The Institute of Mathematical Sciences, Chennai

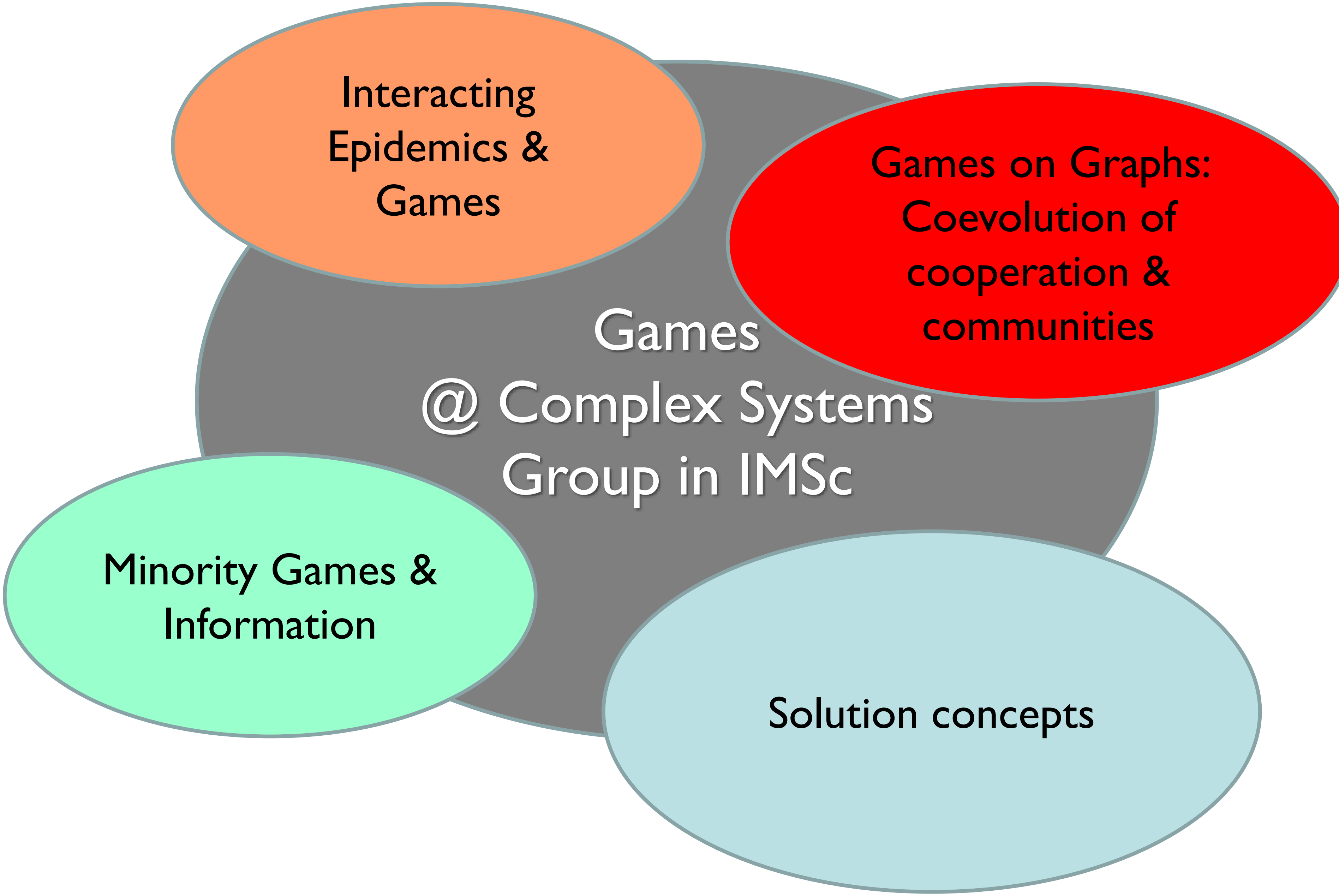
The temptation of Mr Spock

Solution frameworks for non-cooperative games among rational agents

Sitabhra Sinha
The Institute of Mathematical Sciences, Chennai

in collaboration with

V Sasidevan


Interacting Epidemics & Games
Games on Graphs: Coevolution of cooperation & communities
Games @ Complex Systems Group in IMSc
Minority Games & Information
Solution concepts

Why cooperate ?

Elizabeth Pennisi (2005):
"When Charles Darwin was working out his grand theory on the origin of species, he was perplexed by the fact that animals from ants to people form social groups in which most individuals work for the common good. This seemed to run counter to his proposal that individual fitness was key to surviving over the long term"

July 2005

J B S Haldane: "I will jump into the river to save two brothers or eight cousins" (precursor to Hamilton's Rule)

Natural selection may encourage altruistic behavior among kin as it improves the reproductive potential of the genetically related group as a whole, but unclear why unrelated individuals should help each other.

Why cooperate ?
How does cooperative behavior evolve ?

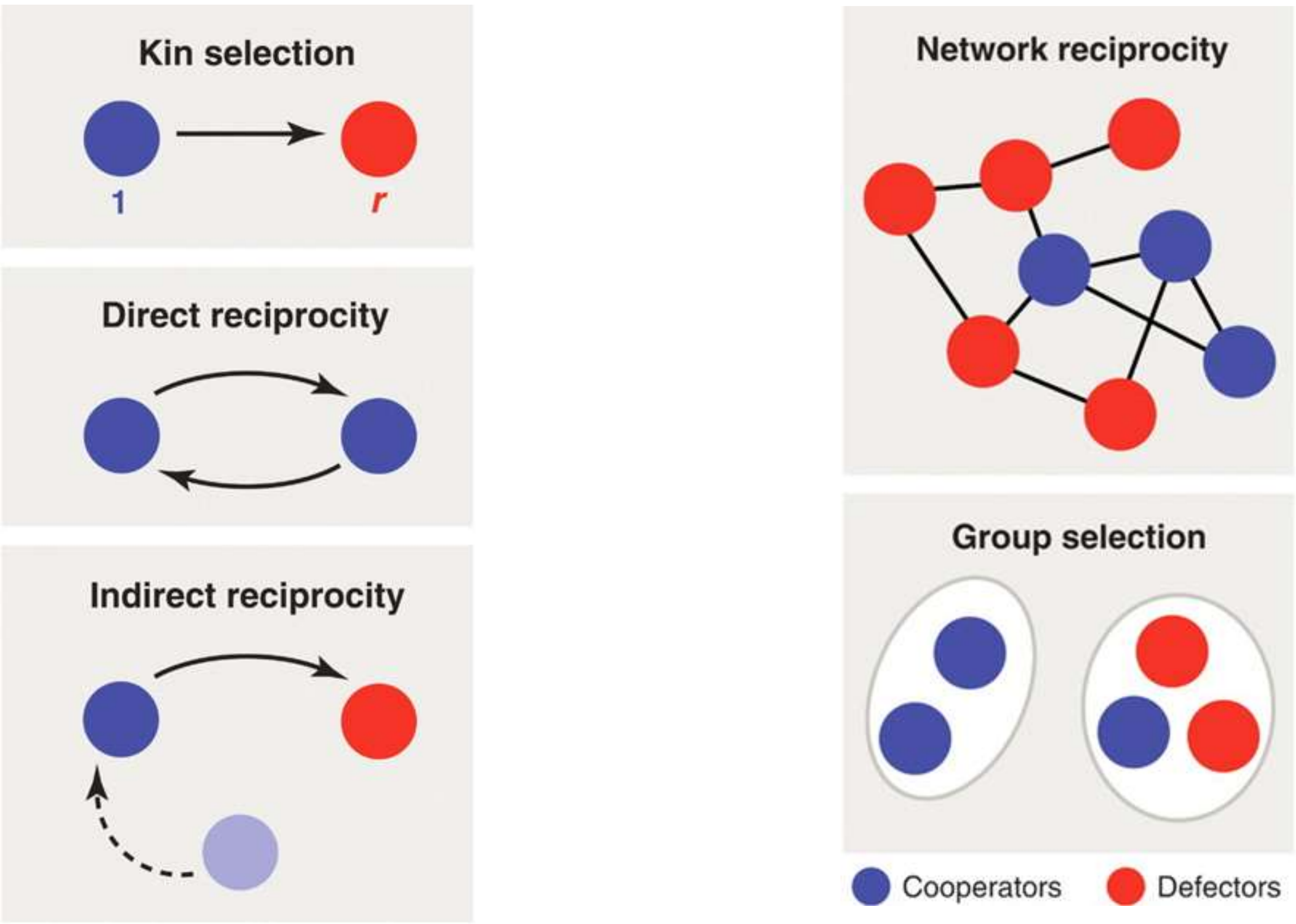
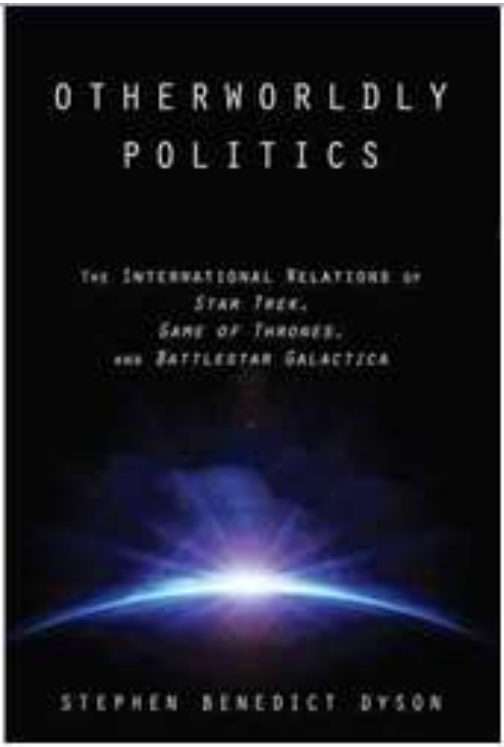


Image: Lisa Burnett, from
Martin A. Nowak, *Science* **314**, 1560-1563 (2006)

Should rational[(*)] agents cooperate ?

[(*)] agents are non-altruistic, only interested in maximizing their own benefit

We need a behavioral science for Vulcans (a.k.a. perfectly rational beings)



"the needs of the many outweigh the needs of the few, or the one"

Mr Spock
in *Star Trek II: The Wrath of Khan*

Or, does the greed of a few (or one) outweigh the needs of many ?

