strategy is to charge the low price.
- Trade wars between countries
  - Free trade benefits both trading countries.
  - Tariffs can benefit one trading country.
  - Imposing tariffs can be a dominant strategy and establish a Nash equilibrium even though it may be inefficient.
- Advertising
  - All firms advertising tends to equalize the effects.
  - Everyone would gain if no one advertised.

Examples...

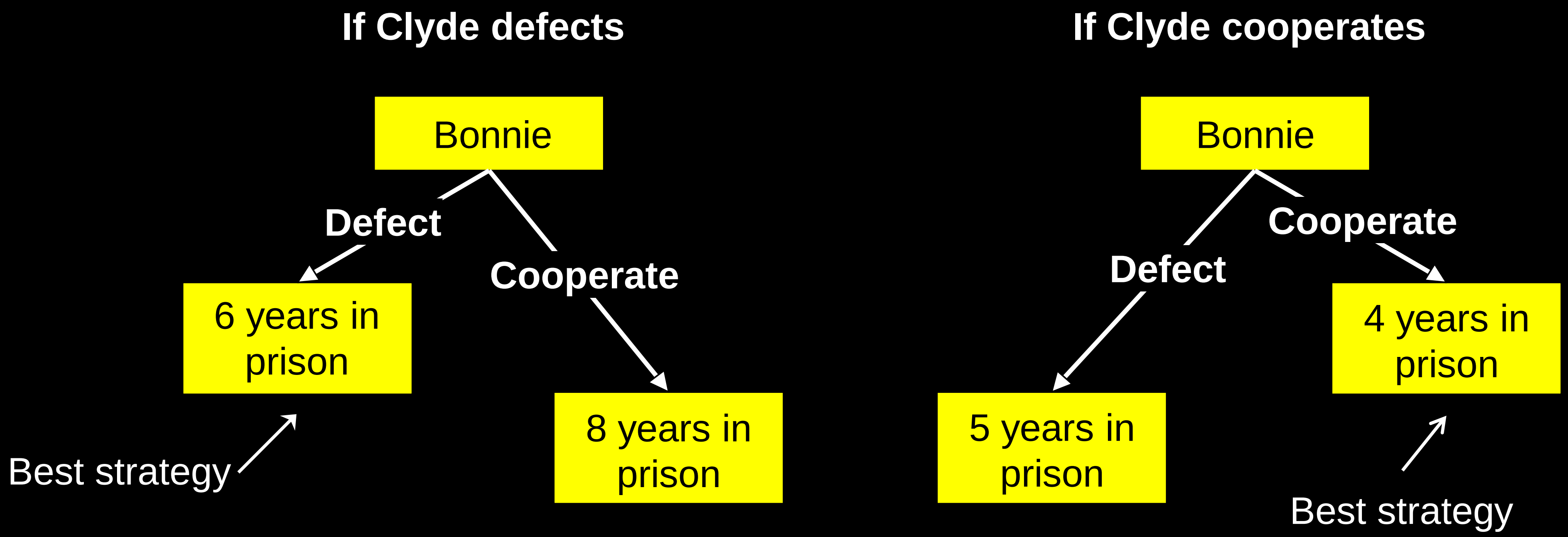

WALLY, THESE PEER REVIEWS ARE LIKE THE FAMOUS "PRISONER'S DILEMMA."
IF YOU RAT ON ME BUT I SAY GOOD THINGS ABOUT YOU, YOU GET THE BIGGEST RAISE. BUT IF WE PRAISE EACH OTHER WE CAN BOTH GET A SMALL RAISE.
I PLAN TO SAY VERY NICE THINGS ABOUT YOU.
NICE TRY, WEASEL-BOY.
MANAGING IS EASY WHEN YOU HATE THE EMPLOYEES.

Games without dominant strategies

In many games the players have no dominant strategy – often a player's strategy depends on the strategies of others.

