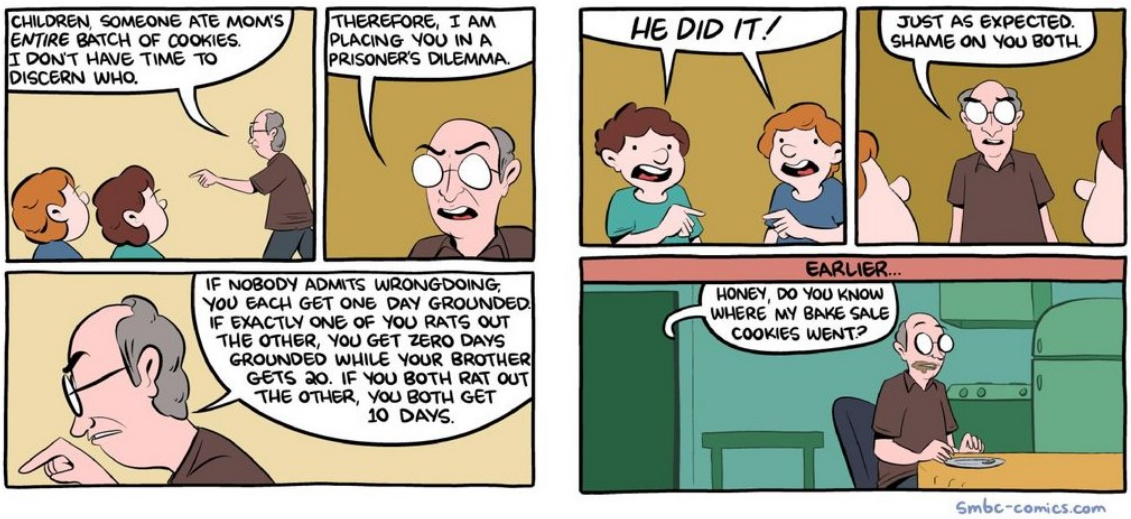
A Summer of Science report on

Game Theory



by
Akkapaka Saikiran

Mentor
Kusumit Ghoderao

Contents

Acknowledgements

I'd like to first express gratitude to my mentor Kusumit Goderao who kept badgering us when deadlines were near to keep us going on. He was also the one who informed us about the coursera course and gave us reference material midway and even met us after the semester was over before SoS had started.

A big shoutout to Nicky Case for developing the amazing game which was "The Evolution of Trust". It was so well I couldn't help but pick images and parts of writing directly from it.

Thanks to my co-mentee Amal Matthew, for being the guy in the same boat.

To my parents and my brother for just being there.

To Wikipedia, which has become so comprehensive, versatile and most importantly, so reliable in recent days.

And thanks lastly to you reader, for you made me put my mind and heart into this report, in the hope you'll get something from this.

1 Introduction

Game theory is the science of strategy, i.e of optimal decision making of "players" in a strategic setting. This strategic setting is what we refer to as a game. Game Theory originated as a theory in mathematics and economics but now has evolved to have applications in many social sciences and computer science.

The aim of this SoS reading project is to familiarise myself with the formal notions of strategy and to develop a mindset for the same. I will also try to touch upon ideas from game theory which are more pertinent for computer science, which is my branch of study here.

A word on pronouns and gender. Male pronouns have been used to refer to nameless agents. A fair deal of thought was given to this but in the end the "standard" male convention was settled upon rather than the reverse female convention or the grammatically-dubious "they". The reader is urged not to read patriarchal intentions between the lines.

2 Games in Normal Form

A N player normal form game consists of: I) a finite set of N players, II) action spaces for the players:$A_1, A_2, \ldots A_N$; III) Payoff or utility functions for the players: $u_i : A = A_1 \times \ldots \times A_N \mapsto \mathbf{R}$.

The action space A_i refers to a collection of actions which the player can choose from and play. Payoff (or utility) is what each player "gets" from the game. It is the entity everyone tries to maximise. It is a function of every player's choice of action/strategy.

Games are represented in the normal form using a matrix. Almost all the games we'll be exploring in this report will be 2-player games so the matrix will be 2 dimensional. The column player is by default player 1 and the row player is player 2. The payoff is a pair whose first number is the payoff to player one and likewise. A few examples follow.

Coin machine

You have one choice. In front of you is a machine. If you put a coin in the machine, the other player gets three coins – and vice versa. You both can either choose to cooperate (put in coin), or cheat (don't put in coin). This game is an instance of the celebrated Prisoner's Dilemma.

	cheat	cooperate
cheat	0,0	3,-1
cooperate	-1,3	2,2

Table 1: the normal form representation of the coin machine game.

The ideal penalty shootout

The players are a shooter and a keeper. If the player kicks in a direction and the keeper dives to the same direction then the goal is saved and keeper gets a payoff of 1 and player 0 - vice versa if not.

This sort of a game is called a constant-sum game as the sum of the payoffs is constant in every outcome of play. A special case is the *zero-sum game*, wherein as the name suggests the sum of payoffs of each player is 0. Constant-sum games can be modified into zero-sum games by adding a “dummy player” whose payoff is the negative of the constant sum.

	left	right
left	0,1	1,0
right	1,0	0,1

Table 2: the ideal penalty shootout.

Movie dilemma

Two friends decide to go to watch a movie. Sadly, they have different choices. If neither of them compromises then they're both sad as they wanted to go together and both of them get a payoff of 0. If however, one of them decides to fall in, then he gets a payoff of 1 while his friend gets 2. The movies have been innovatively named X and Y respectively by me.

	X	Y
X	2,1	0,0
Y	0,0	1,2

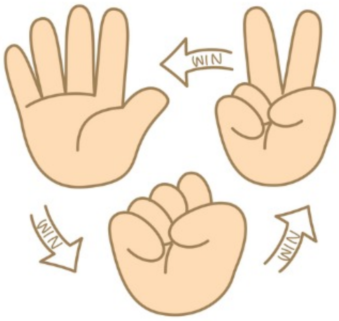
Table 3: the movie dilemma

Stone, paper, scissors!

This is a popular hand game usually played between two people, in which each player simultaneously forms one of three shapes with an outstretched hand. Stone beats scissor, scissor beats paper, paper beats stone. If both players do the same gesture then it's a draw. Inappropriate gestures lead to direct disqualification. This is also a zero-sum game.

	stone	paper	scissors
stone	0,0	1,-1	-1,1
paper	1,-1	0,0	1,-1
scissors	-1,1	1,-1	0,0

Table 4: stone, paper, scissors!



3 Towards Equilibrium

Consider a game where there are a large number of players (say a survey is being taken in your school). Every player is asked to choose an integer in the interval 1-64. The player whose choice is closest to half of the average wins. Ties are broken randomly.

Think about how you would reason if you were part of such a survey. If everyone picked their choice randomly then 32 would be the average. So you should pick 16, half of 32. But then you realise that everyone will reason similarly (or at least most of them) so the average will drop (ideally to 16) below 32. So you should pick something closer to 8. But then everyone will think the same . . . and so 1 is what everyone end up choosing and some random guy is the winner. But this is the best you could've done. This is an example of what is called a *Nash Equilibrium.*

This is the outcome you get when everyone plays rationally and responds the best he can to what he thinks the others will play. He tries to maximize his payoff assuming best responses from everyone else. Actually let's have a look at a couple of definitions.

Definition : Best Response

$a_i^* \in BR(a_{-i})$ iff $\forall a_i \in A_i$, $u_i(a_i^*, a_{-i}) \geq u_i(a_i, a_{-i})$. BR is the set of best responses, which is a function of a_{-i}, the action profile of every player except i. a_i^* is called a *best response.*

Definition : Nash Equilibrium

$a = (a_1, \ldots, a_n)$ is a Nash action profile, iff $\forall i \in \{1, \ldots, n\}$, $a_i \in BR(a_{-i})$.