Paraphrasing a "Occupy Wall Street" poster

Enter the theory of "Games"

- *Games*: Strategic interactions between agents
- *Agents*: Variety of entities, ranging from human beings/ animals/ cells to organizations and nations, or even, computer programs.
- Each agent receives a *payoff* depending upon the strategy choice made by all agents including herself
- Agents aim to maximize their payoff by choosing optimal strategies, taking into account that other agents will also be doing the same

Early attempts at using the concept of games for analyzing strategic thinking: **Kriegsspiel**, a war-game used for training Prussian officers

Games: Equilibrium and Solution Concept

A key concept in the study of games is that of an

- ❑ Equilibrium: a state of affairs where each agent has decided her strategy for the game, which is arrived at by using a
- ❑ Solution concept: the process by which the agents pick their equilibrium strategy – i.e., a formal rule for predicting how a game will be played between agents

A particular solution concept employs certain assumptions regarding agent's behavior.

Standard game theory makes several assumptions about the agent's behavior:

- ❑ Agents are fully rational,
- ❑ Agents would like to optimize their payoff,
- ❑ Agents can perfectly execute their strategies

Nash Equilibrium

J. F. Nash, "Equilibrium Points in n-Person Games"
Proc. Natl. Acad. Sci. USA, 1950

An important solution concept for
Non-cooperative games in which players make decisions independently

While the actions of players may result in cooperation, it must be self-organized

Informally
Nash equilibrium is a state where after every agent has selected their optimal 'Nash' strategies, none of the agents can improve their payoff by **unilaterally** deviating from it.

A game can have multiple Nash equilibria – in which case one needs to employ additional refinement (selection) criteria to decide which equilibrium agents will choose

John F Nash
(1928-2015)

The Nash equilibrium of a game may sometimes be inferior to an alternate choice of actions by the agents in which all the parties get higher payoff … gives rise to "social dilemmas" such as

Prisoners Dilemma

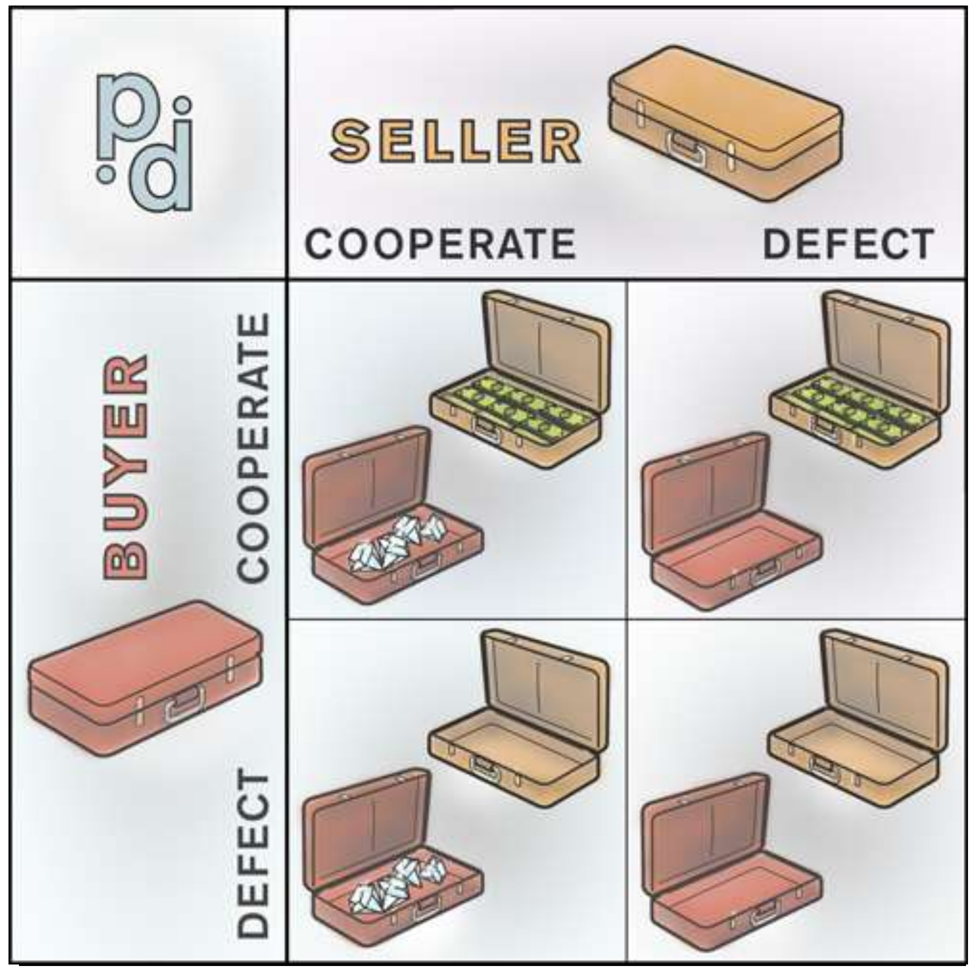
originally framed by Merrill Flood and Melvin Dresher at RAND (1950)

Payoff Matrix

B / A	Cooperate	Defect
Cooperate	R, R	S, T
Defect	T, S	P, P

T: Temptation to defect
R: Reward for cooperation
P: Punishment for mutual defection
S: Sucker's payoff

For PD, $T > R > P > S$
Nash equilibrium for PD : both defect



Evolutionary Games Infographics Project

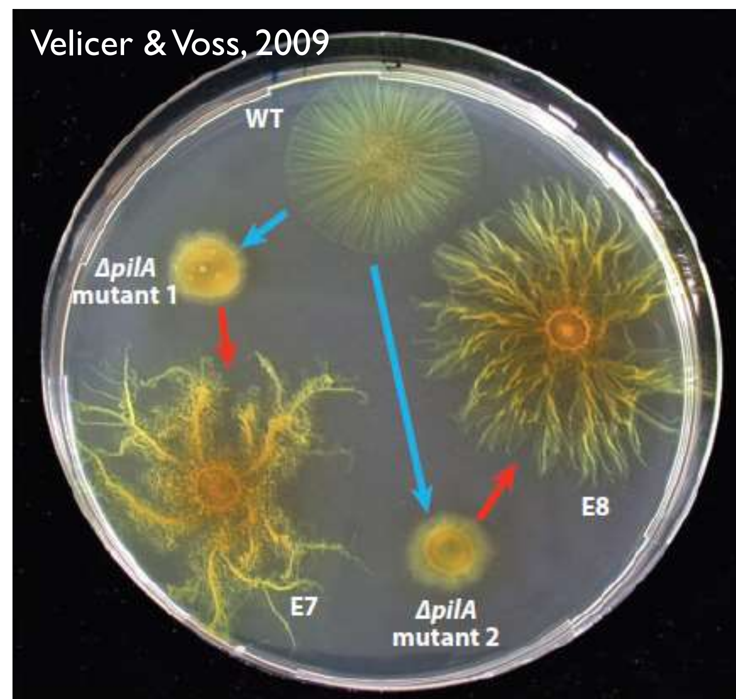
Prisoners Dilemma in Biology

Cleaner fish *Labroides dimidiatus*

Cooperation: eating ectoparasites off clients
Cheating/Defection: eating client mucus
As clients often leave in response to cheating the benefits of cheating can be gained by only one cleaner during a pair inspection

Male–female cleaner pairs jointly inspect "client" fish

Myxobacteria *Myxococcus xanthus*



M. xanthus collectively preys upon a broad range of microorganisms facilitated by gliding movement across solid surfaces generated by a **cooperative** social motility system requiring cell-cell contact to function
Cheaters can exploit cooperative production of exoenzymes, motility surfactants or signals, developmental signals, or secondary metabolites

Nash equilibrium for PD : both defect each getting payoff P
But mutual cooperation will result in higher payoffs R (> P) for both.

Puzzle: Rational action by individual agents result in undesirable collective outcome for all

Do such "social dilemmas" suggest game theory applies only to "rational fools" ?

Amartya Sen (1977) "Rational Fools: A Critique of the Behavioral Foundations of Economic Theory", *Philosophy & Public Affairs* **6**, 317

Results of experimental realizations of PD and other non-cooperative games incorporating such dilemmas also show deviation from Nash solutions…
Humans tend to be much "nicer" than what rational game-theoretic models would tend to suggest

Example:
PD played with prisoners

Image: Giulia Forsythe

Khadjavi and Lange put PD to test with a group of prisoners in Lower Saxony's primary women's prison, as well as students through both simultaneous and sequential versions of the game.

Prisoners don't betray one another.
In fact, they betray one another far less than college students do.
Students cooperated 37% of the time, while the prisoners cooperated 56% of the time.

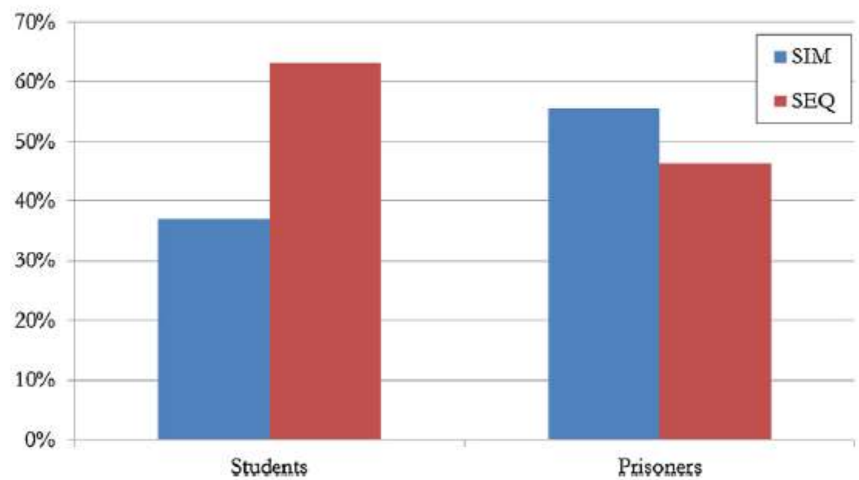


Fig. 1. Cooperation rates of students and prisoners, simultaneous and sequential PD.

Image: Khadjavi and Lange, Journal of Economic Behavior & Organization, 92, 163–175 (2013)

Puzzle:

Is cooperation seen in human societies "irrational" ?

