The Evolution of Cooperation

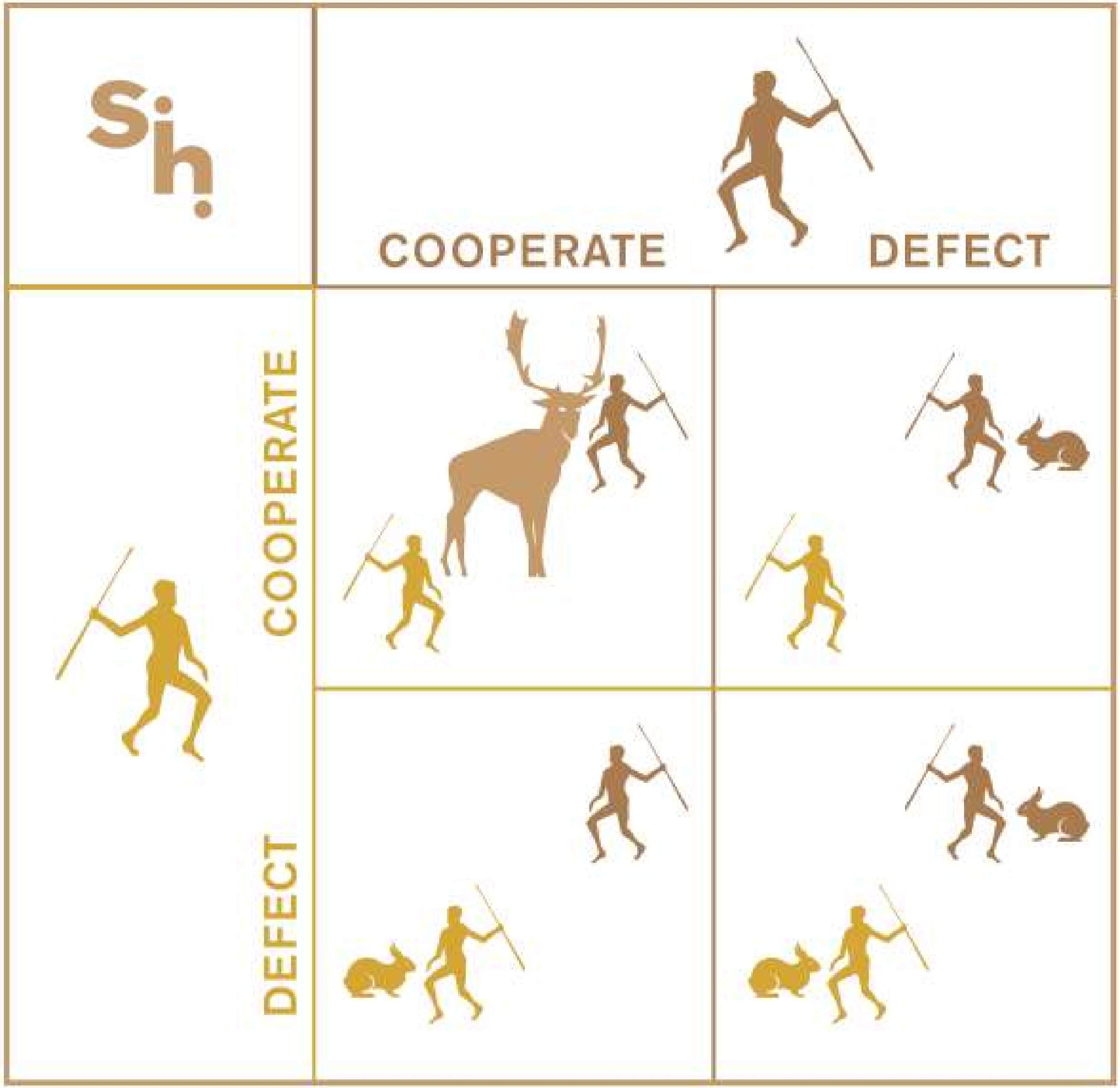
Robert Axelrod:

Under what conditions will cooperation emerge in a world without central authority?

In situations where each individual has an incentive to be selfish, how can cooperation even develop?

Stag Hunt (Jean Jacques Rosseau):

(Wikipedia)Jean-Jacques Rousseau described a situation in which two individuals go out on a hunt. Each can individually choose to hunt a stag or hunt a hare. Each player must choose an action without knowing the choice of the other. If an individual hunts a stag, he must have the cooperation of his partner in order to succeed. An individual can get a hare by himself, but a hare is worth less than a stag. This is taken to be an important analogy for social cooperation.