That is to say it is the situation wherein everybody is best responding. A better way of understanding a Nash Equilibrium is that no one will want to change their choice of action, provided no one else changes. Let's examine the Nash Equilibrium in a few example games. Remember, the Nash Equilibrium doesn't really maximise payoffs always. It's just an equilibrium situation.

Coin machine

Observe your payoffs. If he decides to cheat, it is better for you to cheat as well ($0 > -1$). If he cooperates, it is still better for you to cheat ($3 > 2$). Both of you realise this and so the Nash Equilibrium is both of you cheating. Sad, isn't it? If only there was a bit of trust you both could have cooperated resulting in an equally better payoff for both (a win-win, i.e.).

Movie dilemma

(X,X) and (Y,Y) are the two Nash Equilibria. It should be easy to verify this.

There are no Nash Equilibria in the ideal penalty game and stones-paper-scissors.

Dominant strategies

Let a strategy mean "choice of action" for now. Let s_i and $s_i^{'}$ be two strategies. s_i *strictly dominates* $s_i^{'}$ if for every s_{-i} it gives a strictly greater payoff. If the inequality isn't strict, then we call it very weak dominance.

If one strategy dominates all others, we say it's a dominant strategy. The player will play that strategy regardless of what everyone else does. If every player has a strictly dominant strategy they'll play that and this situation will give us the unique Nash Equilibrium of the game. This was the case in the Coin Machine game (in a prisoner's dilemma, to be more general).

Pareto Optimality

This is kind of like an outsider's view of what the best outcome for a game would be. If an outcome O is at least as good as $O^{'}$ for everyone, but strictly better for at least one player, then we say O *pareto-dominates* $O^{'}$.

O^* is Pareto-optimal if there's no other outcome which pareto-dominates it. It doesn't have to pareto-dominate any outcome. It just must not be pareto-dominated by any other outcome. For example, consider a game where the outcomes are (3,0),(2,1),(1,2), and (0,3). No outcome pareto-dominates any other. All outcomes are pareto-optimal though. In fact, in zero-sum (or constant-sum) games, all outcomes are pareto-optimal.

There can be many pareto-optimal outcomes in a game, but every game has at least one because if not then every outcome is pareto-dominated by some outcome and that doesn't make sense.

Have a look at the coin machine game and try to figure out which outcomes are pareto-optimal. We've already looked at this game's Nash equilibrium. Now we understand why the prisoner's dilemma garners such fame. The paradox of the prisoner's dilemma is that the dominant strategy Nash Equilibrium is the only non pareto-optimal outcome!

4 Mix and Match

In many games (for example the penalty shootout) it is usually a bad idea to play a deterministic move because the the other player can use that to his advantage (which may lead to our loss). In such cases players can try and randomize or mix and match over the set of available actions. Let us now properly define what a strategy (s_i) is.

Definition : Strategy (s_i)

s_i is a probability distribution over the action space A_i.

The set of actions having a non-zero probability is called the *support* of the strategy. A *pure strategy* has a singleton set as the support. A *mixed strategy* has many actions in the support. Set of all s_i is called S_i (strategy space). $S = S_1 \times \ldots \times S_n$.

Utility in mixed strategies

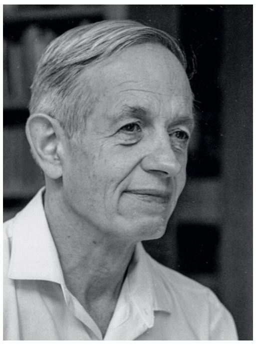
$$u_i = \sum_{a \in A} u_i(a) Pr(a|s)$$

$$Pr(a|s) = \prod_{j \in N} s_j(a_j)$$

where $A = A_1 \times \ldots \times A_n$ and $N = (1, \ldots, n)$.

We have now covered enough ground to be exposed to John Nash's famous existence theorem (1950).

Theorem. *Every finite game has a Nash Equilibrium.*



John Forbes Nash (1928-2015)

The proof of course is not in the scope of this report. The adventurous reader can find a detailed proof at `https://www.cse.iitk.ac.in/users/swaprava/courses/cs711/nash-proof.pdf`.

Computing Nash Equilibrium

The Nash Equilibrium we discussed is called a Pure Strategy Nash Equilibrium (PSNE) as it involves each players playing an action, i.e. a pure strategy. Nash's theorem refers to the generalised Nash Equilibrium which allows for mixed strategies to be played.

It's easy to compute the equilibrium when you can guess the support. Consider the movie dilemma. Because mixed strategies cannot involve a single action, and there are only two action available to the player, those two actions constitute their support.

Let player 2 play X with p, Y with $1-p$. If player 1 best-responds with a mixed strategy, player 2 must make him indifferent between X and Y. This is because if he preferred one over the other, then his mixed strategy would collapse into a pure strategy with the dominant action, which contradicts out assumption. So let us equate player 1's payoffs.

$$u_1(X) = u_1(Y) \Rightarrow 2p + 0 = 0 + (1-p)$$

$$\Rightarrow 3p = 1 \Rightarrow p = \frac{1}{3}.$$

Likewise, player 1 must randomize so as to make player 2 indifferent. Let him play X with a probability of q.

$$u_2(X) = u_2(Y) \Rightarrow q = 2(1-q)$$

$$\Rightarrow 3q = 2 \Rightarrow q = \frac{2}{3}.$$

Thus the mixed strategies $(\frac{1}{3}, \frac{2}{3}), (\frac{2}{3}, \frac{1}{3})$ constitute the mixed strategy Nash equilibrium of this game.

Interpreting mixed strategy equilibrium

What does it mean to play a mixed strategy?

- a mixed strategy describes what would happen in a repeated play
- it is played to confuse opponents
- it is played because of uncertainty about opponent's strategy.
- they can also be said to represent population dynamics.

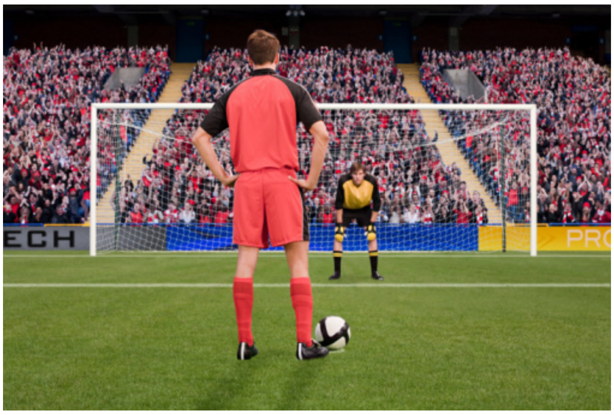
Computing Nash Equilibrium for bigger games

This is actually a very hard problem and let's not go into algorithmic complexity. Again, the adventurous reader can read up on the LCP formulation and the Support Enumeration method in a paper by Carnegie Mellon University at `https://www.cs.cmu.edu/~ggordon/780-spring09/slides/Game%20theory%20lecture2-algs%20for%20normal%20form.pdf`.

Exploring the penalty kick

We have seen the ideal penalty kick. The Nash Equilibrium for that is both of them playing (0.5,0.5). Let us now consider a case where the kicker has a weak right side shoot. By employing our standard method,

$$u_1(L) = u_1(R) \Rightarrow 1 - p = 0.75p$$



$$\Rightarrow 4 - 4p = 3p \Rightarrow p = \frac{4}{7}$$

Similarly,

$$q + 0.25(1 - q) = 1 - q \Rightarrow q = \frac{3}{7}$$

	left	right
left	0,1	1,0
right	0.75,0.25	0,1

Table 5: weak right shoot

Can you notice something weird? The kicker is actually kicking more to the right, his weaker side! Calculate the goalie's payoff in this case, it comes out to be $\frac{4}{7}$ (and because it's a constant-sum game, kicker's payoff is $\frac{3}{7}$). So the goalie *is* exploiting the kicker's weakness. How exactly though?