Basis of several "social dilemmas"

e.g., Traveler's Dilemma (Koushik Basu)

"…majority of people do not use purely rational strategies, but the strategies they do use are demonstrably optimal" (Wikipedia entry on Travelers' Dilemma)

Paradox can be resolved by using a novel solution concept

Co-action Equilibrium

for payoff-symmetric games such as PD, in which the optimal action of rational agents is markedly different from Nash equilibrium

Symmetric games: the identities of the players can be exchanged without changing the payoff matrix – i.e., payoffs for playing a particular strategy depend only on the strategies employed, not on who is playing them.

The co-action solution concept

Agents in the **same information state** will act similarly.

Rationality assumption*: An agent argues that the other agent being equally rational as her and being in the same state as her, will use the same strategy as her.*

Agents in **different information states** will follow the usual Nash-like reasoning.

Implication for non-cooperative games:
Agents tend to behave much more "nicely" towards each other than in the Nash solution

2-person symmetric single-stage games

Each agents (*A* and *B*) has two possible strategies (Action 1 and Action 2)
Each agent receives a payoff corresponding to the pair of choices made by them:

Agent B

Agent A	Action 1	Action 2
Action 1	(R,R)	(S,T)
Action 2	(T,S)	(P,P)

An agent may employ a mixed strategy, in which she randomly selects her options, choosing Action 1 with some probability p and Action 2 with probability $(1 - p)$.
A pure strategy corresponds to p=0 or p=1

Nash solution for 2-person symmetric game

Assuming

- Agent A chooses Action 1 with probability p_1 and Action 2 with probability $(1 - p_1)$ and,
- Agent B chooses Action 1 with probability p_2 and Action 2 with probability $(1 - p_2)$

The expected payoffs are:

$$W_A = p_1(p_2(R + P - T - S) + S - P) + p_2(T - P) + P,$$
$$W_B = p_2(p_1(R + P - T - S) + S - P) + p_1(T - P) + P.$$

If the functions are non-monotonic in p_1, p_2 it is possible that a maximum exists inside the region [0,1], else the maximum will occur at the boundary (either 0 or 1) corresponding to a pure strategy.

The mixed strategy Nash equilibrium (if it exists) is:

$$p_1^* = p_2^* = \frac{P - S}{(R + P - T - S)}.$$

Nash assumptions are inconsistent for games in a completely symmetric setting

Assumptions:

- ❑ Each agent is aware that all others are rational & selfish just as her.
- ❑ Agents can make unilateral deviations in their strategy.

These two assumptions are mutually inconsistent for symmetric games

If the agents are aware that the others are also rational, they should take rational decision-making by the other agents into account.

Pre-cursors

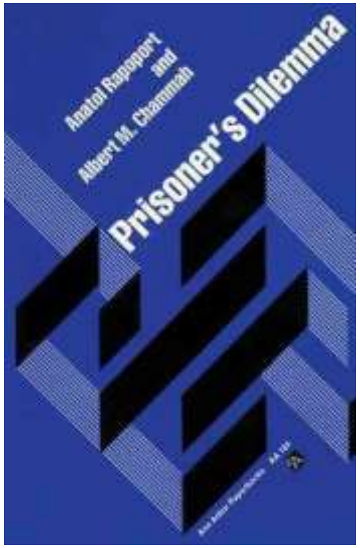
Inconsistency in assumptions of Nash solution for symmetric games alluded to by Rapoport in context of PD:

“Because of the symmetry of the game, rational players will choose the same action - and as it involves a higher payoff, they will always opt for mutual cooperation.” (1963)

Independently by Hofstadter (1983) in context of *N*-person PD.

Conventional response : Approaches rely on constraining the set of feasible outcomes of the game to the main diagonal of the payoff matrix (Binmore 1994).

Anatol Rapoport
(1911-2007)



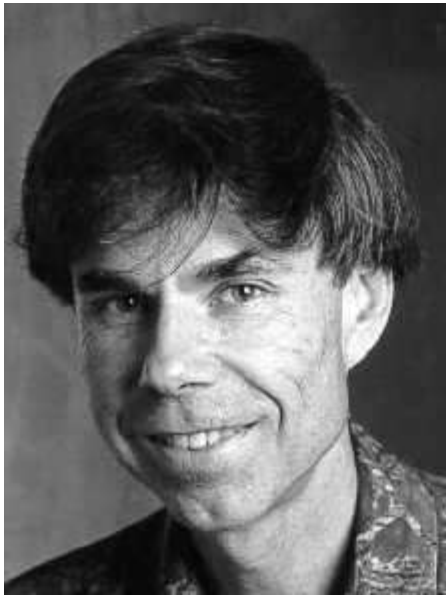
1965

METAMAGICAL THEMAS

The calculus of cooperation is tested through a lottery

by Douglas R. Hofstadter

© 1983 SCIENTIFIC AMERICAN,



Douglas Hofstadter
(1945-)

Co-action soln for 2-person symmetric games

Basis:

- Each agent, by virtue of the symmetry of the game, will realize that whatever complicated processes she employs in arriving at the optimal decision, the other agents will do exactly the same as they are equally rational (and have the same information and capabilities).
- Does not require any communication between the agents
- Does not invoke existence of trust or other extraneous concepts

The expected payoff function for each agent in the co-action solution concept is:

$$W_{A,B} = W = p^2(R+P-T-S)+p(T+S-2P)+P.$$

The mixed strategy co-action equilibrium (if it exists) is:

$$p^* = \frac{2P-(T+S)}{2(R+P-T-S)}.$$

Example: Prisoners Dilemma

T > R > P > S(=0)

The expected payoff function for each agent is

$$W = p^2(R + P - T) + p(T - 2P) + P.$$

Maximum at $\frac{(T - 2P)}{[(T - 2R) + (T - 2P)]}$

When T <= 2R, the optimum choice is $p^* = 1$

⇒ Both agents will cooperate

When T > 2R, there is a optimal mixed strategy:

$$p^* = \frac{T - 2P}{2(T - R - P)}.$$

⇒ as the temptation to defect becomes larger than 2R, the agents randomize between their options and can defect occasionally.

As *T* diverges, *p* converges to 0.5 (with corresponding payoff *T*/4)

Thus, the cooperation probability is between 0.5 and 1 (depending on *T*)

Unlike the Nash solution!

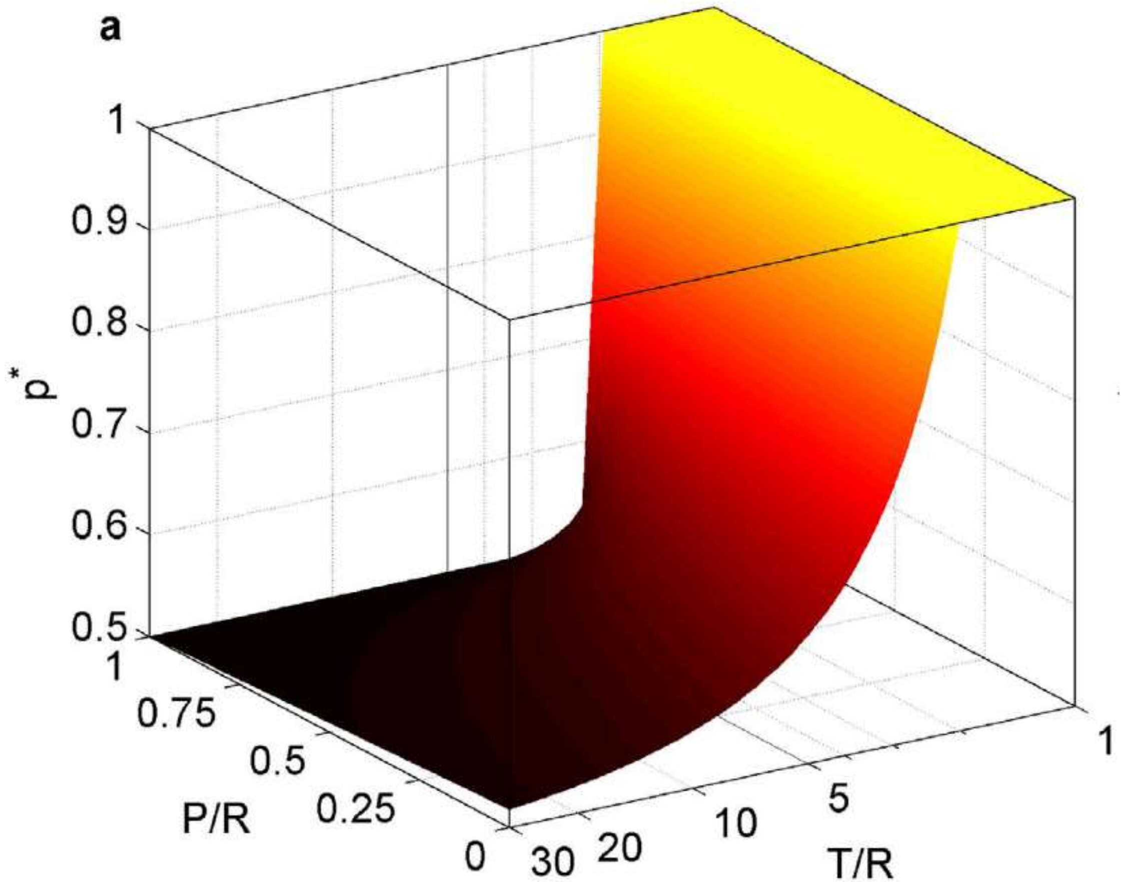
Nor do agents always cooperate (as suggested by Rapoport & Hofstadter)

Co-action allows *probabilistic* cooperation in single-stage PD

As T diverges, $p^* \rightarrow 0.5$ (with corresponding payoff $T/4$) $\Rightarrow$ cooperation probability for any payoff matrix lies between 0.5 and 1 (depending on T)

- Unlike the Nash solution!
- Nor do agents always cooperate (as suggested by Rapoport & Hofstadter)

As T $\rightarrow\infty$, strategy becomes fully random with either action chosen with equal probability.