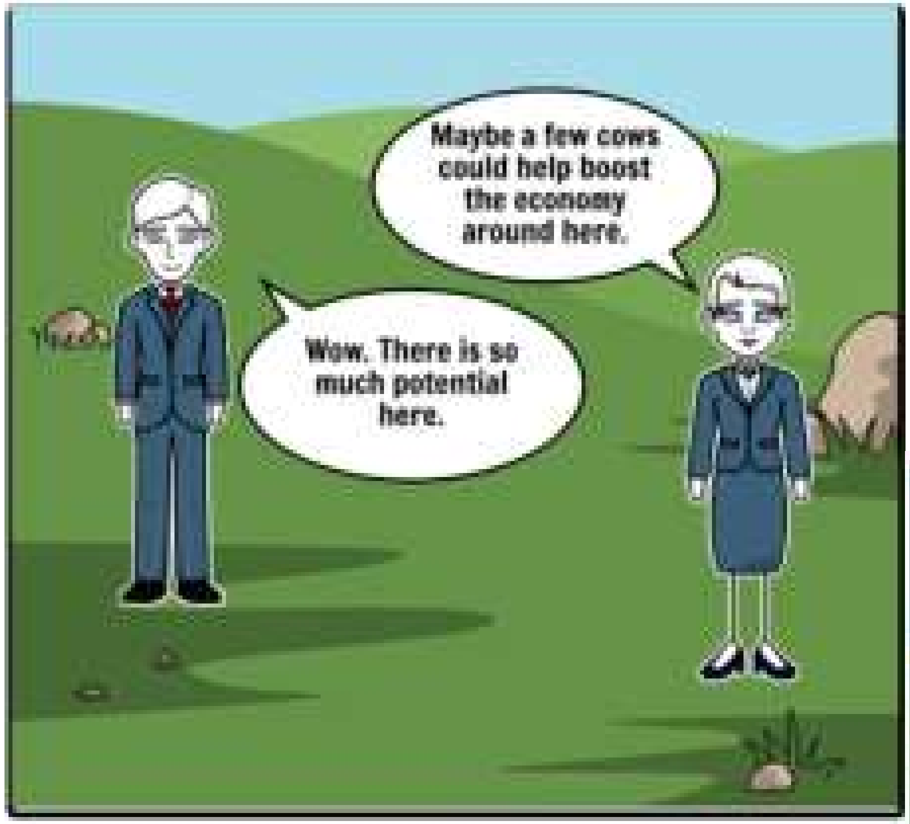

COOPERATE
DEFECT
COOPERATE
DEFECT

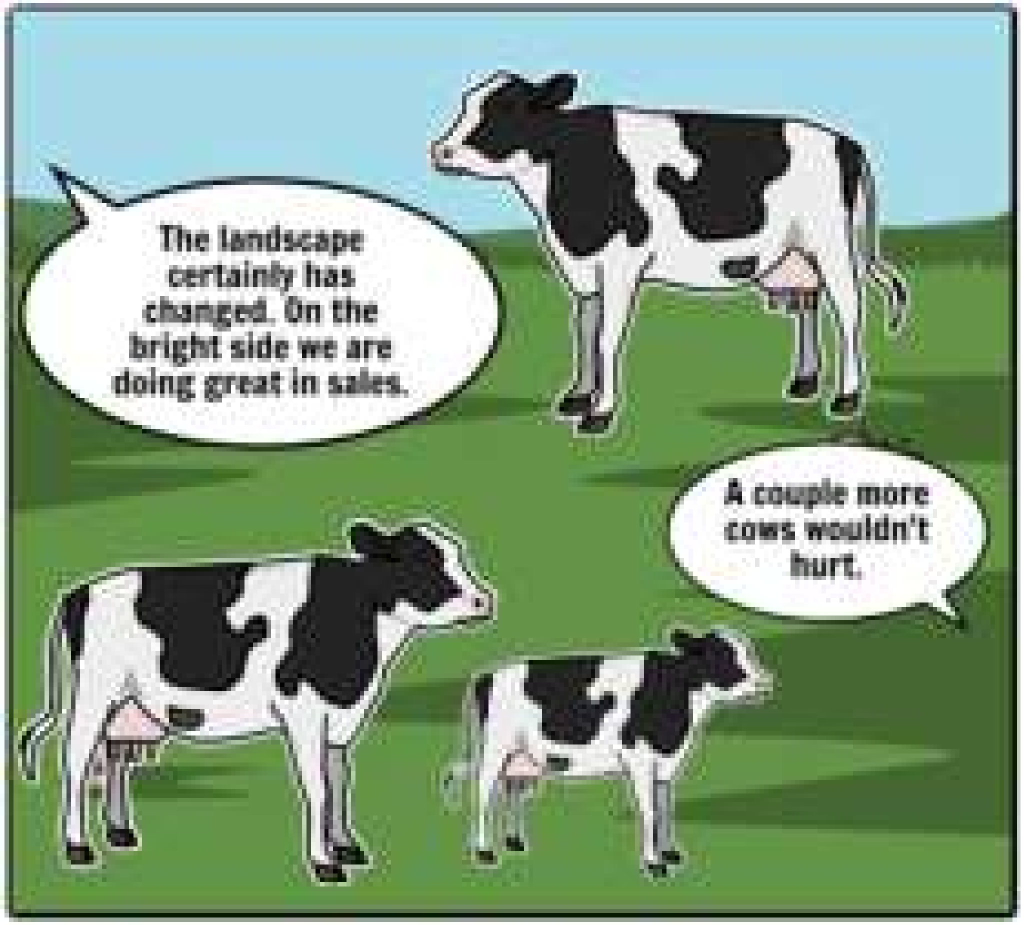
David Hume:

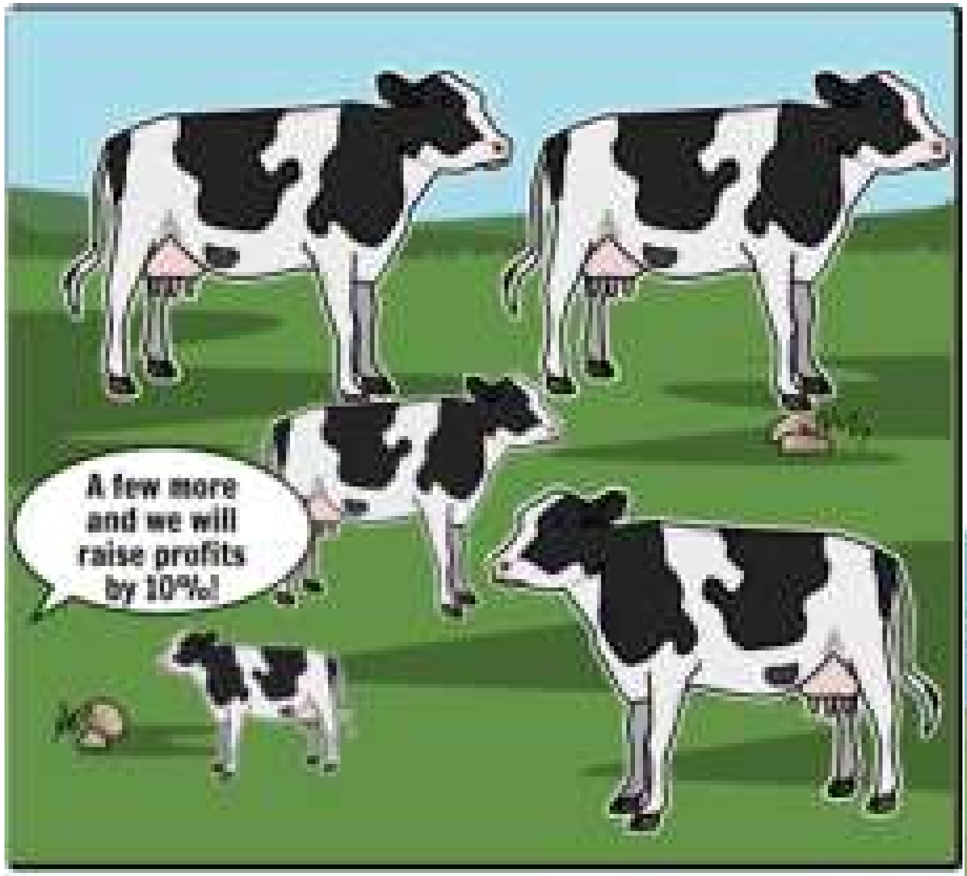
Two individuals who must row a boat. If both choose to row they can successfully move the boat. However if one doesn't, the other wastes his effort.