If the goalie played $(\frac{1}{2}, \frac{1}{2})$, the the kicker could play (1,0), i.e. always left, and get a payoff of 0.5 as opposed to $\frac{3}{7}$. By going left more often (changing his strategy to make kicker indifferent), the goalie forces the kicker to go right more often thereby taking advantage of the weak right shoot and emerging triumphant!

5 Beyond the Nash Equilibrium

Iterative removal of strictly dominated strategies

Reverse the inequality in the definition of strict dominance to arrive at the definition of a strictly dominated strategy.

No sane player will ever play a strictly dominated strategy, so we may as well remove it from the game (i.e. the matrix). Then see if you can find any other strictly dominated strategies (some new ones may have surfaced in the other players' strategy set). Iterate until you can no longer find a strictly dominated strategy.

Consider a game.

For player 2, C strictly dominates R (look at player 2's payoffs). So let's remove R.

	L	C	R
U	3,0	2,1	0,0
M	1,1	1,1	5,0
D	0,1	4,2	0,1

Table 6: original game 1

	L	C
U	3,0	2,1
M	1,1	1,1
D	0,1	4,2

Table 7: R has been removed

Now U strictly dominates M. Then C strictly dominates L. Finally D strictly dominates U. So we arrive at a unique Nash Equilibrium (D,C).

	L	C
U	3,0	2,1
D	0,1	4,2

Table 8: M has been removed

	C
U	2,1
D	4,2

Table 9: L has been removed

	C
D	4,2

Table 10: U has been removed

Consider another example.

	L	C	R
U	3,1	0,1	0,0
M	1,1	1,1	5,0
D	0,1	4,1	0,0

Table 11: original game

R is strictly dominated by L (C also for that matter). So we remove that.

Now M is not strictly dominated by either of U or D. But M is strictly dominated by the *mixed strategy* whose support is (U,D) and the strategy is (0.5,0.5)! Dominance can be exercised by mixed strategies too.

	L	C
U	3,1	0,1
M	1,1	1,1
D	0,1	4,1

Table 12: R has been removed

	L	C
U	3,1	0,1
D	0,1	4,1

Table 13: M has been removed

Now there aren't any more strictly dominated strategies. In fact, none of the outcomes is a PSNE now.

Iterative removal of strictly dominated strategies preserves Nash Equilibria. It can be used as a pre-processing step on the quest to find Nash Equilibria. Games completely solvable by this method are called *dominance solvable*. Order of removal doesn't matter here.

The pig food game

No, the pig isn't the food. There are two pigs in a cage, and one of them is stronger than the other. They need to press a lever for food to arrive but the food is placed in a corner of the cage opposite to the lever. So the pig which presses the lever has to run back to get the food and so it gets lesser food than it would have otherwise. Suppose 10 units of food are given. The game then goes like this. Player 1 is the weaker pig. PS : running consumes 2 units of food's worth energy.

	press	wait
press	1,5	-1,9
wait	4,4	0,0

Table 14: pig game

Notice that pressing is a strictly dominated strategy for the weaker pig. Therefore the Nash Equilibrium is (wait,press), resulting is (4,4). And studies very accurately match this prediction! Pigs don't know game theory, yet they learn and exhibit this behaviour. Evolution and learning are are powerful game theoretic tools indeed.

6 Enter Time : Extensive Form

The normal form which we've discussed so far can only represent games in which moves are simultaneous. It is pretty useful for identifying strictly dominated strategies and Nash Equilibria but to incorporate the sequencing of players' moves we turn to a more elaborate specification of a game.

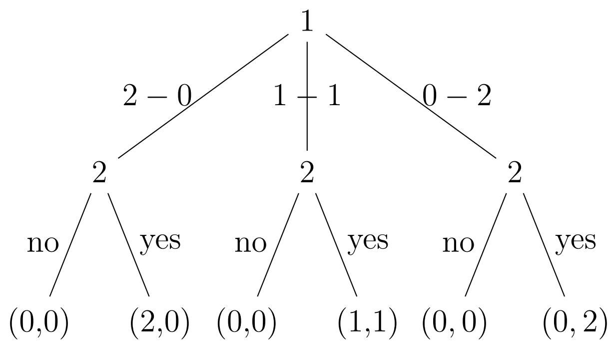
Defining Perfect Information Extensive Form games

It is defined by the tuple $(N, A, H, Z, \chi, \rho, \sigma, u)$.

- $N = (1, \ldots, n)$ is the set of n players.
- A is a single set containing all available actions.
- H is the set of all non-terminal nodes, i.e. choice nodes.
- $\chi : H \mapsto \mathcal{P}(A)$ is the action function. It assigns to each choice node a subset of A, i.e. a set of possible actions.
- $\rho : H \mapsto N$ is the player function. It assigns to each choice node a player.
- Z is the set of terminal nodes.
- $\sigma : H \times A \mapsto H \cup Z$ is the successor function. It maps a choice node and an action to a new choice node or terminal node. It is an injective (one-one) function, which means a node cannot be a successor to multiple nodes.
- $u = (u_1, \ldots u_n)$, $u_i : Z \longmapsto \mathbf{R}$. This is the utility function which assigns a utility to each terminal node.

This definition of extensive form games very naturally gives way to a neat representation using trees. Let us look at an example.

The sharing game



Two siblings are fighting over how to share two toffees they have. Their parents intervene and suggest the following. The older will make an offer which the younger one can either accept or reject. If he rejects, then their parents take away both the toffees. If he accepts then the two chocolates are shared accordingly.

Pure strategies

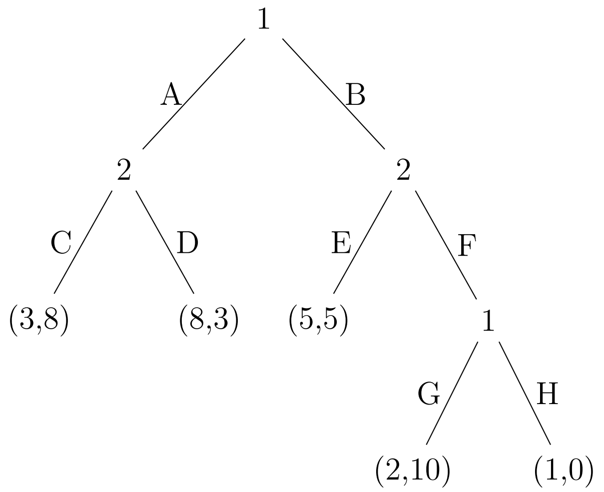
How many pure strategies do you think player 1 has? The answer is 3, which makes sense as he has three choices at his only choice node. What about player 2? Player 2 has 2 choices, "yes" or "no". But the number of pure strategies he has is 8. This is because ...

A pure strategy in a perfect information game is a complete specification of which action to take at which choice node. The set of pure strategies of player i is S, where

$$S = \prod_{h \in H, \rho(h)=i} \chi(h)$$

So every strategy of player 2 will require him to specify whether to say yes or no at each of his choice nodes. That is how we get $2 \times 2 \times 2 = 8$ pure strategies.

An example game



Let us look at the pure strategies of each player.

$$S_2 = \{(C, E), (C, F), (D, E), (D, F)\}$$

$$S_1 = \{(A, G), (A, H), (B, G), (B, F)\}$$

This may feel a bit weird. If player 1 plays A, then he will never reach his second choice node. But it follows from the definition ($\{A, B\} \times \{G, H\}$). Given our current understanding of pure strategies we can reuse our definitions of mixed strategies, best response and Nash Equilibrium.

Theorem. *Every perfect information game in extensive form has a PSNE.*

This makes sense as play is sequential not simultaneous here. So there is no advantage to be had by playing a mixed strategy (no element of surprise).

Induced Normal form

Every extensive form game can uniquely be converted into a normal form game. This is useful as it is easier to find all the Nash Equilibria by looking at the normal form.