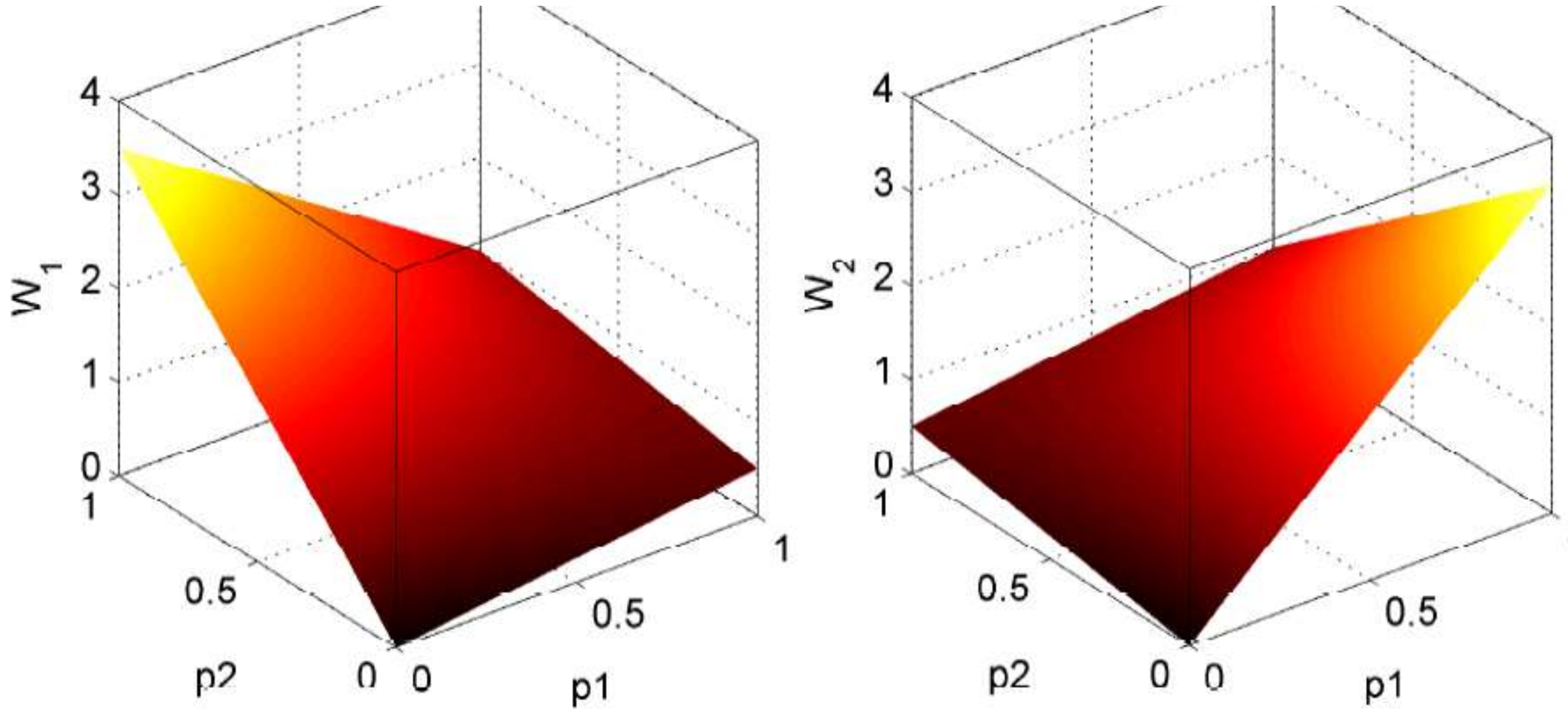


Variation of optimal strategy under co-action solution concept for PD as a function of payoff matrix elements

The optimal strategy also has a very weak dependence on P

A dynamical systems perspective

$$W_A = p_1 p_2 R + p_1(1-p_2)S + (1-p_1)p_2 T + (1-p_1)(1-p_2)P,$$
$$W_B = p_1 p_2 R + p_1(1-p_2)T + (1-p1)p_2 S + (1-p1)(1-p2)P.$$



The payoff functions for two agents A and B playing the game of Chicken, shown as functions of p1 and p2, i.e., the probability of each agent to choose Action 1.
The payoffs are T = 3.5, R = 1, S = 0.5 and P = 0.

A dynamical systems perspective

$$W_A = p_1 p_2 R + p_1(1-p_2)S + (1-p_1)p_2 T + (1-p_1)(1-p_2)P,$$
$$W_B = p_1 p_2 R + p_1(1-p_2)T + (1-p1)p_2 S + (1-p1)(1-p2)P.$$

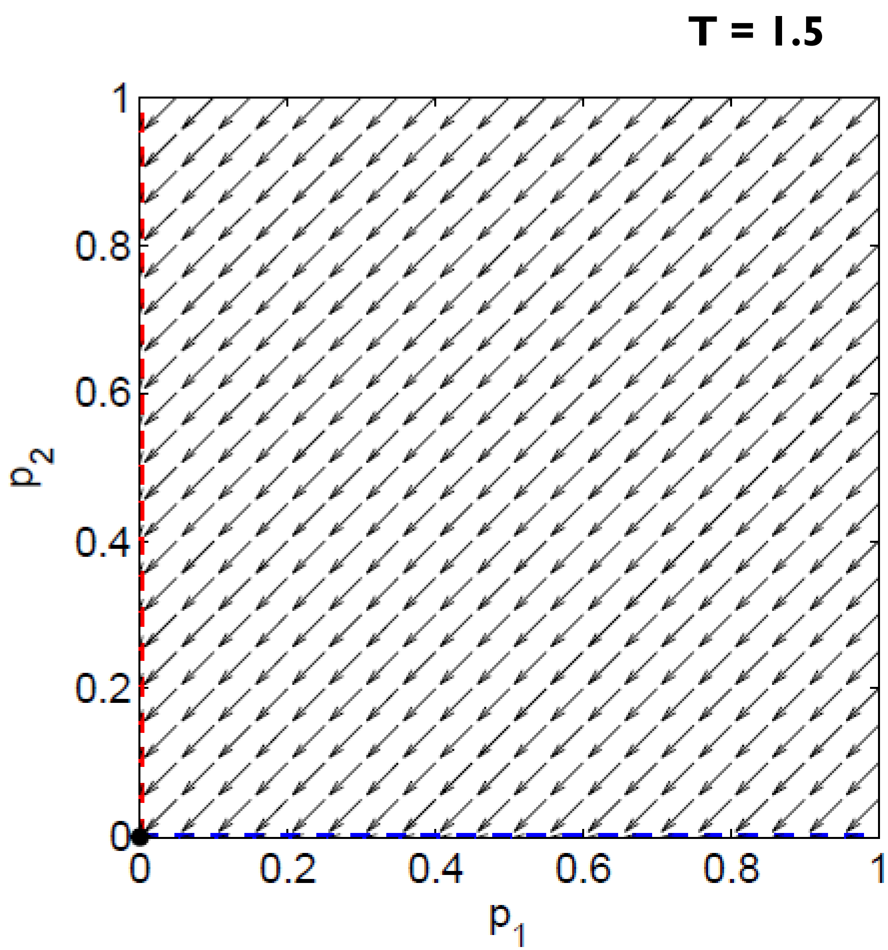
Nash solution:

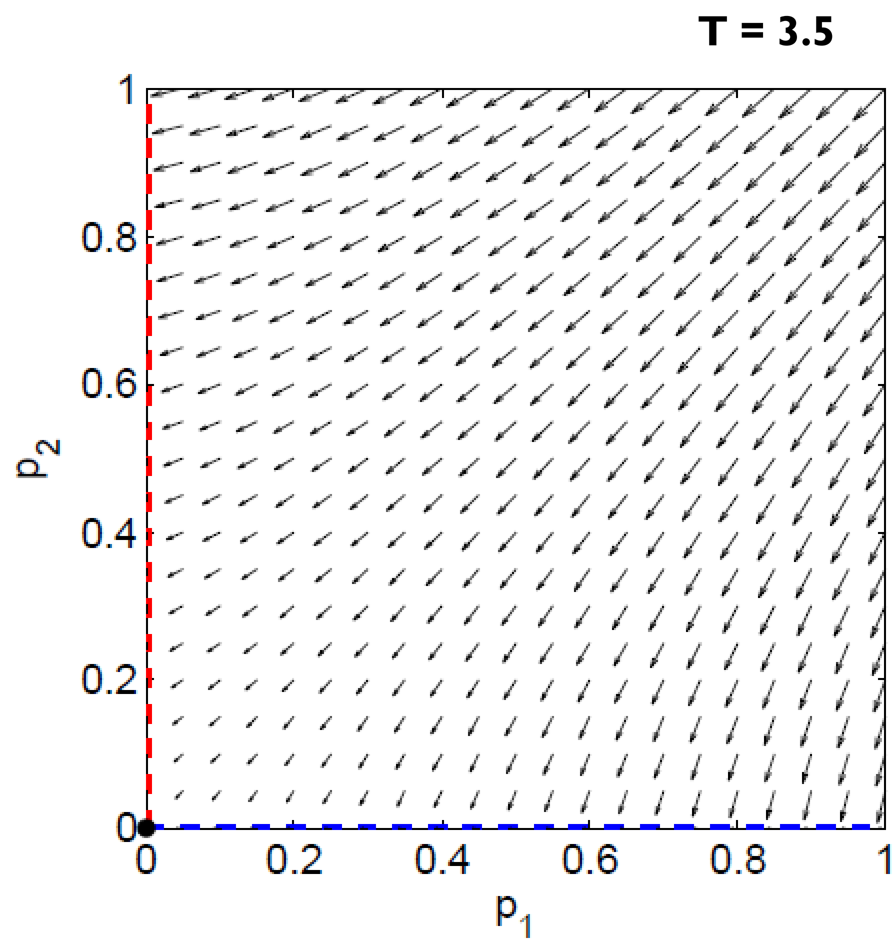
$$\frac{\partial W_A}{\partial p_1} = p_2(R-T) + (1-p_2)S,$$
$$\frac{\partial W_B}{\partial p_2} = p_1(R-T) + (1-p_1)S.$$

Co-action solution:

$$\frac{\partial W_A}{\partial p_1} = 2p_1 R + (1-2p_1)(T+S),$$
$$\frac{\partial W_B}{\partial p_2} = 2p_2 R + (1-2p_2)(T+S).$$