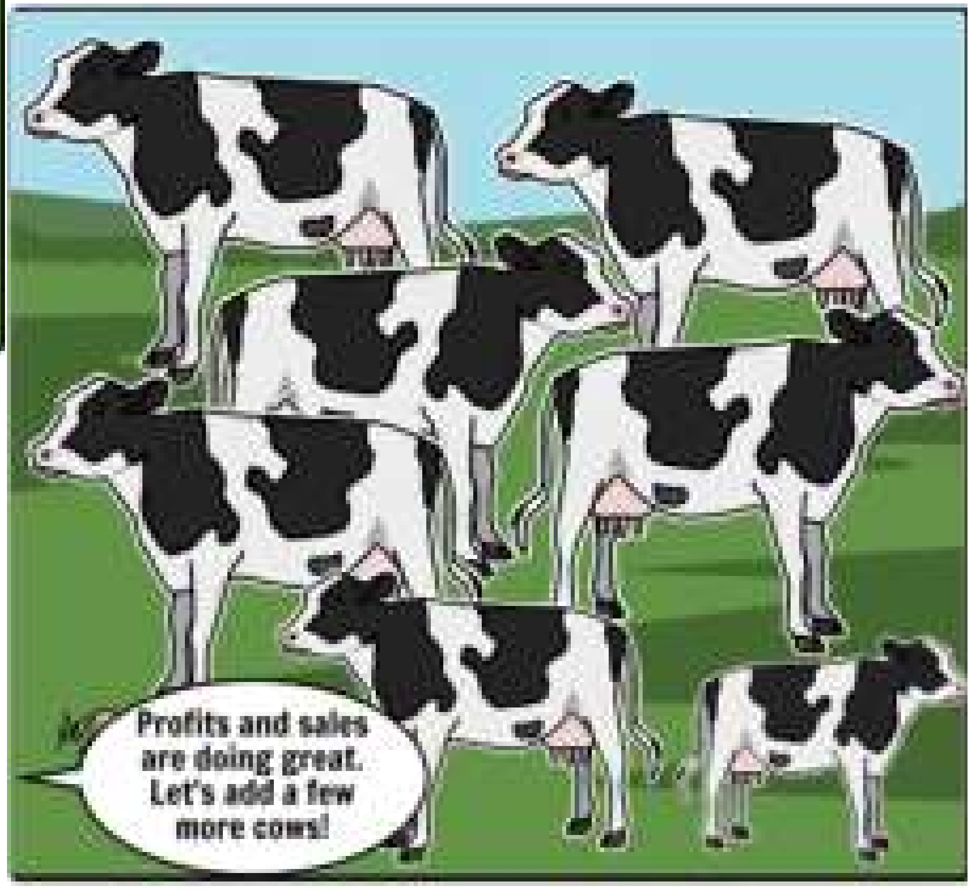
Tragedy of the Commons: (Wikipedia)

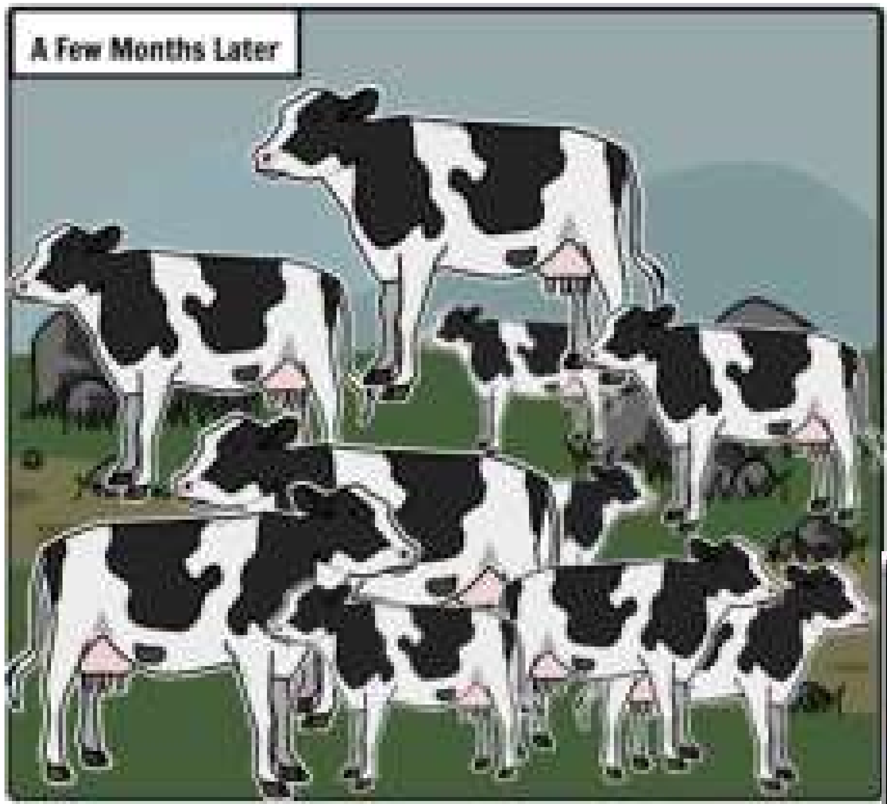
The **tragedy of the commons** is a dilemma arising from the situation in which multiple individuals, acting independently and rationally consulting their own self-interest, will ultimately deplete a shared limited resource, even when it is clear that it is not in anyone's long-term interest for this to happen. This dilemma was described in an influential article titled "The Tragedy of the Commons", written by ecologist Garrett Hardin and first published in the journal *Science* in 1968.