But this transformation shows the lack of compactness of the induced normal form. We got 16 payoff pairs instead of 5. This ratio grows fast as the tree grows in size. Also, Normal form games can't always be converted into extensive form.

What are the PSNE of this game?

	CE	CF	DE	DF
AG	3,8	3,8	8,3	8,3
AH	3,8	3,8	8,3	8,3
BG	5,5	2,10	5,5	2,10
BH	5,5	1,0	5,5	1,0

Table 15: the induced normal form

1. ((A,G),(C,F))
2. ((A,H),(C,F))
3. ((B,H),(C,E))

You can verify using the normal form.

Subgame Perfection

There is something intuitively wrong with the Nash Equilibrium ((B,H),(C,E)) in our example game. G dominates H for him at his second choice node. So why does he choose H? Because if he said G, then player 2 would prefer F over E, which is bad for him. But this is a hollow threat, right? If player 2 actually played F, player 1 would pick G because his real intention (as a rational player anyways) is to maximise his payoff, not minimizing his opponent's or maintaining a lead.

It is at this point we discuss a special type of equilibrium. The restriction of a game to descendants of a particular choice node is called a *subgame*. An equilibrium is called *subgame perfect* if it is a Nash Equilibrium in every subgame of the game.

Thus, non-credible threats are ruled out. Subgame Perfect Equilibrium is therefore a refinement of Nash Equilibrium.

- ((A,G),(C,F)) is subgame perfect.
- ((B,H),(C,E)) is not.
- ((A,H),(C,F)) is also not, despite H being "off path".

Backward Induction

The idea is to identify the equilibria in bottom-most trees and adopt these while moving up the tree. This can be seen as extending the utility function to non-terminal nodes. It is an idea which can be coded pretty easily. But if you aren't familiar with programming don't worry. You don't need to understand the following pseudocode to grasp what it is.

In zero-sum games, backward induction is known as the minimax algorithm. Further, there exists an idea of 'pruning' nodes that'll never be reached in play called *Alpha-beta pruning.*. This is used commonly for machine playing of two-player games (Tic-tac-toe, Chess, Go, etc.).

function BACKWARDINDUCTION (node h) **returns** $u(h)$
if $h \in Z$ **then**
  return $u(h)$
$best\_util \leftarrow -\infty$
forall $a \in \chi(h)$ **do**
  $util_at_child \leftarrow$BACKWARDINDUCTION$(\sigma(h, a))$
  if $util\_at\_child_{\rho(h)} > best\_util_{\rho(h)}$ **then**
    $best\_util \leftarrow util\_at\_child$
return $best\_util$

The subtraction game

There is a pile of n stones. Each player can pick 1,2, or 4 stones from the pile in one turn. The player who plays the last move wins the game. For which values of n will the player playing first win?

This game is a special case of the celebrated game of Nim which has excellent literature available and is an extensively discussed topic in combinatorial game theory. It's high point is the Sprague-Grundy theorem which sort of solves every perfect information impartial game.

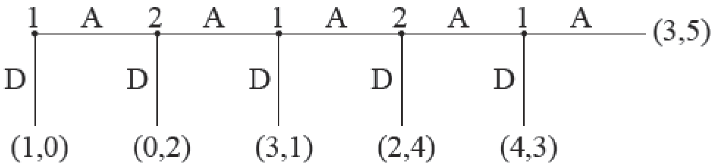
As for this particular game, this can be solved just by backward induction. Let us call the player playing first P and the other player (opponent) O. If $n = 1$ then P picks the stone up and wins. If $n = 2$ then P can pick 1 but then he'll lose so he picks up 2 and wins. This is a core idea. Even if a player has one choice which'll lead him to winning the game or taking the game into a situation advantageous to him he'll take it and win. If $n = 3$ P has to pick either 1 or 2 and then O will pick up either 2 or 1 (respectively) and win the game. If $n = 4$ then P will pick 'em all up and win. For $n = 5$, if P picks up 1 or 4 the game is reduced to $n = 4$ or $n = 1$ with the opponent as P and so the opponent wins. But if P picks 2 the game is now $n = 3$ for O hence P wins for $n = 5$. $n = 6$ means P loses.

And now modular arithmetic comes to our rescue. If $n = 3k + 1$ or $3k + 2$ then P can take the game to $3k$ and then keep it at a multiple of 3 by responding with 2 to 4, with 1 or 4 to 2 and with 2 to 1. And we've seen how facing 3 or 6 is a losing strategy. So P wins in these cases. If on the other hand $n = 3k$ we use the same argument on O, and observe whatever P does O always has a reply which takes him to victory.

There, we've solved (this particular version of) the subtraction game completely. Think about how you'd solve if players were allowed to pick 1, 3 or 4 stones. What about 1,2 or 5?

The centipede game

If both players keep playing A then they would reach the pareto-optimal (3,5). But player 1 gets a higher payoff by playing D on his last move. Player 2 knows this and observes he's better off by playing D in *his* last move. Continuing on this backward induction we find the only subgame perfect equilibrium is (1,0) which is ironically pareto-dominated by all other outcomes except one ((0,2)).



But several studies have shown that this is rarely ever observed. Participants regularly show partial cooperation (playing A for a while) before one of them defects.

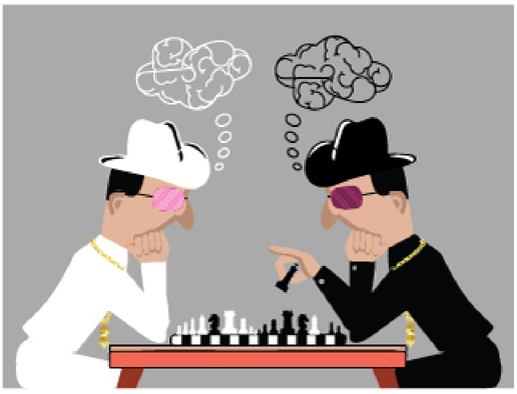
Explanations

Nagel and Tang (1998) suggest altruism as a reason. The basic idea is that if you are playing against an altruist to maximize your payoff you should defect on the last round. If enough people are altruists, sacrificing the payoff of first-round defection is worth the price in order to determine whether or not your opponent is an altruist.

Another possibility involves error. If there is a significant possibility of error in action, perhaps because your opponent has not reasoned completely through the backward induction, it may be advantageous (and rational) to cooperate in the initial rounds.

However, Parco, Rapoport and Stein (2002) illustrated that the larger the incentives are for deviation, the greater propensity for learning behavior in a repeated single-play experimental design to move toward the Nash equilibrium.

Palacios-Huerta and Volij (2009) find that expert chess players play differently from college students. With a rising Elo, the probability of continuing the game declines; all Grandmasters in the experiment stopped at their first chance. They conclude that chess players are familiar with using backward induction reasoning and hence need less learning to reach the equilibrium.

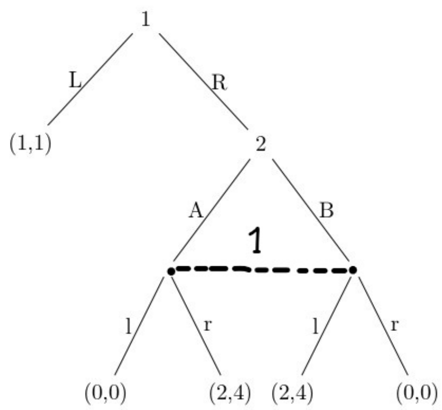


Significance

Like the Prisoner's Dilemma, this game presents a conflict between self-interest and mutual benefit. If it could be enforced, both players would prefer that they both cooperate throughout the entire game. However, a player's self-interest or players' distrust can interfere and create a situation where both do worse than if they had blindly cooperated. Although the Prisoner's Dilemma has received substantial attention for this fact, the Centipede Game has received relatively less.