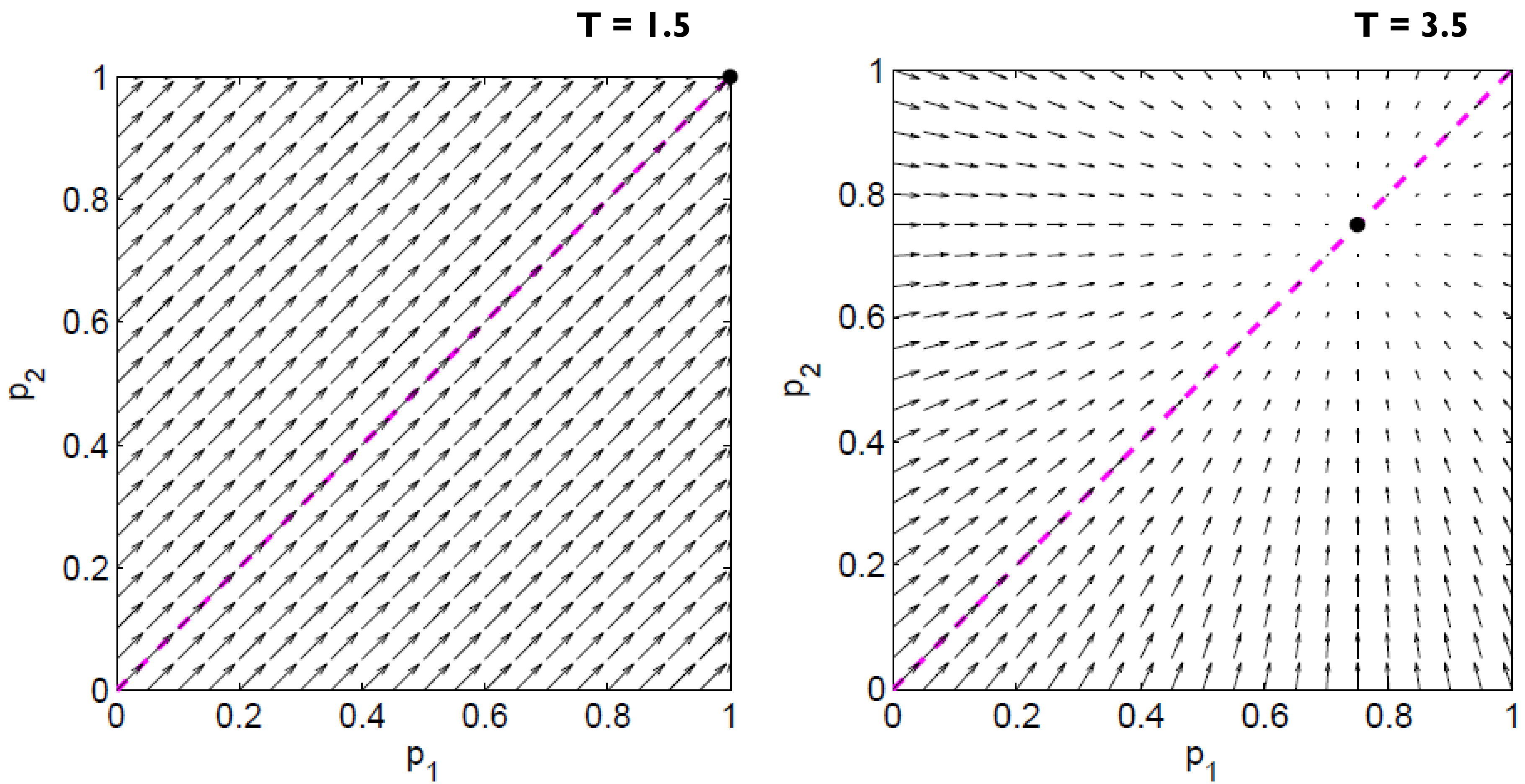
Nash solution of Prisoners Dilemma





Vector flow diagram representation of the Nash solution of the 2-person Prisoners Dilemma

Co-action solution of Prisoners Dilemma

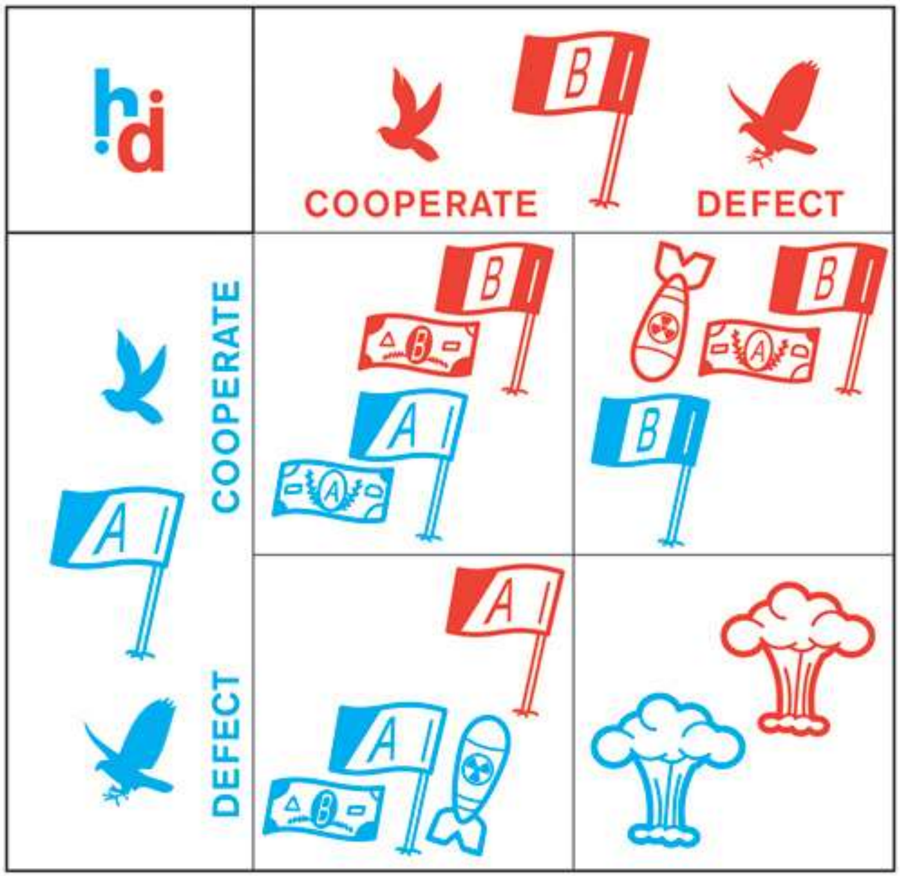


Vector flow diagram representation of the co-action solution of the 2-person Prisoners Dilemma

Example: Hawk-Dove or Chicken

T > R > S > P

Differs from PD in S being higher than P



Evolutionary Games Infographics Project

Strategy defined by probability p of Action 1 [probability of Action 2 is(1-p)]

Represents a strategic interaction between 2 agents choosing either

- ❑ Action 1: being docile, or
- ❑ Action 2: being aggressive

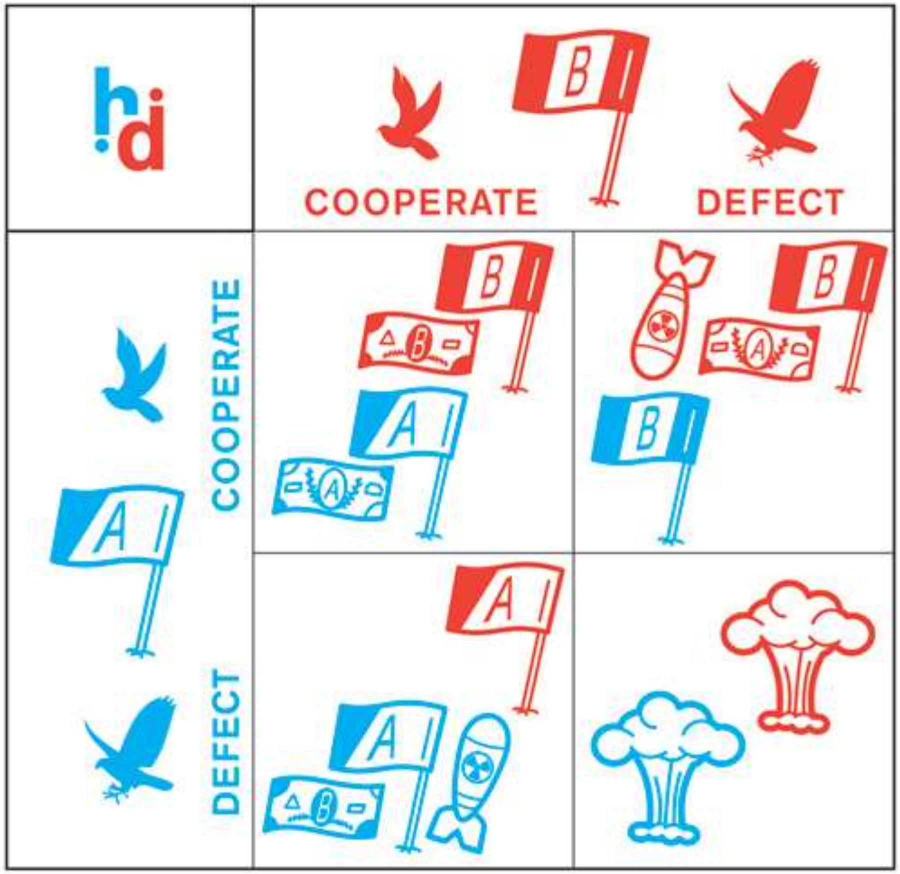
An agent benefits by being aggressive only if the other is docile but is better off being docile otherwise, as cost of mutual aggression is high.