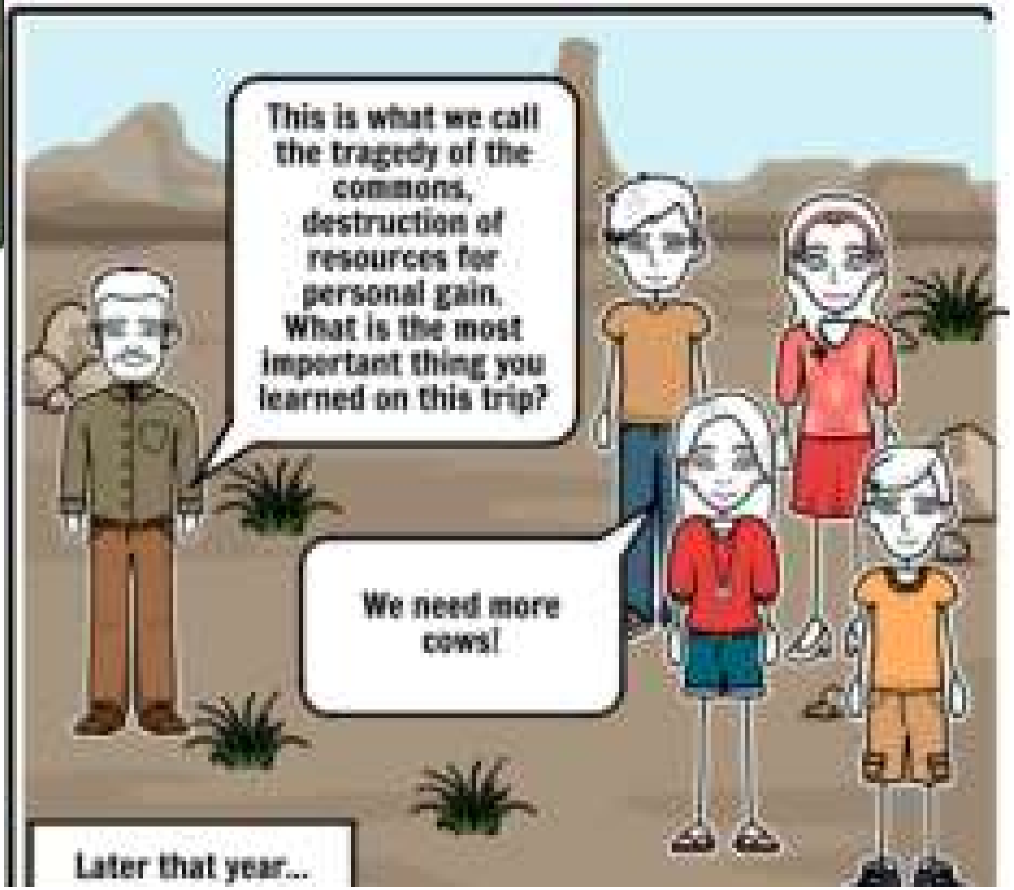

Maybe a few cows could help boost the economy around here.
Wow. There is so much potential here.


The landscape certainly has changed. On the bright side we are doing great in sales.
A couple more cows wouldn't hurt.


A few more and we will raise profits by 10%!

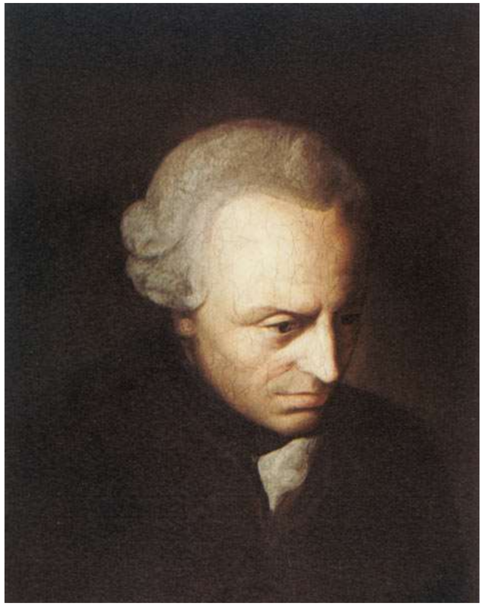

Profits and sales are doing great. Let's add a few more cows!


A Few Months Later


This is what we call the tragedy of the commons, destruction of resources for personal gain. What is the most important thing you learned on this trip?
We need more cows!
Later that year...

Categorical Imperatives of Kant:

"Act only according to that maxim whereby you can, at the same time, will that it should become a universal law."



己所不欲 勿施於人

"Act only according to that maxim whereby you can, at the same time, will that it should become a universal law."

己所不欲 勿施於人

Which one is better? Are they equivalent?

Repeated Prisoner's Dilemma

This is an extremely interesting topic. There is an extensive theory on Repeated Games but we'll not go into that. Students are recommended to read the book "The Evolution of Cooperation" by Robert Axelrod.

Prisoner’s Dilemma

$$\begin{array}{rc} & \begin{array}{cc}\text{Cooperate} & \text{Defect}\end{array} \\ \begin{array}{r}\text{Cooperate} \\ \text{Defect}\end{array} & \begin{pmatrix} R, R & S, T \\ T, S & P, P \end{pmatrix} \end{array}$$

T: Temptation, R: Reward, P: Punishment,

S: Sucker’s payoff.

T > R > P > S ensures <D, D> is a PSE.

2R > T + S promotes Cooperation

Example: Prisoner's Dilemma

$$
\begin{array}{ccc}
 & \text{Cooperate} & \text{Defect} \\
\begin{array}{c}\text{Cooperate} \\ \text{Defect}\end{array} & \multicolumn{2}{c}{\begin{pmatrix} 5,5 & -1,6 \\ 6,-1 & 0,0 \end{pmatrix}}
\end{array}
$$

T=6, R=5, P=0, S= -1

If we view this as a non-cooperative bimatrix game, the solution to this game is clear using the method of elimination by domination. Player I will not play Cooperate and the same for II. Therefore, the outcome of this game is for both to defect. There are many similar situations in real life.

1. Two stores engage in a price war. If the other store does not cut prices, I can attract its customers by cutting prices. If the other store cut prices, I had better cut prices so as not to lose my customers. If both stores reason this way and cut prices, both make lower profits than if neither had cut prices.
2. Two nations engaged in an arms race, the corresponding strategy choices might be “arm” and “don’t arm”, and the same reasoning might prevail.

3. Trade War (Axelrod): Two industrial nations have erected trade barriers to each other's exports. Because of the mutual advantages of free trade, both country would be better off if these barriers were eliminated. But if either country were to eliminate its barriers unilaterally, it would find itself facing terms of trade that hurts its own economy. In fact, whatever one country does, the other country is better off retaining its own trade barriers. Therefore, the problem is that each country has an incentive to retain trade barriers.