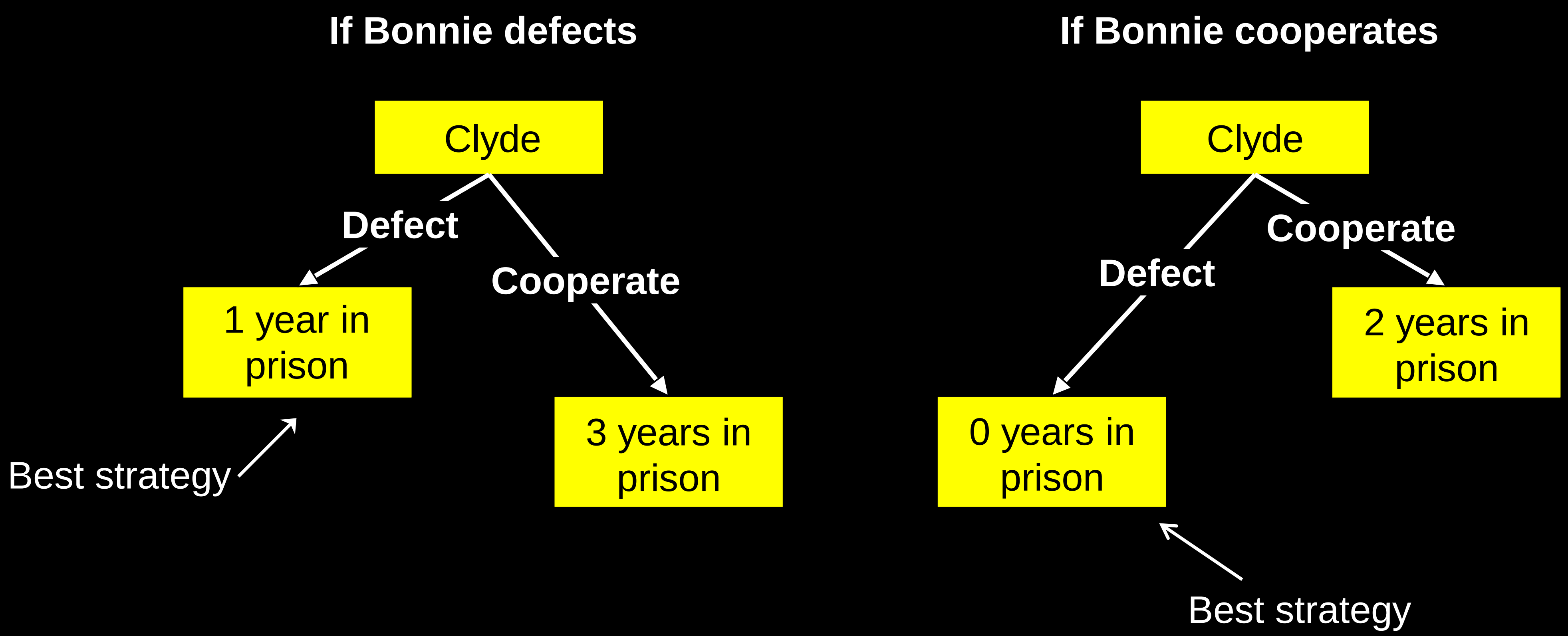
		CLYDE	
		Cooperate	**Defect**
BONNIE	**Cooperate**	4 years for B 2 years for C	8 years for B 0 years for C
	Defect	5 years for B 3 years for C	6 years for B 1 year for C

Bonnie's Decision Tree



Bonnie has no explicit dominant strategy, but there is an *implicit* one!

Clyde's Decision Tree



Clyde has a dominant strategy (confess); thus, Bonnie could assume that he would choose it and choose to confess as well.