7 Enter Confusion : Imperfect Information Games

So far, we've assumed that players in an extensive-form game always know what node they're at and know all prior choices, both theirs and the others'. But sometimes players don't know all the actions the others took or don't recall all their past actions.



To deal with such games we create partitions of a player's choice nodes called equivalence relations or information sets. What do we mean by a partition? Consider a set $S = \{1, 2, 3, 4, 5, 6\}$. $\{\{1, 2\}, \{3, 5, 6\}, \{4\}\}$ is an example of a partition of the same. Now suppose these were how the 6 choice nodes of a player in some game were numbered. Then if the aforementioned partition were to be one of his equivalence relations, then he would have no way of telling nodes 3,5, and 6 apart as they're in the same *equivalence class.*

Put more formally, we augment a perfect information game with $I = (I_1, \ldots, I_n)$ where $I_i = \{I_{i1}, \ldots, I_{ik_i}\}$ is a partition of $\{h \in H : \rho(h) = i\}$ with the property that $\chi(h) = \chi(h^{'})$ whenever there exists a j such that $h \in I_{ij}$ and $h^{'} \in I_{ij}$. Which means for any two choice nodes in the same set created after the partition, the same set of actions available should be the same. Otherwise the player will see a difference.

The set of pure strategies now is

$$S = \prod_{I_{ij} \in I_i} \chi(I_{ij})$$

Essentially we're taking a cartesian product of the action set of all equivalence classes of a player. Here $I_1 = \{I_{11}, I_{12}\}$ and $I2 = \{I_{21}\}$. Player 1's pure strategies are Ll,Lr,Rl,Rr. Player 2's pure strategies are just A or B.

Now if you think about it, we can represent every normal form game as an imperfect information game. The induced normal form is made using exactly the same method as before. Further, due to Nash's theorem every imperfect information extensive form game has a Nash Equilibrium.

Mixed and Behavioral strategies

Till now we have looked at one way of randomizing and that was the concept of a mixed strategy. That involved randomizing over the set of pure strategies, in other words, having a probability distribution over the set of pure strategies.

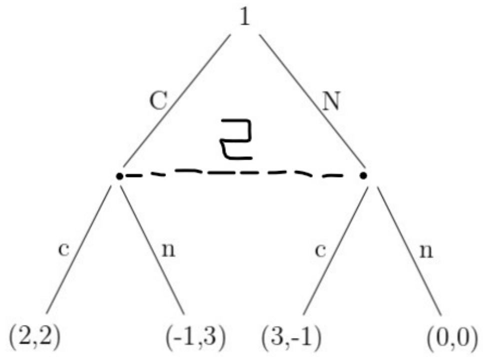


Figure 1: The coin machine game in extensive form (n stands for not cooperate, i.e. cheat)

But we can also randomize at every action node (or information set which is basically a collection of action nodes). Think of this as an independent coin toss every time we have to make a decision.

Consider the example game from the previous section. The set of pure strategies of player 1 was $\{AG, AH, BG.BH\}$. An example of a mixed strategy would be $s = \{0, 0, 0.6, 0.4\}$. Whereas an example of a behavioral strategy would be playing A with probability 0.5 at the first choice node and then playing G with probability 0.4 at the second choice node. This behavioral strategy can be seen as a mixed strategy as $\{0.2, 0.3, 0.2, 0.3\}$. In fact, all behavioral strategies can be represented as mixed strategies. The vice versa, however is not true.

8 Repeated Games

Repeated games can be found everywhere. In political alliances to firms in the marketplace to workers helping each other out.

Utility in repeated games

Given a sequence of payoffs $r_1, r_2, \ldots, r_k$, the average payoff is literally the average payoff, $(\sum_{j=1}^{k} r_j)/k$. In infinitely repeated games we just take the limit of this as $k \to \infty$. This may not always be well defined but let us keep that aside.

There is an alternate definition of utility called the future discounted utility, which is defined as

$$\lim_{k\to\infty} \sum_{j=1}^{k} \beta^j r_j$$

where $0 < \beta < 1$ is the discount factor. The idea here is that people care more about near term well being than some obscure future. Also if we made the game finite then this resembles the calculation in an extensive form game if β is viewed as the probability of the game carrying on at any point.

Stochastic games

These types of games are a generalisation of a repeated game except after every "move" (i.e. after every player picks an action) the game changes state (in the sense the actions and payoffs change). More formally we have $G = \{G_1, G_2, \ldots\}$ as a set of games, each having the same set of players and players jump between games depending on their decisions. The 1-player versionof this is called a Markov Decision Process (MDP). MDPs are useful for studying optimization problems solved via dynamic programming and reinforcement learning.

Fictitious play

This is a learning rule wherein each player initialises some beliefs about his opponents. At each turn he plays a best response to this assumption which he then updates based on his observations of the opponents' moves.

Formally, he maintains count of his opponent's actions ($w(a); a \in A$) and assesses opponent's strategy $\sigma(a)$ using the counts. Basically he calculates the relative frequency with which his opponent played each action so far. Then he best responds to this assessed strategy σ.

$$\sigma(a) = \frac{w(a)}{\sum_{a' \in A} w(a')}$$

9 A Game of Trust

A Christmas Story

It was Christmas 1914 on the Western Front. Despite being involved in a *world war*, British and German soldiers left their trenches and crossed No Man's Land to mingle with each other and exchange food and souvenirs. There were joint burial ceremonies and prisoner swaps, while several meetings ended in carol-singing. Men played games of football with one another giving one of the most memorable images of the truce.

How was this possible? Game theory might give us some insight on why this was dramatic, but neither unique, nor unusual.

The setting of the game

Recall the coin machine game (a specific case of the more general Prisoner's Dilemma). Let us construct a repeated game from it. It will be finite but the players will not know when it is ending.

There are some famous strategies for (practically) infinite games. Here are five of them.

1. COPYCAT (blue): Hello! I start with Cooperate, and afterwards, I just copy whatever you did in the last round.

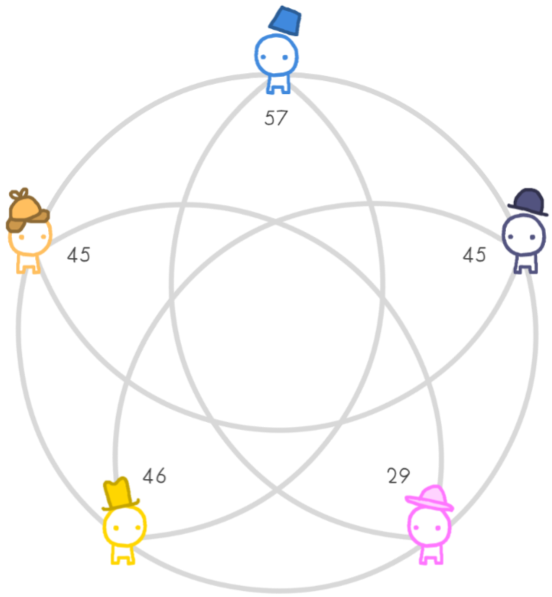
2. ALWAYS CHEAT (violet) : *The strong shall eat the weak.*

3. ALWAYS COOPERATE (pink) : Let's be best friends!

4. GRUDGER (yellow) : I'll start cooperating, and keep cooperating, but if you ever cheat me, I'll cheat you back forever.

5. DETECTIVE (orange) : First: I analyze you. I start: Cooperate, Cheat, Cooperate, Cooperate. If you cheat back, I'll act like Copycat. If you never cheat back, I'll act like Always Cheat, to exploit you.

Suppose there were to be a tournament between these five players, each face-off involving 10 games (the players won't know this). Who'd you place your bets on? Think about this for a moment.



The results

It turns out to be Copycat. Copycat goes by many names. The Golden Rule, reciprocal altruism, tit for tat, or ... live and let live.

Now, let's let our population of players evolve over time. It's a 3-step process.