Tosca’s Dilemma

www.**ryerson.ca**/~dgrimsha/courses/cps720/prisoner.html

One of the world's most famous operas is Tosca, by Giacomo Puccini. Its plot nicely illustrates the Prisoner's Dilemma. The opera takes place during the time of Napoleon. There are 3 main characters, Tosca, Cavaradossi, an artist and revolutionary in love with Tosca, and Scarpia, chief of police, and also in love with Tosca. Tosca is in love with Cavaradossi, and hates Scarpia whom she sees as an ugly lecher.

Scarpia catches Cavaradossi and plans to execute him by firing squad. He also opens "negotiations" with Tosca. If she will make love to him, he will free Cavaradossi. She accepts his offer. There is a catch. For appearance sake, the execution must go ahead, but Scarpia will order that the bullets of the firing squad will be blanks. Cavaradossi will have to pretend to be dead, but afterwards, he and Tosca can escape.

Tosca comes to Scarpia's chambers. He embraces her, and she stabs him to death using a stiletto. This episode has become famous as "Tosca's kiss". At this moment the firing squad is heard firing.

Tosca runs to fetch her lover. He is lying on the ground. Of course, he really is dead. Scarpia had defected too. The firing squad used live bullets. In despair, Tosca leaps to her death off the battlements of the prison.

Will the situation change if the game is played twice?

The keyword is backward induction. At the 2nd time, both will Defect as this is the last game, there is no future. Knowing that both will Defect at the 2nd game, both will Defect in the 1st game. Therefore, <DD, DD> is a SE. There is no room for cooperation.

One can then extend the above reasoning to playing Prisoner's Dilemma 100 times etc.

Then, how about playing Prisoner's Dilemma infinite number of times?

- Question 1. Is it a realistic situation of playing a game infinite number of times?
- Question 2 How to compute the total payoff of a game played infinite number of times?

A solution to Question 2 is to introduce a discount factor β. \$$x$ at the n^{th} game is worth \$$\beta^{n-1}x$. Then, it makes sense to compute total payoff.

With this interpretation of total payoff, it is called a discounted game.

Interpretation of the discount factor.

1. Present Value: If the interest rate is r, then \$1 at state 1 is equal to \$$(1+r)^{n-1}$ at stage n. \$x at the n^{th} game has a present value of \$$(1+r)^{-n+1}$ x.

2. At the end of each stage, we roll a dice. The next game will be played with

probability β.

Question: What is a strategy of a game played infinite number of times?

The answer to the question is very complicated. We are only interested in some simple situations.

Finite Automata

1. Input: Some "history" of the game. Player II's action at the n^{th} stage.
2. Output: Player I's action at the $(n+1)^{th}$ stage.
3. A finite automata can only remember a finite number of things.

Example: 1-memory automata

The player can only remember what happened in the previous move and make decision based on that. We have to input the first move. Then, the player will make a move based on the 4 possibilities: CC, CD, DC, DD, where the first letter denotes the move of the player concerned and the second letter denotes the move of the opponent.

Then, a pure strategy with 1-memory is given by
f: $\{CC, CD, DC, DD\} \rightarrow \{C, D\}$.

There are 16 such pure strategies.

We can also make mixed strategies based on these 16 pure strategies.

Examples:

1. All D: Defect all the times.

First move=D, f(CC)=D, f(CD)=D, f(DC)=D, f(DD)=D

2. PR: Permanent Retaliation

Cooperate, until, if ever, the opponent defects, then defect for ever.

First move=C, f(CC)=C, f(CD)=D, f(DC)=D, f(DD)=D

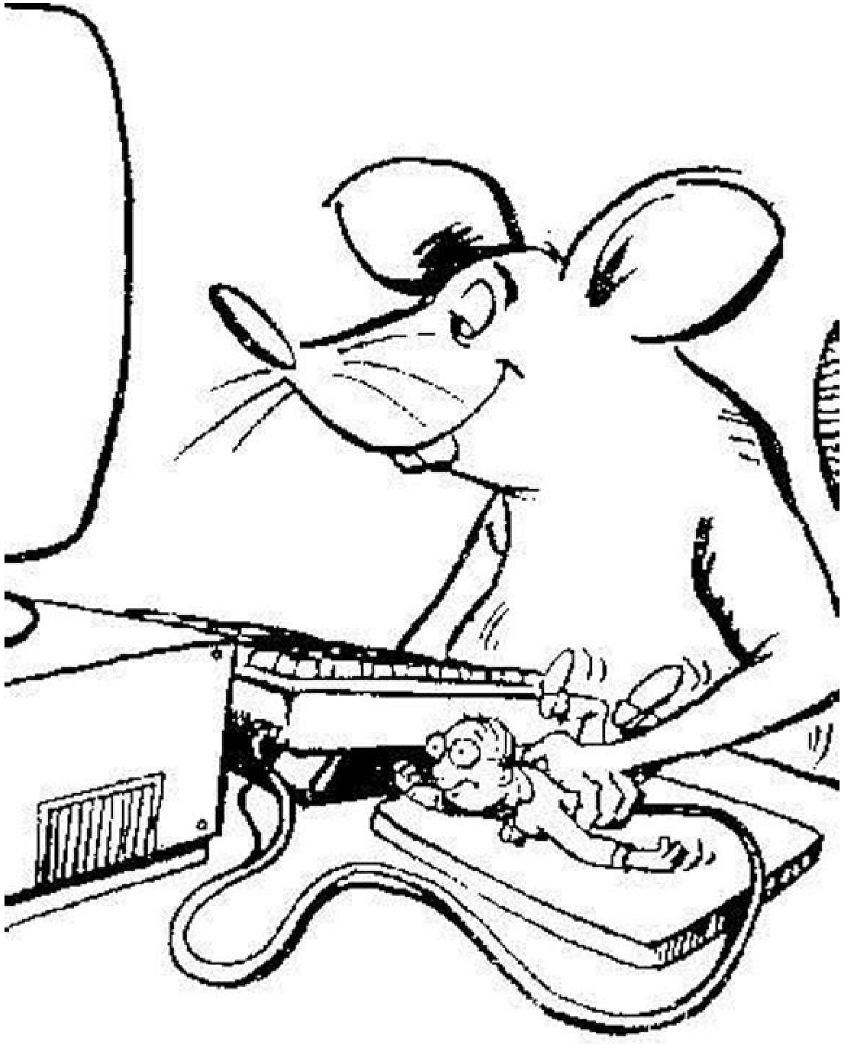
3. Tit-for-Tat (以牙還牙)： Cooperate first, then do what your opponent's previous move.