Some games have no simple solution

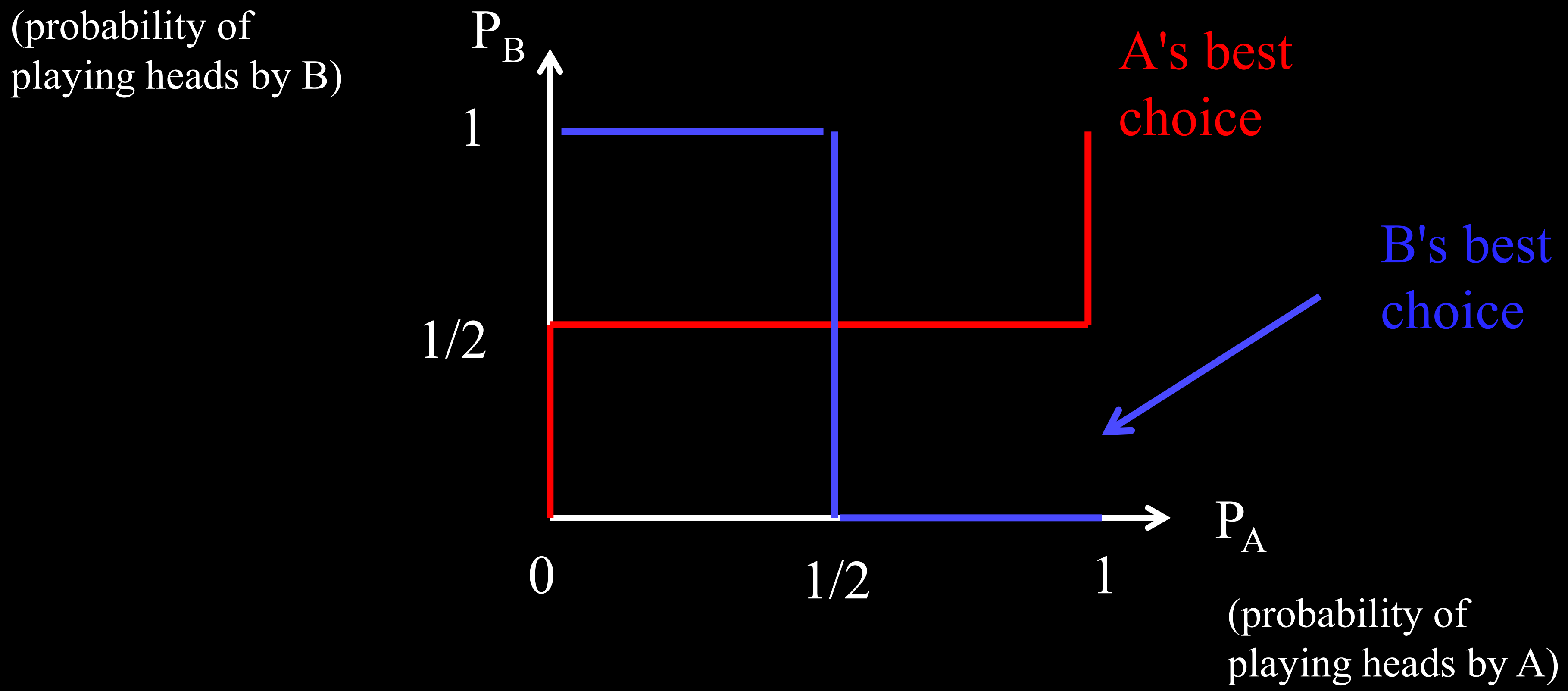
In *matching pennies*, neither player has a dominant strategy:

		Player B: Heads	Player B: Tails
Player A	Heads	1, -1	-1, 1
	Tails	-1, 1	1, -1

This game has no pure strategy Nash equilibrium since there is no pure strategy (heads or tails) that is a best response to a best response.

Mixed (non-deterministic) strategy

How should the players play, possibly in a randomized way, to maximize their expected reward?



Randomized Nash equilibrium!

Randomized Nash equlibrium

These best response functions have only one intersection: (0.5, 0.5). Thus there is only one pair of choices for (P_A, P_B) where no player has an incentive to deviate unilaterally. If, e.g. (P_A, P_B) = (0.4, 0.5), then B has an incentive to deviate from 0.5 because changing PB to 1 will increase his chance of winning. Only the intersection of the two best response functions is such that no player wants to deviate unilaterally: this is Nash Equilibrium. Thus, the game of Matching Pennies has a unique Nash Equilibrium, and it is randomized.

Payoff matrix for the generic symmetric two-person dilemma game

(A’s payoff, B’s payoff)

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

M. J. Guyer and A. Rapoport (1966)

Payoffs

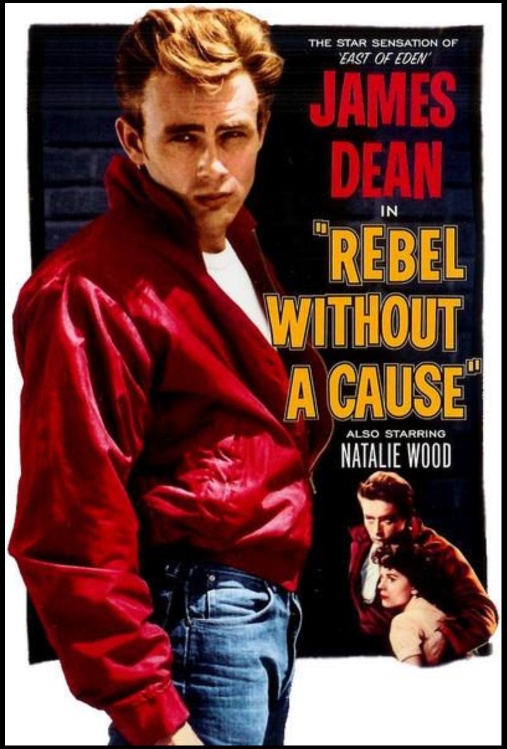
Four payoffs involved:

- CC: both players cooperate,
- CD: you cooperate but other defects (aka "sucker's payoff"),
- DC: you defect and other cooperates (aka "temptation to defect"),
- DD: both players defect.