Has been extensively investigated in the context of the study of social interactions and evolutionary biology

Solutions for Hawk-Dove/Chicken

T > R > S > P

Differs from PD in S being higher than P



Evolutionary Games Infographics Project

Strategy defined by probability p of Action 1 [probability of Action 2 is(1-p)]

Three Nash equilibria:

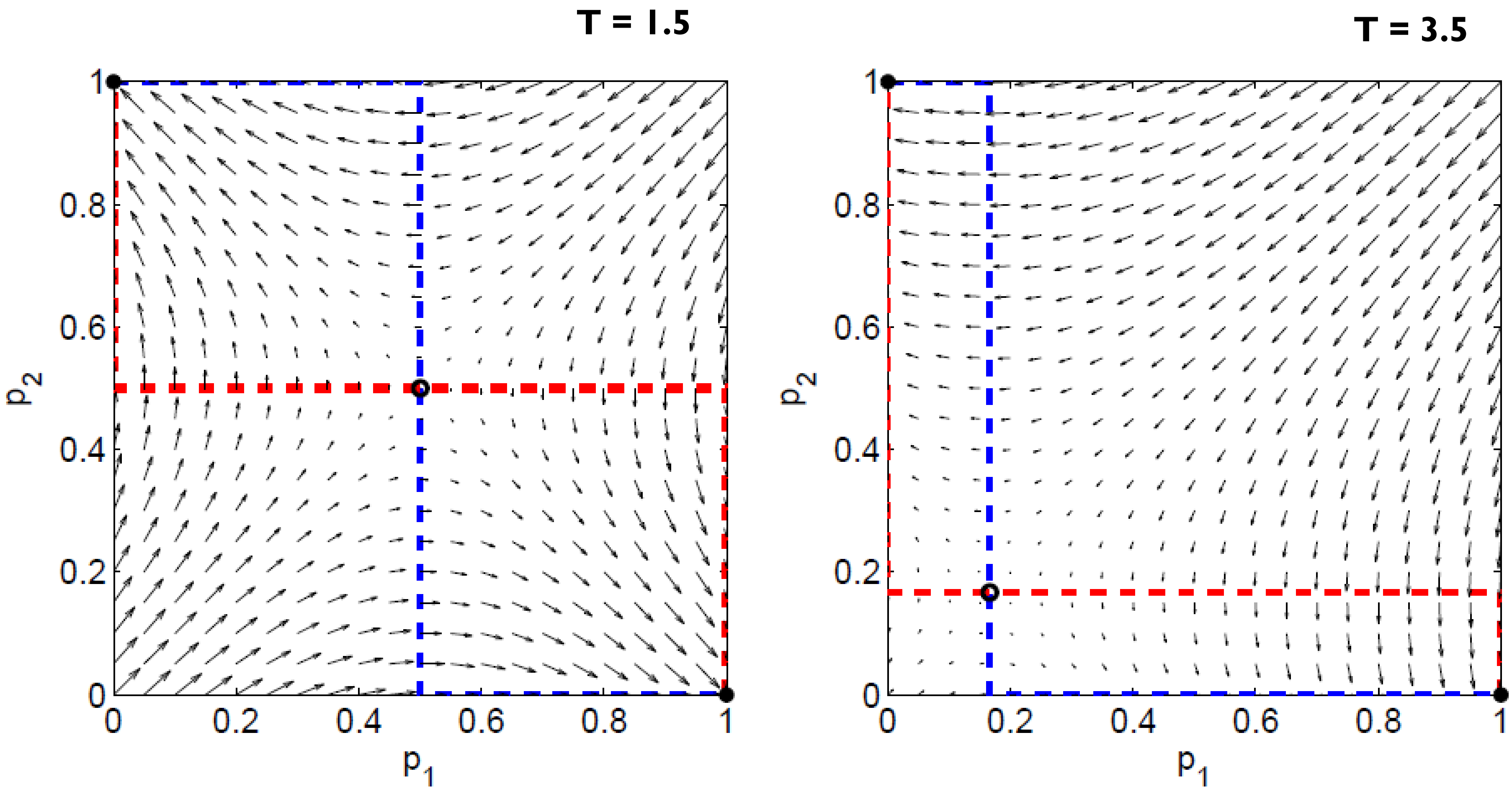
- a mixed strategy

$p^* = S/(T + S - R)$
[assuming P=0]
which is ESS

- pure strategy #1: $p^* = 1$
- pure strategy #2: $p^* = 0$

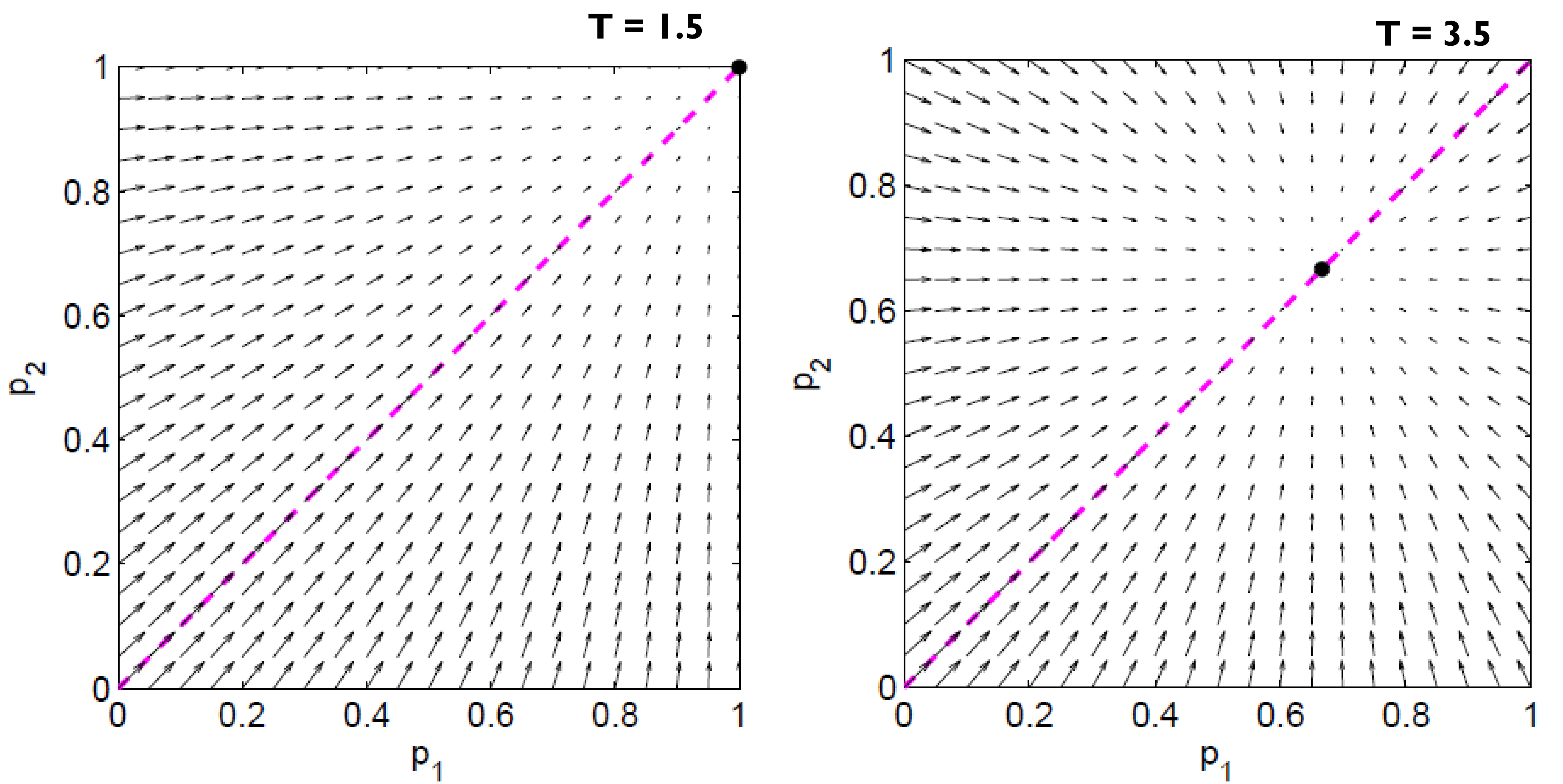
Single co-action equilibrium:
$p^*=(T+S)/2(T+S-R)$, when
$2R<T+S$
and
$p^*=1$, otherwise

Nash solution of Hawk-Dove or Chicken



Vector flow diagram representation of the Nash solution of the 2-person Snowdrift or Chicken

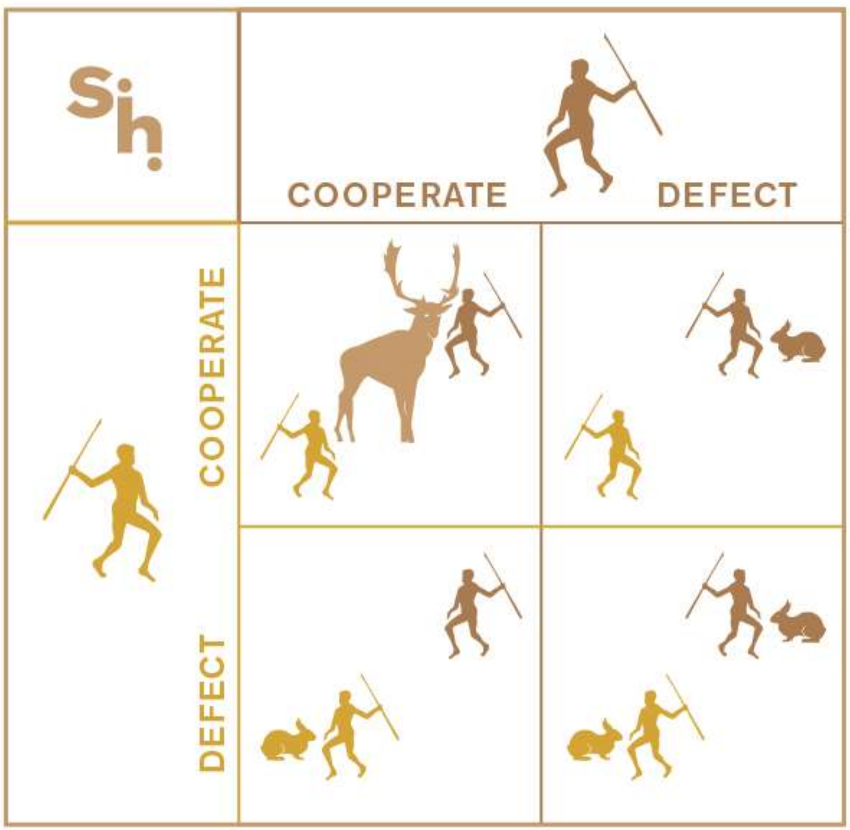
Co-action solution of Hawk-Dove or Chicken



Vector flow diagram representation of the co-action solution of the 2-person Snowdrift or Chicken

Example: Stag-Hunt

R > T ≥ P > S



Evolutionary Games Infographics Project

Strategy defined by probability p of Action 1 [probability of Action 2 is(1-p)]

Represents a strategic interaction between 2 agents choosing either :

- Action 1: a high-risk strategy having potentially large reward, viz., hunting for stag or
- Action 2: a relatively low-risk, but poor-yield, strategy, viz., hunting for hare .