First move=C, f(CC)=C, f(CD)=D, f(DC)=C, f(DD)=D

4. AltDC: Start with D and then alternatively playing C and D.

First move=D, f(CC)=D, f(CD)=D, f(DC)=C, f(DD)=C

Tit-for-Tat



Expressing Mixed Strategies: When both players are using a 1-memory automata, we can use the technique of transition matrix in expressing the strategies. Each row of the matrix is a probability vector giving the probability of transiting to one of the 4 possible states (CC, CD, DC, DD).

	CC	CD	DC	DD
CC	p_{11}	p_{12}	p_{13}	p_{14}
CD	p_{21}	p_{22}	p_{23}	p_{24}
DC	p_{31}	p_{32}	p_{33}	p_{34}
DD	p_{41}	p_{42}	p_{43}	p_{44}

Let P denote the 4x4 transition matrix.

Let $v^T = (p_1, p_2, p_3, p_4)$ be a probability vector of the outcomes from the previous game.

Then, the result for the next game will be

$(v^T P)^T = P^T v$.

Then, if we play the game n times, the result will be $(P^T)^n v$.

Example: When Player I uses TFT and Player II uses TFT. We have the following transition matrix.

	CC	CD	DC	DD
CC	1	0	0	0
CD	0	0	1	0
DC	0	1	0	0
DD	0	0	0	1

Example: When Player I uses TFT and Player II uses PR. We have the following transition matrix.

	CC	CD	DC	DD
CC	1	0	0	0
CD	0	0	0	1
DC	0	1	0	0
DD	0	0	0	1

Theorem 1: <All D, All D> is a SE.

Proof: If Player I knows II defects all the time, I defects too.

If Player II know I defects all the time, II defects too.

Theorem 2: <PR, PR> is a SE if β is large enough ($T > R > P > S$, $2R > T + S$) .

Proof: Suppose that I knows II is using PR.

Payoff to I if I uses PR.

I: C C C ....

II: C C C ....

$R+R\beta+R\beta^2\ldots$

$=R(1-\beta)^{-1}$

$$\begin{array}{ll} \text{Cooperate} & \multirow{2}{*}{$\begin{pmatrix} R,R & S,T \\ T,S & P,P \end{pmatrix}$} \\ \text{Defect} & \end{array}$$

If Player I defects at the nth game, Player II will defect from the (n+1) th game on. Then, Player I must as well defect from the (n+1)th game on.

I: C… DDD…

II: C… CDD…

$$\begin{array}{ll} \text{Cooperate} & \\ \text{Defect} & \end{array} \begin{pmatrix} R,R & S,T \\ T,S & P,P \end{pmatrix}$$

The payoff to Player I is:

$$R + R\beta + \ldots + R\beta^{n-2} + T\beta^{n-1} + P\beta^{n} + P\beta^{n+1} + P\beta^{n+2} \ldots$$

$$= R(1-\beta^{n-1})(1-\beta)^{-1} + T\beta^{n-1} + P\beta^{n}(1-\beta)^{-1}$$

For PR to be best response to PR, we need

$$R(1-\beta)^{-1} \geq R(1-\beta^{n-1})(1-\beta)^{-1} + T\beta^{n-1} + P\beta^{n}(1-\beta)^{-1}$$

This means

$\beta \geq (T-R)/(T-S)$.

When $\beta \geqslant (T-R)/(T-S)$, the best response to Player II's PR is PR.

The same reasoning goes for Player II.

Therefore, <PR, PR> is a SE when β is large enough.

Theorem 3: When β is large enough,

<TFT, TFT> is a SE

$(T > R > P > S\ ,\ 2R > T + S)$.

Proof: Suppose now that I defects at the $(k+1)^{th}$ game and continue to defect up to and including the $(k+t)^{th}$ game. At $(k+t+1)^{th}$, Player I plays C. The play of the game is then as follows

- I: C ....CDD .... DC
- II: C ....CCD .... DD

k+1 k+t+1

$$\begin{array}{c} \text{Cooperate} \\ \text{Defect} \end{array} \begin{pmatrix} R,R & S,T \\ T,S & P,P \end{pmatrix}$$

If both use TFT, the play of the game is then as follows

- I: C ....CCC CC
- II: C ....CCC CC

k+1 k+t+1

For the payoff from this course of play to be better than that of the previous one, it suffices if the sum of payoffs from the $(k+1)^{\text{th}}$ and $(k+t+1)^{\text{th}}$ is better than the corresponding sum of the previous one (Note: R>P).

$$\begin{array}{c} \text{Cooperate} \\ \text{Defect} \end{array} \begin{pmatrix} R,R & S,T \\ T,S & P,P \end{pmatrix}$$

Therefore, it suffices to require

$$R\,\beta^k + R\,\beta^{k+t} \geq T\,\beta^k + S\,\beta^{k+t}$$

$$R + R\,\beta^t \geq T + S\,\beta^t$$

$$(R\text{-}S)\,\beta^t \geq (T\text{-}R)$$

Note that (R-S)>(T-R) follows from 2R>T+S. Hence, when $\beta>(T\text{-}R)/(R\text{-}S)$ TFT of Player I is a BR to Player II's TFT.

$$\begin{array}{c} \text{Cooperate} \\ \text{Defect} \end{array} \begin{pmatrix} R,R & S,T \\ T,S & P,P \end{pmatrix}$$

The same reasoning goes for Player II, we then have proved the theorem that

<TFT, TFT> is a SE when β is large enough ($\beta>(T-R)/(R-S)$ in this case).

TFT strategy in WWI

(Wikipedia)

Live and let live is the spontaneous rise of non-aggressive co-operative behavior that developed during the First World War particularly during prolonged periods of Trench Warfare on the Western Front.

It is a process that can be characterized as the deliberate abstaining from the use of violence during war. Sometimes it can take the form of overt truces or pacts negotiated locally by soldiers. At other times it can be a tacit behaviour - sometimes characterized as "letting sleeping dogs lie" - whereby both sides refrain from firing or using their weapons or deliberately discharging them in a ritualistic or routine way that signals their non-lethal intent.

TFT strategy in WWI

This behaviour was found at the small-unit level, sections, platoons or companies, usually observed by the "other ranks" e.g. privates and non-commissioned officers. Examples were found from the lone soldier standing sentry duty, refusing to fire on exposed enemy soldiers, through to snipers, machine-guns teams and including field-artillery batteries.