1. Play a tournament :
   Let them all play against each other and tally up ther scores.

2. Eliminate losers :
   Get rid of the 5 worst players. (if there's a tie, pick randomly between them)

3. 3. Reproduce winner :
   Clone the 5 best players. (if there's a tie, pick randomly between them)

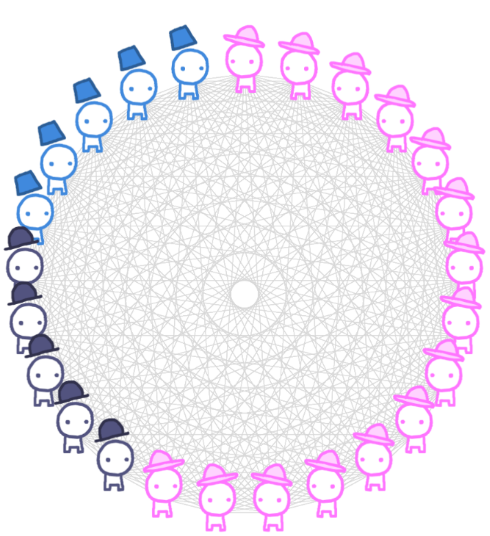
Note: you don't have to wait for people to literally die and reproduce for culture to evolve - all that's needed is that "unsuccessful" behaviors go away, and "successful" behaviors are imitated.

Say we start with the following population of players: 15 Always Cooperates, 5 Always Cheats, and 5 Copycats (We'll ignore Grudger and Detective for now). We're going to follow the three steps of evolution a dozen times or so. Let's make another bet! Who do you think will win the first tournament?

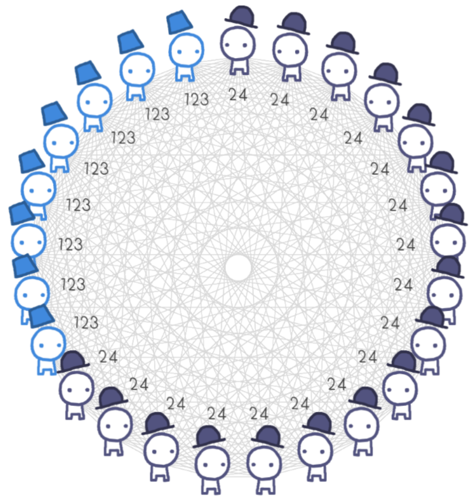
On running the simulation we see that initially the Always Cheats take advantage of the huge number of Always cooperate and end up gaining in numbers until all the Always Cooperators have been squashed off, leaving the equation as 5 Copycats and 20 Always Cheats.

But then the Always Cheats become a victim of their own success! They exploited the naive Always Cooperators, but once they ran out of them, they had to face the Copycats: who are nice, but not naive. By simply copying the other player's moves, Copycats can play nice with each other, while Always Cheats just cheat themselves! Not only that, but it also means Copycat can give Always Cheat a taste of their own medicine.

So in the long run Copycat wins! Always Cheat may have won in the short run, but its exploitativeness was its downfall.



Initial situation



The first round after the Always Cooperators have all gone

And the result is similar even if we put Grudger and Detective back in. So, it seems the math of game theory is telling us something: that Copycat's philosophy, "Do unto others as you would have them do unto you", may be not just a moral truth, but also a mathematical truth.

The evolution of Distrust

However ... there is one parameter which we've assumed to be fixed so far changing which makes a huge difference to who wins. It is the number of interactions (rounds) between two opponents in a tournament. That number (n) is currently 10. On toggling that number we observe that if you don't have enough rounds, (here: 5 or less) Always Cheat dominates.

There is yet another way to breed distrust. By changing the payoffs of the game. Suppose we change the *both cooperate* payoff to +1 instead of +2. Even though +1 is more than the punishment for *both cheat* (0) Always Cheat emerges victorious in this setting. The same thing happens when we keep *both cooperate* at 2 but change *both cheat* to -1. With a lower "win-win" reward, Always Cheat takes over. This is due to two powerful ideas Game Theory has about this.

Zero-sum game : This is the sadly common belief that a gain for "us" must come at a loss to "them", and vice versa.

Non-zero-sum game : This is when people make the hard effort to create a win-win solution! (or at least, avoid a lose-lose) Without the non-zero-sum game, trust cannot evolve.

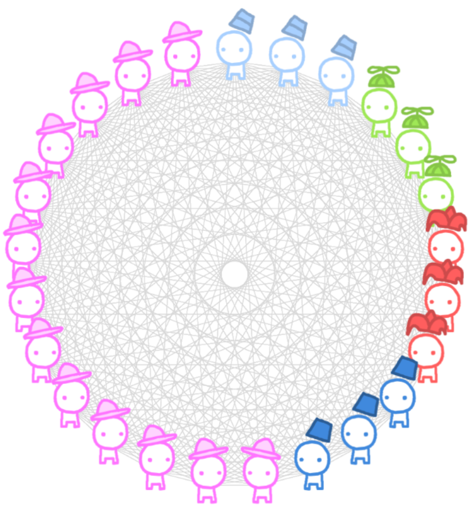
~~Mistaekes~~ Mistakes

There is yet another huge barrier to trust. Consider two copycats playing each other. They start off by cooperating and will continue to do so till the end of time. However suppose one of them fell on his way to the coin machine. Mistakes, miscommunication, misinterpretations - accidents happen all the time in real life. But if the other person doesn't think it was an accident, being a copycat he has to retaliate. Thus these two Copycats will spiral into an endless cycle of vengeance ... that started over a single mistake, long ago. Tragic. But there are some characters which can deal with mistakes. Introducing ...

1. COPYKITTEN (light blue) : I'm like Copycat, except I Cheat back only after you Cheat me twice in a row. After all, the first one could be a mistake!

2. SIMPLETON (green) : I start by cooperating. If you cooperate back, I'll do the same thing as my last move, even if it was a mistake. If you cheat back, I'll do opposite thing as my last move, even if it was a mistake.

3. RANDOM (red) : (Just plays Cheat or Cooperate randomly with a 50/50 chance)

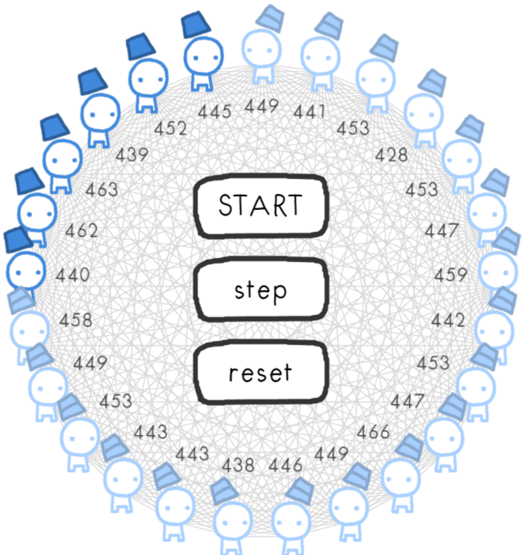


Let's start with a dozen Always Cooperates, versus our old winner, the fair Copycat, and our three new characters: the forgiving Copykitten, the dull Simpleton, and the silly Random.

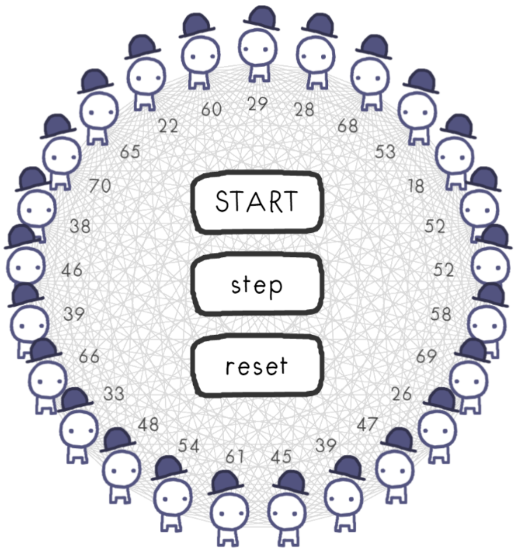
In each round of a match, players have a small chance of making a mistake. (let's say, 5%) Who do you think will come out on top? It turns out to be Simpleton! This is because Simpleton is actually capable of exploiting Always Cooperate. They both start cooperating, but if Simpleton makes a mistake and cheats, since Always Cooperate never retaliates, it'll keep cheating them.

