Can selfish rational agents achieve cooperation ?

V Sasidevan

Department of Theoretical Physics
IMSc, Chennai.

March 15 2014, ECONOPHYS Kolkata VIII

Plan of the talk

1. To cooperate or not: Prisoner's Dilemma
2. The co-action concept
3. Solution to Iterated Prisoner's Dilemma
4. N-player IPD
5. Conclusion

- The free rider problem:

Bob is writing a paper co-authoring with his colleague.

It is better for both to work on it and finish fast.

But it is even better for Bob if the other person writes up the paper and he free rides.

If both think in this way, the job will never get done.

- If a “collective good” is involved, individuals have little incentive to work towards achieving that good.
- Makes sense for others to do the work and sit back and reap the benefits of their labor.
- Then no one will do the work and the collective benefit would not be achieved.

In practice, we find that some degree of co-operation is achieved.

- Prisoner's Dilemma (PD): A widely employed metaphor for problems associated with co-operative behavior among selfish individuals.

		Player 2	
		Cooperate (C)	**Defect (D)**
Player 1	**Cooperate (C)**	**R** = 3, **R** = 3 Reward for mutual co-operation.	**S** = 0, **T** = 5 Sucker's payoff and Temptation to defect.
	Defect (D)	**T** = 5, **S** = 0 Temptation to defect and sucker's payoff.	**P** = 1, **P** = 1 Punishment for mutual defection.

Figure: Prisoner's Dilemma (PD): Strategies and payoffs.

For PD, we require $T > R > P > S$.

Player 1 \ Player 2	Cooperate (C)	Defect (D)
Cooperate (C)	**R** = 3, **R** = 3 Reward for mutual co-operation.	**S** = 0, **T** = 5 Sucker's payoff and Temptation to defect.
Defect (D)	**T** = 5, **S** = 0 Temptation to defect and sucker's payoff.	**P** = 1, **P** = 1 Punishment for mutual defection.

- Defection is a strictly dominant strategy for either player.
  Mutual defection is the only Nash equilibrium.
- Dilemma: Agents could have done better had they co-operated.
  Rational selfish agents do not co-operate even though it is in their best interest to do so.
- Also true if game is repeated a finite number of times.

- Co-operative behavior is observed quite regularly in experiments involving PD, especially when interaction is a repeated one (Andreoni 1993).
- Attempts to explain the observed co-operation:
  Conventional economist/game theorist say: People are behaving irrationally,
  Alternate view point: The assumptions in conventional theory are wrong. "rational fools" (Amartya Sen 1977).

  ϵ equilibrium (Radner 1986), assumption of incomplete information (Kreps 1982), bounded rationality (Neyman 1985).

  Changes the nature of the information available or the game in some way.

  Assume simple behavioral rules like: reciprocal altruism (tit-for-tat strategy)

- The co-action 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 make the same decision as her. (Sasidevan and Dhar 2014)*

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

  Solution to the one-shot PD: If a player believes that her opponent is in the 'same state' as her, she should co-operate.

  Similar ideas proposed in the context of PD: E.g. Super-rational agents (Hofstadter 1983), symmetry of agents (Goldberg 2004).

- Iterated Prisoner's Dilemma (IPD): The two players play PD for several rounds.
  in addition to $T > R > P > S$, we require $2R > T$.
- We consider memory-one strategies; also set $S = 0$
  Four possibilities can arise during the course of the game:

  $$CC, DD, CD, DC$$

- Denote the state of a marked agent X, by $|C, n>$ or $|D, n>$:
  n - How many agents were cooperating in total on last step.
  First letter denotes whether X was cooperating (C) or defecting (D) in the last time step.
- Let p_1, p_2, p_3, p_4 denote respective switching probabilities to opposite action from states $|1\rangle = |C, 1\rangle$, $|2\rangle = |C, 2\rangle$, $|3\rangle = |D, 0\rangle$ and $|4\rangle = |D, 1\rangle$.

- Expected payoffs of the marked agent in these states are,

$$
\begin{aligned}
W_1 &= p_1\left(P + p_4\left(T - P - R\right)\right) + p_4 R,\\
W_2 &= R - p_2(2R - T) - p_2^2(-1 - R + T),\\
W_3 &= P + p_3(T - 2P) + p_3^2(R + P - T),\\
W_4 &= T - p_4\left(T - R - p_1\left(T - R - P\right)\right) - p_1\left(T - P\right).
\end{aligned}
$$

Solution:

- From $|C,2\rangle$: W_2 is a functions of only p_2; Optimum value is $p_2^* = 0$.
- From $|D,0\rangle$: W_3 is a functions of only p_3; Optimum value is $p_3^* = 1$.

- From $|C,1\rangle$ or $|D,1\rangle$: Here agents are in different states. They follow Nash-like reasoning and select $p_1^* = 1$ and $p_4^* = 0$ (they go to mutual defection).

$$CC \to CC, \quad CD \to DD, \quad DC \to DD, \quad DD \to CC$$

Co-operation is the steady-state outcome.

Termed as win-stay lose-shift (or pavlov) strategy (Nowak 1993).

Has advantages over the widely considered tit-for-tat strategy (Axelrod 1984) for IPD.

- N players: on each round, each player plays two-player PD with everybody else.

  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,

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

- When there is a mixture of cooperators and defectors: The expected payoffs of these two groups of agents can be represented by a payoff matrix,

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

The "row player" is the group of agents, i in number, who cooperated in the last time step.
The "column player" is the group of agents, $(N-i)$ in number, who defected in the last time step.

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 $T/R < (N-1)/(N-i)$; otherwise will switch their respective choices.
- Special case: When N is even and $i = N/2$ and $T/R > 2(N-1)/N$, there are multiple equilibriums possible.

An example

- E.g. $N = 5$

$$
\begin{aligned}
|C,1> &\rightarrow |D,4> \\
|C,2> &\rightarrow |C,5> (4R > 3T) \\
&\rightarrow |D,3> (4R < 3T) \\
|C,3> &\rightarrow |C,5> (4R > 3T) \\
&\rightarrow |C,3> (4R < 3T) \\
|C,4> &\rightarrow |C,4> \\
|C,5> &\rightarrow |C,5> \\
|D,0> &\rightarrow |C,5>
\end{aligned}
$$

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.

Conclusion

- Adapting the co-action concept makes rational agents co-operate in an IPD.
- In the special two-player case, rational agents following the solution concept will adopt a win-stay lose-shift strategy.
- In the general case of N-player IPD: A group of majority cooperators and defectors may co-exist depending upon the ratio of the temptation payoff T to the reward payoff R.
- Results are true even for a finitely repeated game.

Future work: Probabilistic strategies in N-player IPD, on networks, other social dilemmas...

THANK YOU