Tony Ashworth in his book *Trench Warfare 1914-1918: The Live and Let Live System* researched this topic based upon diaries, letters, and testimonies of veterans from the war. He discovered that Live and Let Live was widely known about, at the time, and was common usually at specific times and places. It was often to be found when a unit had been withdrawn from battle and was sent to a rest sector.

Remark: We should ask the question whether TFT is evolutionarily stable. Will a group using TFT will be invaded by some “mutants”?

In the following we study a population with three possible strategies AllD, AltDC and TFT.

We can write down the payoff matrix of

{AllD, AltDC, TFT} vs {AllD, AltDC, TFT}.

$$
\begin{array}{c}
\text{AllD} \\
\text{AltDC} \\
\text{TFT}
\end{array}
\begin{pmatrix}
\dfrac{P}{1-\beta} & \dfrac{P+T\beta}{1-\beta^2} & T+\dfrac{P\beta}{1-\beta} \\
\dfrac{P+S\beta}{1-\beta^2} & \dfrac{P+R\beta}{1-\beta^2} & \dfrac{T+S\beta}{1-\beta^2} \\
S+\dfrac{P\beta}{1-\beta^2} & \dfrac{S+T\beta}{1-\beta^2} & \dfrac{R}{1-\beta}
\end{pmatrix}
$$

TFT will be an ESS whenever

$$(*)\frac{R}{1-\beta} > T+\frac{P\beta}{1-\beta}$$

$$(**)\frac{R}{1-\beta} > \frac{T+S\beta}{1-\beta^2}$$

(*) is valid whenever $\beta > (T-R)/(T-P)$

(**)is valid whenever $\beta > (T-R)/(R-S)$

Thus, TFT cannot be invaded by AllD and AltDC whenever

$$\beta > \max\left((T-R)/(T-P),\ (T-R)/(R-S)\right).$$

A strategy is said to be

- Nice: Start cooperating and never the first to defect
- Retaliatory: It should reliably punish defection by its opponent.
- Forgiving: Having punished defection, it should be willing to try to cooperate again.
- Clear: It’s pattern of play should be consistent and easy to predict.

TFT possesses all the above properties.

Theorem 4: Let S be a nice strategy and that <S, S> is a SE. Then β is sufficiently large.

Proof: (Exercise)

Theorem 5: Let S be a nice strategy and that <S, S> is a SE. Then S should be provoked (when the opponent plays D, S must retaliate by playing D at some later move) by the defection of the opponent.

Proof: The payoff for <S, S> is $R/(1-\beta)$. If S cannot be provoked then the payoff to Player II using AllD is $T/(1-\beta)$ which is larger than that of <S, S>. This is a contradiction! Therefore, S must be provoked.

In 1980, Robert Axelrod, professor of political science at the University of Michigan, held a tournament of various strategies for the prisoner's dilemma. He invited a number of well-known game theorists to submit strategies to be run by computers.

In the tournament, programs played games against each other and themselves repeatedly. Each strategy specified whether to cooperate or defect based on the previous moves of both the strategy and its opponent.

- Some of the strategies submitted were:
- Always defect: This strategy defects on every turn. This is what game theory advocates. It is the safest strategy since it cannot be taken advantage of. However, it misses the chance to gain larger payoffs by cooperating with an opponent who is ready to cooperate.
- Always cooperate: This strategy does very well when matched against itself. However, if the opponent chooses to defect, then this strategy will do badly.
- Random: The strategy cooperates 50% of the time.

All of these strategies are prescribed in advance. Therefore, they cannot take advantage of knowing the opponent's previous moves and figuring out its strategy. The winner of Axelrod's tournament was the TIT FOR TAT strategy.

The strategy cooperates on the first move, and then does whatever its opponent has done on the previous move. Thus, when matched against the all-defect strategy, TIT FOR TAT strategy always defects after the first move. When matched against the all-cooperate strategy, TIT FOR TAT always cooperates.

This strategy has the benefit of both cooperating with a friendly opponent, getting the full benefits of cooperation, and of defecting when matched against an opponent who defects. When matched against itself, the TIT FOR TAT strategy always cooperates.

- Several variations to TIT FOR TAT have been proposed. TIT FOR TWO TATS is a forgiving strategy that defects only when the opponent has defected twice in a row. TWO TITS FOR TAT, on the other hand, is a strategy that punishes every defection with two of its own.
- TIT FOR TAT relies on the assumption that its opponent is trying to maximize his score. When paired with a mindless strategy like RANDOM, TIT FOR TAT sinks to its opponent's level. For that reason, TIT FOR TAT cannot be called a "best" strategy.
- It must be realized that there really is no "best" strategy for prisoner's dilemma. Each individual strategy will work best when matched against a "worse" strategy. In order to win, a player must figure out his opponent's strategy and then pick a strategy that is best suited for the situation.

How to promote cooperation?

(According to Axelrod):

1. Enlarge the shadow of the future:

The importance of the next encounter between the same two individuals must be great enough to make defection an unprofitable strategy.

Example: Once the US and USSR knew that they will be dealing with each other indefinitely, the necessary preconditions for cooperation will exist. The foundation of cooperation is not really trust, but the durability of the relationship.

2. Teach reciprocity:

There is no need to assume players are rational or trust between the players. The use of reciprocity can be enough to make defection unproductive. No central authority is needed. Cooperation based on reciprocity can be self-policing.

3.Teach people to care about each other

4. Change the payoffs

5. Improve recognition abilities

According to Martin Nowak the 5 mechanisms for the evolution of cooperation are:

Kin Selection

Direct Reciprocity

Indirect Reciprocity

Spatial Selection

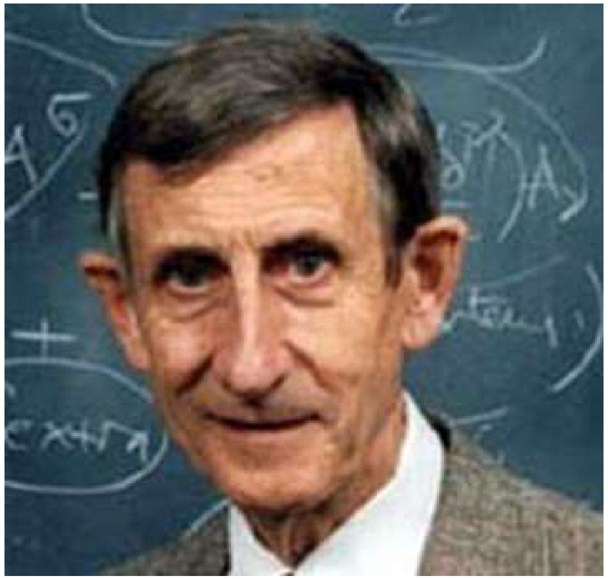
Group Selection

Robert Axelrod:
The key to doing well lies not in overcoming others, but in eliciting their cooperation.