Now let's try the the same thing as before, except instead of half-Always Cooperate, it's half-Always Cheat. It's a much less forgiving, more hostile environment. Let's play! Copykitten wins this time! That's surprising that with an even meaner starting population, Copykitten, a more forgiving version of Copycat, was the most successful! (note: Copykitten is so forgiving it doesn't even entirely wipe out Copycat. it shares room).

In this case, a bit of "miscommunication" (5% chance of mistake each round) could lead to more forgiveness. But is this true for all levels of miscommunication? Let us try toggling the chance of mistake and see what happens. The results turn out something like this: At 0%, the fair Copycat wins! From 1% to 9%, the forgiving Copykitten wins! From 10% to 49%: the unfair, unforgiving Always Cheat wins. At 50%, nobody wins ever.



5% miscommunicaton results



10% miscommunication results

This is why "miscommunication" is such an interesting barrier to trust: a little bit of it leads to forgiveness, but too much and it leads to widespread distrust! It is plausible that our modern media technology, as much as it's helped us increase communication... has increased our miscommunication much more.

Conclusion

Game theory has shown us three things we need for the evolution of trust:

1. REPEAT INTERACTIONS : Trust keeps a relationship going, but you need the knowledge of possible future repeat interactions before trust can evolve.
2. POSSIBLE WIN-WINS : You must be playing a non-zero-sum game, a game where it's at least possible that both players can be better off – a win-win.
3. LOW MISCOMMUNICATION : If the level of miscommunication is too high, trust breaks down. But when there's a little bit of miscommunication, it pays to be more forgiving

Of course, real-world trust is affected by much more than this. There's reputation, shared values, contracts, cultural markers, blah blah blah. But if there's one big takeaway from all of game theory, it's this:

What the game is, defines what the players do. Our problem today isn't just that people are losing trust, it's that our environment acts against the evolution of trust.

That may seem cynical or naive, that we're "merely" products of our environment. But as game theory reminds us, we are each others' environment. In the short run, the game defines the players. But in the long run, it's us players who define the game. So, do what you can do, to create the conditions necessary to evolve trust. Build relationships. Find win-wins. Communicate clearly. Maybe then, we can stop firing at each other, get out of our own trenches, cross No Man's Land to come together.

Footnotes

The Christmas Truce

Not every trench joined in the peace, but it was pretty widespread. Many front-lines came up with the idea independently, and again: despite specific, strict orders not to. And in fact, even before Christmas, several front-lines already had established an unofficial, secret peace. They called it: the "live and let live" system. Basically, you don't shoot me, I don't shoot you. And this worked, in a lot of places! Still, most soldiers don't spontaneously form peace with the enemy. What's so special about trench warfare? Well, here's what's unique about the trenches: unlike almost every other form of war, you have to face the same specific soldiers every day. It's a repeated game. And that makes all the difference.

Not knowing when the game ends

In the repeated game of trust (also known as Iterated Prisoner's Dilemma), it's important that neither player knows when the last round is. Why? Think about it - on the last round, both players would know their action has no consequence, so they'd both cheat. But that means in the second-last round, their actions can't change the next round, so they'd also both cheat. But that means in the third-last round... etc etc.

Copycat

This strategy is better known in game theory as Tit For Tat. It was created by Anatol Rapoport in 1980, for Robert Axelrod's game theory tournament. "Tit For Tat" was not used here because it sounds mean, although it's a nice and fair strategy.

Copykitten

Just like how Copycat's original name was Tit For Tat, Copykitten's original name is Tit For Two Tats. Same rule: Cooperate, unless the other players cheats twice in a row. There's another forgiving variant of Tit For Tat called Generous Tit For Tat. It's got a similar but slightly different rule: Cooperate, but when the other player cheats, forgive them with a X% chance. This design, with the variable "X", lets you set different "forgiveness" levels for the player.

Simpleton

The learning rule bases its decision only on the outcome of the previous play. Outcomes are divided into successes (wins) and failures (losses). If the play on the previous round resulted in a success, then the agent plays the same strategy on the next round. Alternatively, if the play resulted in a failure the agent switches to another action.

A large-scale empirical study of players of the game rock, paper, scissors shows that a variation of this strategy is adopted by real-world players of the game, instead of the Nash equilibrium strategy of choosing entirely at random between the three options.

10 Bayesian Games

Auctions are everywhere. From fish markets in Tokyo to auctionig seized horses by the US Marshalls. From IPL auctions for cricketers to Ebay auctions for "Virgin Mary in a grilled cheese sandwich" (seriously, look this up) to silent auctions, which closely resemble a game in the form we know it. Auctions are basically games where a player is unsure about other players' utility.

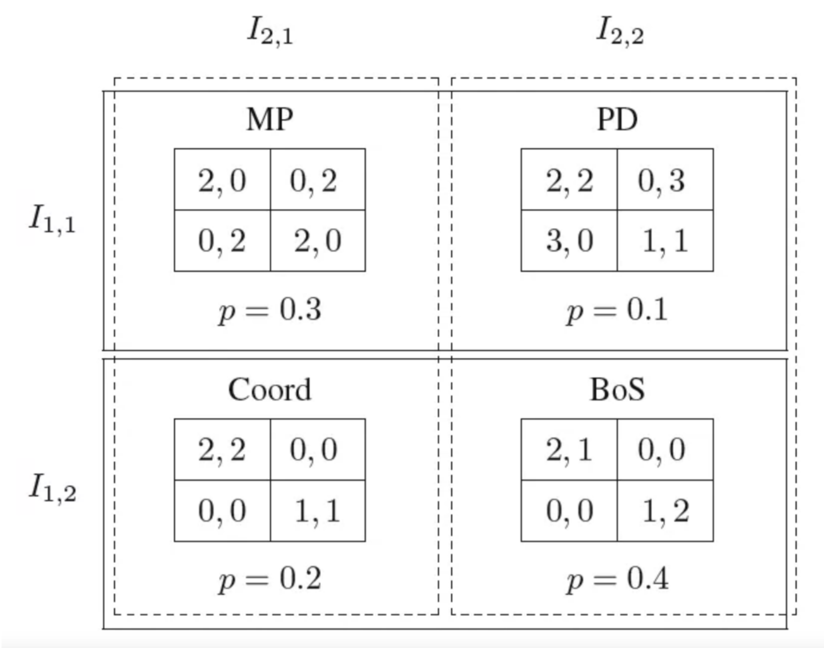
So far, we've assumed that each player knows the number of players (N), actions available to each player (A), and payoff associated with each action vector $u(a)$. Now let's relax this a bit. We shall have multiple games, each having the same set of players and action spaces but different payoffs. Players have beliefs over these payoffs which are modified based on individual private signals, which are again actually just information sets (partitions).

Definition 1

It is defined by the tuple (N, G, P, I).

- N is the set of players
- G is a set of games. Recall that a game g is defined as a tuple (N, A, u). If $g, g^{'} \in G$ then for every player strategy space in g is same as that in $g^{'}$.
- $P \in \prod(G)$ is a common prior where $\prod(G)$ is the set of all probability distributions over G.
- $I = (I_1, I_2, \ldots, I_n)$ is a set of partitions (information sets) of G, one for each agent.

Example game



Here we have four sub-games (not to be confused with the subgame of the extensive form) and players having two information sets each. The probabilities $(0.3, 0.1, 0.2, 0.4)$ constitute the *common prior*. Action set for player 1 is (top,bottom) and that for player 2 is (left,right).