Some classical dilemma games:

- ➢ Prisoner's dilemma: DC > CC > DD > CD
- ➢ Chicken: DC > CC > CD > DD
- ➢ Stag Hunt / Trust dilemma: CC > DC > DD > CD

Chicken



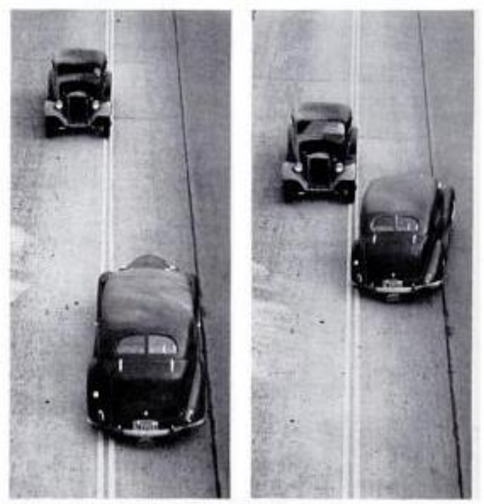
Rebel Without a Cause (1955)

Chicken

DC > CC > CD > DD

Cooperation: swerving

Defecting: not swerving



- The worst thing that can happen is for both players *not* to swerve.
- The best thing that can happen, is to show your machismo by not swerving and letting the other driver swerve.
- Being chicken is the next to worst outcome, but still better than dying.

There is a cooperative outcome in chicken. It's not so bad if both players swerve. Both come out alive, and no one can call the other a chicken.

Example payoff matrix

		Player B	
		swerve	drive straight
Player A	swerve	2, 2	1, 3
	drive straight	3, 1	0, 0

- No dominant strategy: the best is to do the opposite to what the other player does.
- Playing safe? Swerving has the maximum minimum.
- Chicken vs the prisoner's dilemma: mutual defection (the crash when both players drive straight) is the most feared outcome in chicken.

Trust Dilemma

CC > DC > DD > CD

		Player B: haircut	Player B: no haircut
Player A	haircut	5, 5	0, 4
	no haircut	4, 0	2, 2

- The best possible outcome would be for both of you to get the haircut.
- It would be awful to be the *only* one to show up with the haircut.
- If you *didn't* get the haircut, but the friend did, and looked like a real jerk, that would almost be as good as if both of you got the haircut.
- It wouldn't really be so bad if *no one* got the haircut.

Trust Dilemma

CC > DC > DD > CD

		Player B: haircut	Player B: no haircut
Player A	haircut	5, 5	0, 4
Player A	no haircut	4, 0	2, 2

risk dominant strategy - no haircut to minimize the risk

payoff dominant strategy - haircut to get the maximum reward

The more uncertainty players have about the actions of the other players, the more likely they will choose the risk dominant strategy.