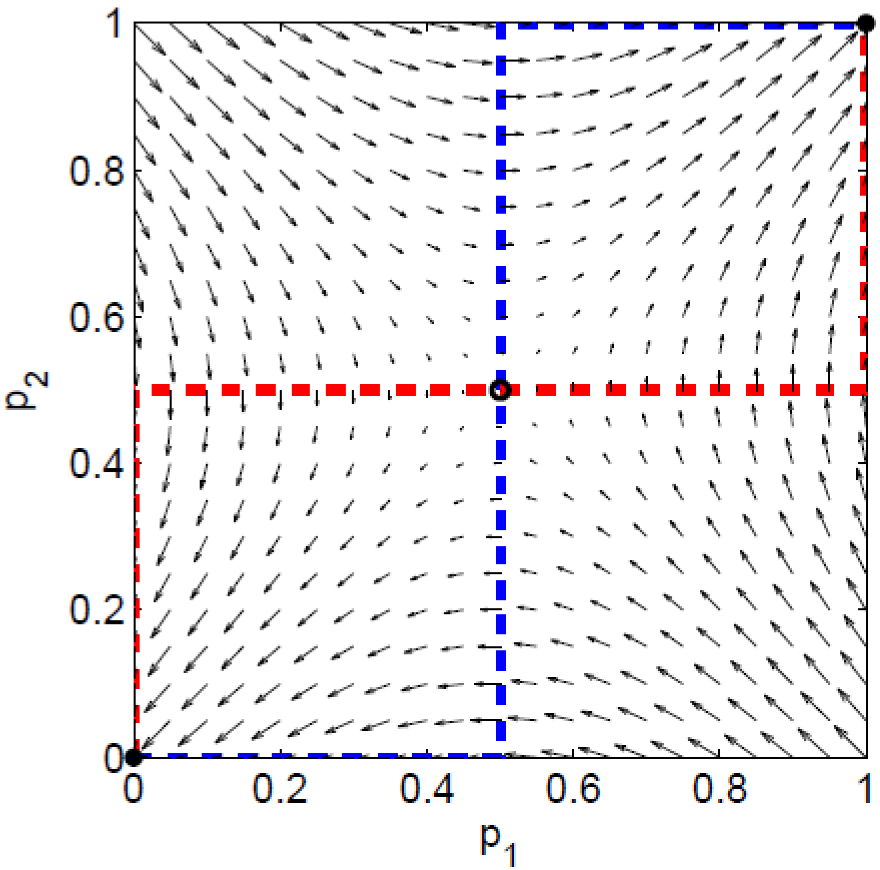
Describes many social situations where cooperation is required to achieve the best possible outcome

Three Nash equilibria: two pure strategies ($p^*=1$ and 0) which are also ESS and a mixed strategy $p^* = P/(P - R + T)$ [assuming S=0]

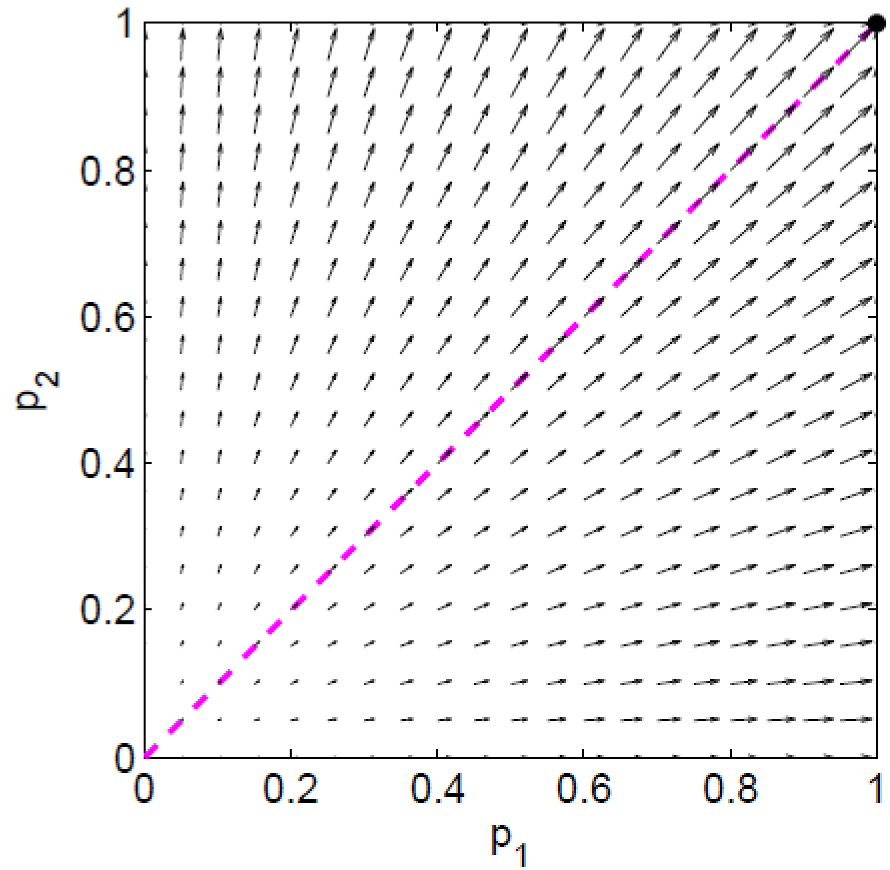
Single co-action equilibrium: $p^*=1$

Example: Stag-Hunt

Nash



Co-action



Vector flow diagram representations of 2-person Stag-Hunt (T = 0.5, S = – 0.5)

From one-shot to repeated interactions

- Symmetry is a crucial ingredient for co-action to apply.
- Such symmetry is more likely to be realized among members of a given community who share the same beliefs and a common identity.
- It is observed that cooperation is more common within an in-group than between agents belonging to different groups
- Significant levels of cooperative behavior reported in experimental realizations of social dilemmas can be explained by players ascribing to other players the same reasoning process as themselves and therefore resorting to co-action-like thinking.
- Players can become aware of symmetry through repeated interactions
- Let us look at iterated games

How does cooperation evolve through repeated interactions ?

Has been sought to be studied by the game of

2-person Iterated Prisoners Dilemma

Finite number of iterations always lead to Nash solution corresponding to Mutual Defection
Indefinitely continuing PD has no clear solution

Computer tournaments organized by Axelrod show the success of strategies such as Tit-for-tat in building cooperation

We shall consider

- Two agents play PD for several rounds

in addition to $T > R > P > S$, we assume $2R > T$

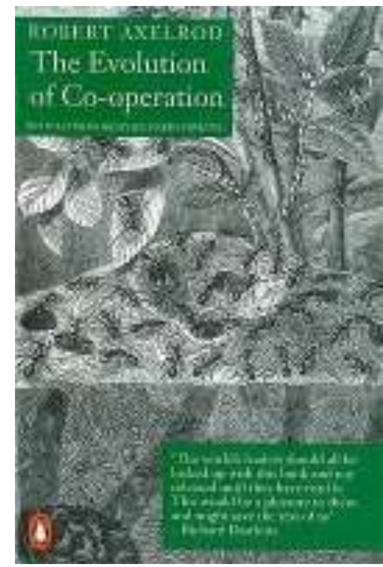
Rules out the possibility of agents alternately playing *C* and *D*

- consider memory-one strategies

Four possibilities can arise during each round of the game

CC,DD,CD,DC

Robert Axelrod
(1943 –)



1984

Possible strategies for Iterated Prisoners Dilemma

Utilizing information of player actions in the preceding round (memory 1): 16 possibilities

PRIOR PLAY			STRATEGIES			
Last move	Opponent's move	Outcome	Always Cooperate	Always Defect	Tit-for-Tat	Pavlov
C	C	"REWARD"	C	D	C	C
C	D	"LOSER'S PAYOFF"	C	D	D	D
D	C	"TEMPTATION"	C	D	C	D
D	D	"PUNISHMENT"	C	D	D	C

Image: Lisa Burnett, from
Martin A. Nowak, Robert M. May and Karl Sigmund, *Scientific American* **272**(6), 76-81 (1995)

2-person Iterated Prisoners Dilemma

Denote the state of a specific agent X, by $|C, n>$ or $|D, n>$

n: # agents cooperating in last round

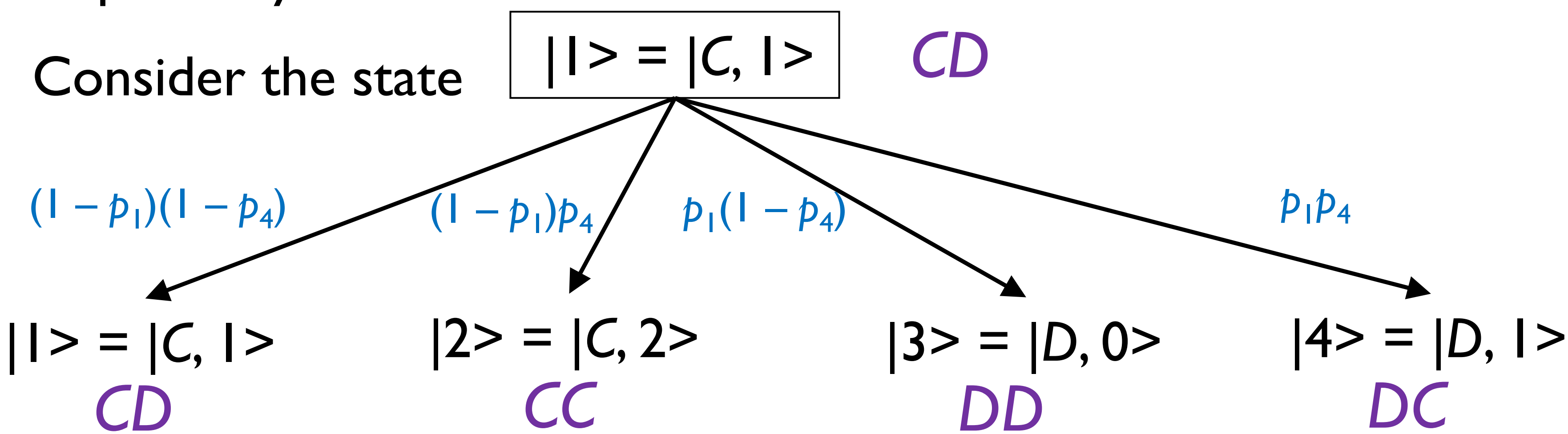
C,D: whether X was cooperating (C) or defecting (D) in last round.

p_1, p_2, p_3, p_4: switching probabilities to opposite action from states

$|1> = |C, 1>$, $|2> = |C, 2>$, $|3> = |D, 0>$, $|4> = |D, 1>$

respectively

Consider the state



Expected payoffs of the agent X in this state is:

$$W_1 = p_1 p_4\, T + p_1(1 - p_4)P + (1 - p_1)p_4\, R + (1 - p_1)(1 - p_4)S$$
$$= p_1\,(P + p_4\,(T - P - R)) + p_4 R$$

For simplicity, set $S = 0$

2-person Iterated Prisoners Dilemma

Expected payoffs of the specific agent X from states $|1\rangle = |C, 1\rangle$, $|2\rangle = |C, 2\rangle$, $|3\rangle = |D, 0\rangle$ and $|4\rangle = |D, 1\rangle$ respectively are

For simplicity, $S = 0$

$$W_1 = p_1 (P + p_4 (T - P - R)) + p_4 R,$$
$$W_2 = R - p_2 (2R - T) - p_2^2 (T - P - R),$$
$$W_3 = P + p_3 (T - 2P) + p_3^2 (R + P - T),$$
$$W_4 = T - p_4 (T - R - p_1 (T - R - P)) - p_1 (T - P)$$

Solution:

From $|C, 2\rangle$:

W_2 is a function of only $p_2 \rightarrow$ optimum value is $p_2^* = 0$.