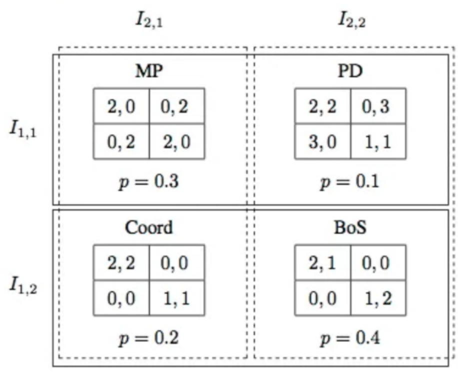
Nature decides which game will be played. Let's say the game happens to be BoS, the most likely one. Now each player is made aware of their types. So Player 1 knows they're in $I_{1,2}$ but doesn't know if the game is BoS or Coord. Similarly player 2 knows the game is either one of BoS or PD because that's what his information set ($I_{2,2}$) tells him.

Definition 2

This definition is mathematically identical to the previous ones but represents uncertainty over utility function using the notion of epistemic *types*.

It is a tuple (N, A, Θ, P, u) where

- N is the set of players
- A is the set of action sets
- $\Theta = (\Theta_1, \Theta_2, \ldots, \Theta_n)$ is the set of typespaces. A typespace is pretty similar to an information set except here it decides payoffs not games. A typespcae of a player contains typs which tell narrow down on the set of possible payoffs for him as well as other players. It acts like a private signal.
- $P : \Theta \mapsto [0.1]$ is the common prior, a probability distribution over typespaces.
- $u_i : A \times \Theta \mapsto \mathbf{R}$ is the utility function



a_1	a_2	θ_1	θ_2	u_1	u_2
U	L	$\theta_{1,1}$	$\theta_{2,1}$	2	0
U	L	$\theta_{1,1}$	$\theta_{2,2}$	2	2
U	L	$\theta_{1,2}$	$\theta_{2,1}$	2	2
U	L	$\theta_{1,2}$	$\theta_{2,2}$	2	1
U	R	$\theta_{1,1}$	$\theta_{2,1}$	0	2
U	R	$\theta_{1,1}$	$\theta_{2,2}$	0	3
U	R	$\theta_{1,2}$	$\theta_{2,1}$	0	0
U	R	$\theta_{1,2}$	$\theta_{2,2}$	0	0

a_1	a_2	θ_1	θ_2	u_1	u_2
D	L	$\theta_{1,1}$	$\theta_{2,1}$	0	2
D	L	$\theta_{1,1}$	$\theta_{2,2}$	3	0
D	L	$\theta_{1,2}$	$\theta_{2,1}$	0	0
D	L	$\theta_{1,2}$	$\theta_{2,2}$	0	0
D	R	$\theta_{1,1}$	$\theta_{2,1}$	2	0
D	R	$\theta_{1,1}$	$\theta_{2,2}$	1	1
D	R	$\theta_{1,2}$	$\theta_{2,1}$	1	1
D	R	$\theta_{1,2}$	$\theta_{2,2}$	1	2

Types deciding payoffs in the original example game

Strategy and expected utility

Given a bayesian game (N, A, Θ, P, u) with finite sets of players, actions, and types a pure strategy is defined as

$s_i : \Theta_i \mapsto A_i$:- a choice of pure action for every type of the player.

A mixed strategy is $s_i : \theta_i \mapsto \prod(A_i)$, ($\prod$ stands for 'probability distribution over'). $s_i(a_i|\theta_i)$ is the notation used for probability of action a_i in strategy s_i, given type θ_i.

There are three types of utilities in bayesian games.

1. ex-ante : the player knows nothing about anyone's type, not even his own.
2. interim : the player knows something about his own type, but nothing about others'.
3. ex-post : the player knows everything about everyone's types. This case actually boils down to a complete information game.

Now there are expressions used to evaluate utilities in these cases but I didn't understand them well so I'll just leave put them down here, if you're interested you can look them up.

- interim expected utility

$$EU_i(s|\theta_i) = \sum_{\theta_i \in \Theta_i} p(\theta_i|\theta_{-i}) \sum_{a \in A} \big(\prod_{j \in N} s_j(a_j|\theta_j) \big) u_i(a, \theta_i, \theta_{-i})$$

- ex-ante expected utility

$$EU_i(s) = \sum_{\theta_i \in \Theta_i} p(\theta_i|\theta_{-i}) EU_i(s|\theta_i)$$

The sheriff game

A sheriff is faced with an armed suspect and they must each simultaneously decide whether to shoot the other or not. The suspect is guilty with probability p and innocent with $1 - p$. The sheriff would rather not shoot, unless the suspect shoots him.

If the suspect were a criminal he'd rather shoot (irrespective of the sheriff's choice) so as to escape. But if the suspect were a civilian he'd rather not shoot, even if shot at. He may be too stunned to react or he might not wanna have blood on his hands (red on his ledger).

We can model this whole situation using two games (i.e. one bayesian game). You can have a nice look at the tables and make sense of the payoffs keeping the above discussion in mind. Player 1 is the suspect and player 2 is the sheriff.

good	shoot	not shoot
shoot	-3,-1	-1,-2
not shoot	-2,-1	0,0

bad	shoot	not shoot
shoot	0,0	2,-2
not shoot	-2,-1	-1,-1

There are two types available to the suspect. θ_G for good, i.e. innocent and θ_B for bad, i.e. guilty. So $\Theta_1 = \theta_{1,1}, \theta_{1,2}$. $\Theta_2 = \theta_{2,1}$ as the sheriff has only one type. In this game the suspect will know his type. If hes bad, shooting is a strictly dominant strategy for him and if he's good not shooting is strictly dominant. Knowing this, the sheriff figures out what'd be best for him. So if he were to shoot, he'd get a payoff of $p \times (0) + (1-p) \times (-1)$. If he didn't shoot he'd get a payoff of $p \times (-2) + (1-p) \times 0$. When is not shooting better?

$$-2p > -1(1-p) \Rightarrow 2p < 1-p$$
$$\Rightarrow 3p < 1 \Rightarrow p < \frac{1}{3}$$

So even if there's a one-third chance of the suspect being a criminal, it is better for the sheriff to shoot.

11 Conclusion

I'll use this section to summarize what we saw in this short summer trip to Game Theory land (or maybe it should be through Game Theory land).

We started out by defining the normal form and browsed through some example games. Then we saw an example where reasoning a bit gave us an unexpected result which we saw was what a Nash Equilibrium is. Then we got down to defining *Best Response* and *Nash Equilibrium* properly and applied them to the examples seen before. Two more ideas were understood, namely those of *dominant strategies* and *pareto optimality*.

Next we saw mixed strategies and Nash's famous theorem (whose proof was beyond us). Exploring the penalty kick led to some very interesting results of how the goalie took advantage of the kicker's weak foot in a kinda counter-intuitive way. Then we iteratively removed dominant strategies and saw how pigs know *game theory for dummies*.

Then we incorporated sequential movement into our games by turning to a different representation of games and saw how subgame perfection and backward induction were easy but powerful ways of solving games. Two interesting example games were understood before incorporating confusion into our games via equivalence relations and seeing how now we could freely interchange between normal form and extensive form (though not uniquely so). We also found about a new way of randomizing before moving on the short exposition on repeated games

And then we came to my favorite part of the report, The Game of Trust. It packs in lots of useful little things and ends on a moral note. It has been summarized then and there so I don't think there's a need to say anything here. And then finally we had a peek into two ways bayesian games can be defined and looked at an interesting sheriff dilemma.

12 Bibliography

- Coursera course on game theory offered by Stanford University and The University of British Columbia
- Nicky Case's *The Evolution of Trust* at https://ncase.me/trust
- Wikipedia, the holy grail of knowledge