More examples of the GT in real life

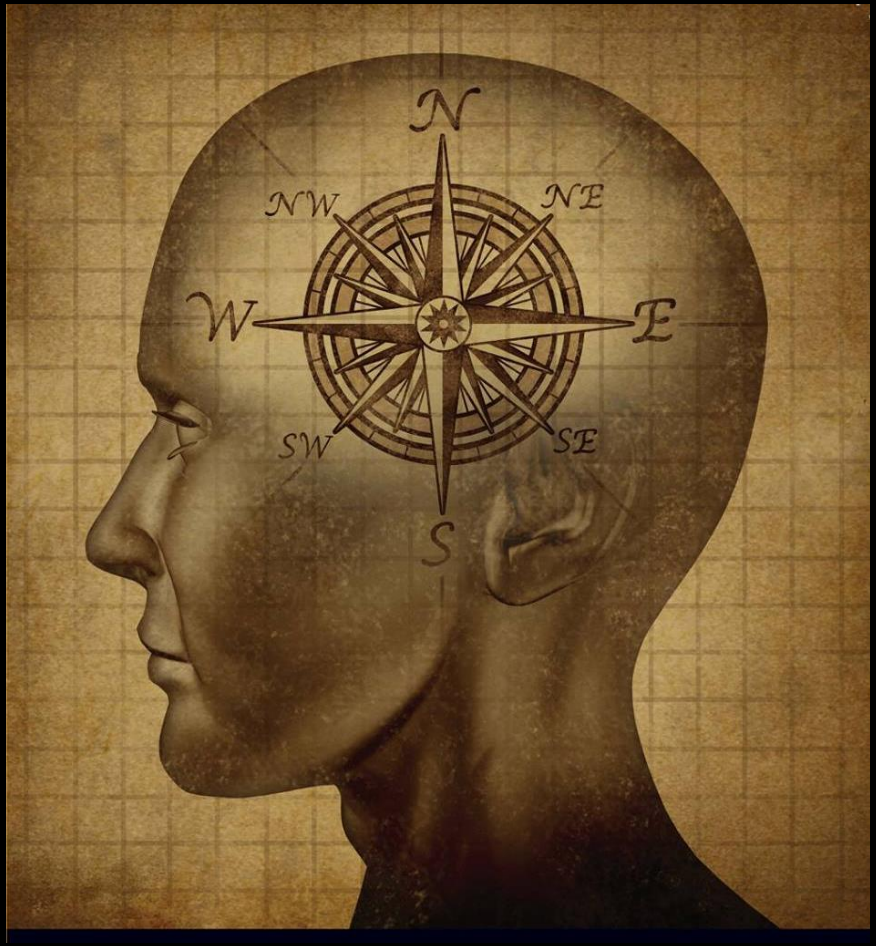
Communal coffeepot

- Cooperate by making a new pot of coffee if you take the last cup.
- Defect by taking the last cup and not making a new pot, depending on the next coffee seeker to do it.
- DC > CC > DD > CD

Class team project

- Cooperate by doing your part well and on time.
- Defect by slacking, hoping the other team members will come through and sharing the benefit of a good grade.
- DC > CC > DD > CD

How to save morality?



Iterated Prisoner’s Dilemma

Game theory shows that a rational player should always defect when engaged in a prisoner’s dilemma situation. We know that in real situations, people don’t always do this.

Why not? Possible explanations:

- people aren’t rational,
- morality,
- social pressure,
- fear of consequences,
- evolution of species-favoring genes.

In frames of game theory, some of these ideas can be explored by playing the iterated version of the games.

Iterated Prisoner's Dilemma

Key idea: in many situations, we play more than one "game" with a given player.

Players have complete knowledge of the past games, including their choices and the other player's choices. Your choice in future games when playing against a given player can be partially based on whether he has been cooperative in the past.

Axelrod & Hamilton tournament

Round-robin computer tournament between strategies submitted by colleagues.

Each game between competing strategies consisted of 200 moves, yielding possible scores from 0 to 1000, so that a "good performance" score was judged to be 600 points (complete mutual cooperation) and a "poor performance" score would be 200 points (complete mutual defection).

The overall winner with an average of 504 points per game, submitted by game theorist Anatol Rapoport, was so called Tit-for-Tat: a player cooperates on the first move and then does exactly what the other player did on the previous move.

While never actually attaining a higher score in any particular game, this strategy can lead to many rounds of cooperation between opponents with its commensurate rewards, whereas the more defecting strategies do poorly.

Robert Axelrod and William D. Hamilton. "The evolution of cooperation" *Science* 211.4489 (1981): 1390-1396.

Axelrod on Tit-for-Tat

"What accounts for Tit-for-Tat's robust success is its combination of being nice [not being the first to defect], retaliatory, forgiving, and clear. Its niceness prevents it from getting into unnecessary trouble. Its retaliation discourages the other side from persisting whenever defection is tried. Its forgiveness helps restore mutual cooperation. And its clarity makes it intelligible to the other player, thereby eliciting long-term cooperation" (Axelrod [1984]).