From $|D, 0\rangle$:

W_3 is a function of only $p_3 \rightarrow$ optimum value is $p_3^* = 1$.

2-person Iterated Prisoners Dilemma

From |*C*, 1> or |*D*, 1>:
Agents are in different states and can follow Nash-like reasoning → select $p_1^* = 1$ and $p_4^* = 0$ (i.e., choose mutual defection).

Leads to the Pavlov strategy (Kraines & Kraines, 1989)

$$CC \rightarrow CC, CD \rightarrow DD, DC \rightarrow DD, DD \rightarrow CC$$

⇒ Co-operation is the steady-state outcome

- ❑ equivalent to the **win-stay lose-shift** strategy (Nowak & Sigmund, 1993)

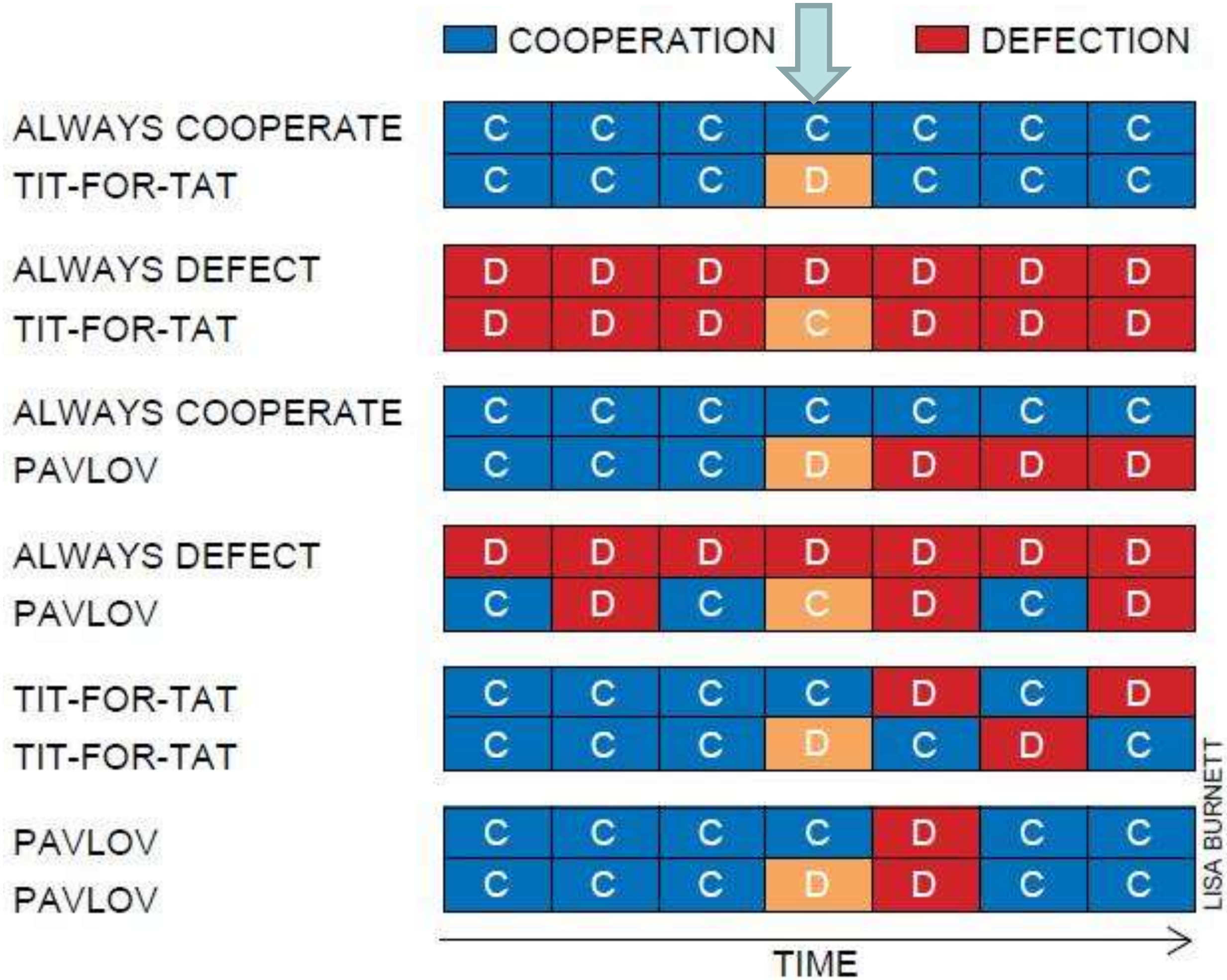
  Sustains behavior that brings reward but changes behavior that brings punishment → recalls simple conditioned response of **Pavlov**'s dogs

- ❑ Robust against noise – uncertainty, errors & misunderstanding
- ❑ Advantage over the tit-for-tat strategy (Axelrod 1984) for Iterated PD

The Importance of being "Pavlov"

Pairing off strategies against each other

"Mistake"



In presence of randomness, Pavlov is better at cooperating with itself than TFT

Image: Lisa Burnett, from
Martin A. Nowak, Robert M. May and Karl Sigmund, Scientific American 272(6), 76-81 (1995)

Iterated Prisoners Dilemma with N agents
Homogeneous Population

- *N* players: on each round, each player plays two-player PD with every other agent.

- Payoff of an agent is the sum total of payoffs from the two-player games.

- Deterministic, memory-one strategies are considered.

Also assume $P = S = 0$.

Under co-action equilibrium,

$$|C,N> \rightarrow |C,N>$$
$$|D,N> \rightarrow |C,N>$$

Iterated Prisoners Dilemma with N agents

Heterogeneous popn of cooperators & defectors

Expected payoffs of the two groups of agents can be represented by the payoff matrix

	$C_{N\text{-}i}$	$D_{N\text{-}i}$
C_i	$(N-1)R, (N-1)R$	$(i-1)R, iT$
D_i	$(N-i)T, (N-i-1)R$	0,0

- "Row player" is the group of *i* agents who cooperated in last round
- "Column player" is the group of (*N-i*) agents who defected in last round

Iterated Prisoners Dilemma with N agents

Generalizes Pavlov to contests involving multiple players

A dominant strategy analysis gives,

- A state in which everybody co-operate ($i = N$) or all but one agent co-operate ($i = N - 1$) is a stable state.
- States in which majority of the agents are co-operators will be stable when $T/R > (N - 1)/(N - i)$
- States in which minority of the agents are co-operators and majority are defectors will
  - go to all co-operation if $T/R < (N - 1)/(N - i)$,
  - will switch their respective choices otherwise.
- Special case: When N is even and $i = N/2$ and $T/R > 2(N - 1)/N$, multiple equilibria are possible.

Example: $N = 3$

If all three agents had chosen the same action (*C* or *D*) in the previous round, all of them cooperate in the next round.

In all other cases, the system converges to the state corresponding to two cooperators and one defector.

Clearly distinguishes the co-action approach from the conventional Nash solution, which would have corresponded to all three defecting.

A notable feature of the co-action solution is the stable coexistence of cooperators and defectors (as in state *S*2).

Example: $N = 5$

$$|C,1> \;\rightarrow |D,4>$$
$$|C,2> \;\rightarrow |C,5> \quad (\text{for } 4R > 3T)$$
$$\rightarrow |D,3> \quad (\text{for } 4R < 3T)$$
$$|C,3> \;\rightarrow |C,5> \quad (\text{for } 4R > 3T)$$
$$\rightarrow |C,3> \quad (\text{for } 4R < 3T)$$
$$|C,4> \;\rightarrow |C,4>$$
$$|C,5> \;\rightarrow |C,5>$$
$$|D,0> \;\rightarrow |C,5>$$

Solution also depends on future time-horizon of agents. Larger future time-horizon gives co-operation as the steady state outcome regardless of T/R.

Conclusions

- Co-action solution framework of games takes into account the symmetry of the situation, rectifying an inherent contradiction of the Nash solution approach
- The co-action equilibrium corresponds, in general, to nicer strategies (e.g., co-operation in single-stage Prisoners Dilemma).
- A dynamical interpretation in terms of vector fields
- In the Iterative Prisoners Dilemma, rational agents can become aware of symmetry through repeated interactions
- 2-player IPD co-action solution ≡ Pavlov strategy.
- In the general case of *N*-player IPD: Cooperators form the majority and coexist with defectors
- Co-action solution generalizes Pavlov to contests involving multiple players
- Tag-based cooperation among "similar" individuals (Riolo, Cohen & Axelrod) seen in biology can arise naturally in the co-action framework

THANKS

- V. Sasidevan & S. Sinha, *Scientific Reports* **5**, 13071 (2015)
- V. Sasidevan & S. Sinha, in *Econophysics and Data-Driven Modelling of Market Dynamics* (Springer, 2015) pp 213-223. (arxiv1501.05458)
- V. Sasidevan & S. Sinha, *Scientific Reports* **6**, 30831 (2016)
- S. N. Menon, V. Sasidevan & S. Sinha, in *Network Theory and Agent-Based Modeling in Economics and Finance* (Springer, 2019) pp 265-281. (arxiv1906.11683)