Recent breakthroughs in Iterated Prisoner's Dilemma (Here they are concerned with the long term average payoffs, not total payoff with a discounted factor.)

Iterated Prisoner's Dilemma contains strategies that dominate any evolutionary opponent

By William Press and Freeman Dyson



Abstract: The two-player Iterated Prisoner's Dilemma game is a model for both sentient and evolutionary behaviors, especially including the emergence of cooperation. It is generally assumed that there exists no simple ultimatum strategy whereby one player can enforce a unilateral claim to an unfair share of rewards. Here, we show that such strategies unexpectedly do exist. In particular, a player X who is witting of these strategies can (*i*) deterministically set her opponent Y's score, independently of his strategy or response, or (*ii*) enforce an extortionate linear relation between her and his scores. Against such a player, an evolutionary player's best response is to accede to the extortion. Only a player with a theory of mind about his opponent can do better, in which case Iterated Prisoner's Dilemma is an Ultimatum Game.

Cooperative Games

In one version of the noncooperative theory, communication between the players is allowed but the players are forbidden to make binding agreements.

In the cooperative theory, we allow communication of the players and also allow BINDING AGREEMENTS to be made. This requires some outside mechanism to enforce the agreements.

In the noncooperative theory, the only believable outcome would be some Nash equilibrium because such an outcome is self-enforcing: neither player can gain by breaking the agreement.

With the extra freedom to make enforceable binding agreements in the cooperative theory, the players can generally do much better.

For example in the prisoner’ s dilemma, the only Nash equilibrium is for both players to defect. In the cooperative theory, they can reach a binding agreement to both use the cooperate strategy, and both players will be better off.

		Prisoner 2	
		Not Confess	Confess
Prisoner 1	Not Confess	$(5,5)$	$(-1,6)$
	Confess	$(6,-1)$	$(0,0)$

The cooperative theory is divided into two classes of problems depending on whether or not there is a mechanism for transfer of utility from one player to the other.

Transferable utility (TU) or **Nontransferable utility** (NTU)

Remark: For TU, we are assuming the utility scale for players are comparable. Note that it is an important topic in Economics concerning interpersonal comparison of utilities.

Feasible Sets of Payoff Vectors. For any game it is important to know the set of feasible payoffs vectors for the players.

For a noncooperative game, this set is called the noncooperative feasible set.

Let [A, B] be a bimatrix game. Then, the noncooperative feasible set is the set of all points of the form $(p^T A q, p^T B q)$, where p, q are mixed strategies of I, II respectively.

In general, it is not easy to plot the noncooperative feasible set.

- Example: Plot the noncooperative feasible set for Battle of Sexes

$$\begin{pmatrix} 2,1 & -1,-1 \\ -1,-1 & 1,2 \end{pmatrix}$$

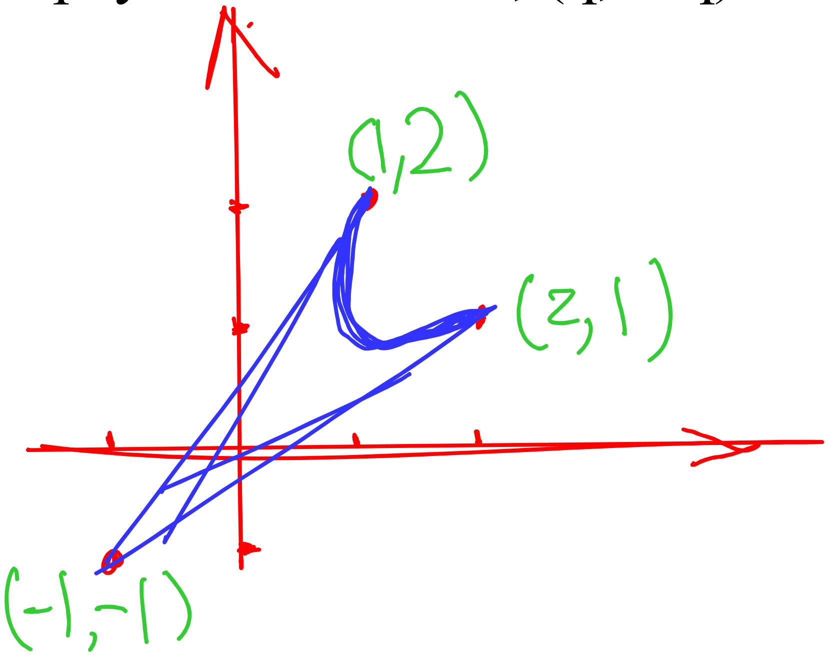
Example: Plot the noncooperative feasible set for Battle of Sexes.

Solution: For strategies <Row 1, (q, 1-q)> we get the line segment going from (-1,-1) to (2,1) as q goes from 0 to 1.

For strategies <Row 2, (q, 1-q)> we get the line segment going from (1,-2) to (-1,-1) as q goes from 0 to 1.

For fixed q, for strategies <(p,1-p), (q,1-q)> when p goes from 0 to 1 is a segment goes from payoff of <Row 2, (q, 1-q)> to that of <Row 2, (q, 1-q)>.

$2,1$	$-1,-1$
$-1,-1$	$1,2$



One of the main features of cooperative games is that the players have freedom to choose a joint strategy. This allows any probability mixture of the payoff vectors to be achieved.

The set of payoff vectors that the players can achieve if they cooperate is called the Cooperative Feasible Set.

.

When players cooperate in a bimatrix game with matrices [**A**,**B**], they may agree to achieve a payoff vector of any of the mn points, (a_{ij}, b_{ij}) for i = 1, . . . , m and j = 1. . . , n. They may also agree to any probability mixture of these points.

Therefore, the set of all such payoff vectors is the convex hull these mn points. Without a transferable utility, this is all that can be achieved.

Definition. The NTU feasible set is the convex hull of the mn points, (a_{ij}, b_{ij}) for $i = 1, \ldots, m$ and $j = 1 \ldots, n$.

Example: Plot the NTU feasible set of Battle of sexes.

2, 1	−1, −1
−1, −1	1, 2