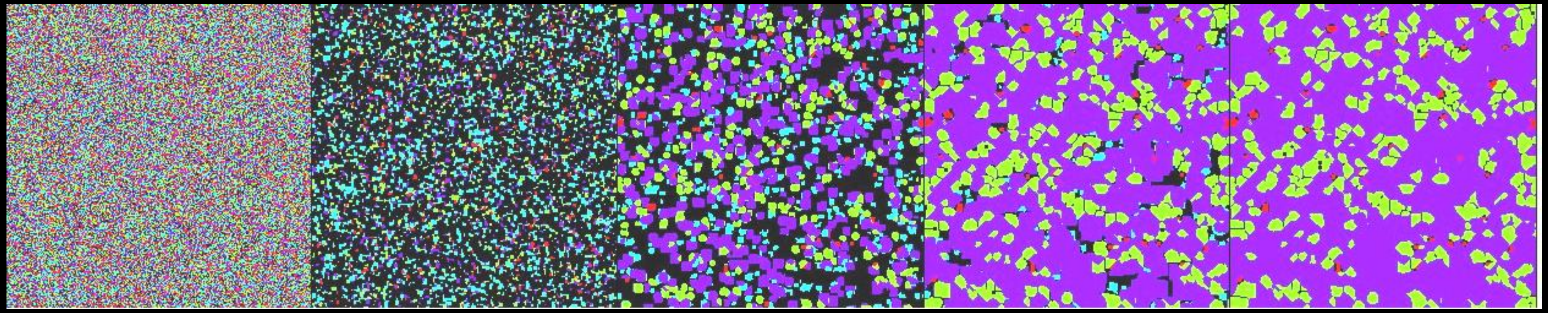
Spatial version of PD

Cooperation is easier to maintain in a sedentary population: defectors can thrive in an anonymous crowd, but mutual aid is frequent among neighbors.

As a model of the above, one can consider PD played on a lattice with either Moore or von Neumann neighbourhood.

- Various strategies are allowed to compete against each other.
- Initially, strategies are randomly assigned to the cells.
- Each cell then competes for a fixed number of rounds with each of its neighbors.
- At the end of each session a cell will adopt the strategy of its most successful neighbor in terms of highest total score.

Spatial version of PD



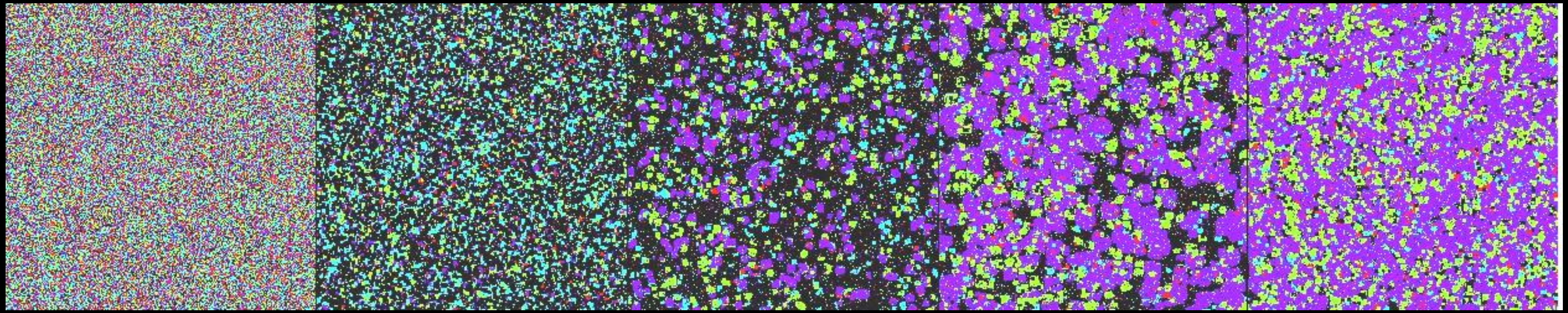
The evolution of the spatial iterated prisoner's dilemma with five competing strategies from an initial random configuration. Each cell's strategy plays off five times against its eight neighbors and then adopts the strategy of its most successful neighbor. Images are taken at the initial state, after 2, 5, 10, and 50 iterations.

- Always Cooperate
- Tit-for-Tat
- Pavlov
- Always Defect
- Random

Pavlov - repeats its former choice whenever it wins a round (C/C or D/C) and switches otherwise (if winning, hold fast; if not, change course).

TfT seems to win, but this depends on initial conditions!

The same with noise



A certain level of noise introduced, that is, a cell will make a random move in a single round, thus increasing the dynamics of the evolution.

Always Cooperate

Tit-for-Tat

Pavlov

Always Defect

Random

Unpredictable social behaviour?

Remember the Turing halting problem?

In SPD, will a single strategy prove triumphant in the sense of progressively conquering more and more territory without opposition, or will an equilibrium of some small number of strategies emerge?

Spatial PD systems are in general unpredictable...

P. Grim. "The undecidability of the spatialized prisoner's dilemma" *Theory and Decision* 42.1 (1997): 53-80.

Calvin and Hobbes by Bill Watterson