ECE495 Game Theory and Networked Systems

Lecture 10: Fictitious Play and Dynamic Game

1 Level k Behavior

In real life, due to limited cognitive ability, people either don't recognize the setting of strategic game, or do not actively calculate the best responses (or eliminations), i.e, not fully rational. An alternatively way of modeling human behavior in lieu of Nash equilibrium is to classify people based on how much effort they put in to analyzing the strategic setting. On the high level, we can define iteratively different level of rationality in the following:

1. Player of level 0 plays a random strategy.
2. Player of level 1 assumes that other players are level 0, and best responds.
3. Player of level 2 assumes that other players are level 1, and best responds.
   ...
4. When k approaches ∞, it is the Nash Equilibrium

In real life, most people are of levels 0, 1 or 2. This is partly due to lack of computation resources. In these settings, Nash Equilibrium may not be the best representation of the real life scenario.

2 More on Behavioral Games

Consider the following hunting game of Stag or Hare. In this example, a stag can't be captured by one player, so if only one player aims for stag and the other aims for hare, the player for stag will get nothing. The hare on the other hand can be captured by either one or two players. In case without communication, risk averse players will go for hare, since they have a guaranteed payoff of 7 regardless of what the other player does. This tendency makes (H, H) the more likely outcome, although there are 2 Nash Equilibria (H, H) and (S, S).

		P_2	
		S	H
P_1	S	(9,9)	(0,8)
	H	(8,0)	(7,7)

Table 1: The stag and hare example

We note that the prediction changes when one player can message. For example, player 1 tells player 2 to catch stag together. In this case, the player 1 will (or perhaps) choose stag. In real life, the choice of player 1 depends on the relationship between players. If they are cooperating for long term, player 1 will choose stag. If they are short term or they don't trust each other, player 1 will likely to choose hare. Perception and risk taste of individual will dominate how games are played in reality. More details on this topic can be found at here.

3 Fictitious Play

We next give a dynamic rule of iterative best response called *fictitious play*, where the game is played many times and each player counts the opponent's strategy distribution and best responds to the empirical distribution of strategies. We illustrate this through the following example.

Suppose we are given the payoff matrix as in Table 2:

		$Player_2$	
		L	R
$Player_1$	U	(3,3)	(0,0)
	D	(4,0)	(1,1)

Table 2: Fictitious play game example

In fictitious play, we introduce 2 concepts: *counting and belief.* Counting: a player counts how many times each strategy has been used by the other user. Define η_i^k a vector representing $Player_i$'s counting during iteration k. For example, $\eta_1^0 = (3, 0)$ denotes that for $Player_1$'s counting in iteration 0 that $Player_2$ has used L 3 times and R 0 times. The counts should normally be integer, but in initial stage, it can be any nonzero vector.

The belief is the probability of each strategy that will be used by the other player, formed based on the counting vector. Define μ_i^k as $Player_i$'s belief that in iteration k, the other player's probability of using each strategy. For example, $\mu_1^0 = (0.4, 0.6)$ denotes that $player_1$'s counting in iteration 0, the $player_2$ will use L with empirical distribution of 40% times and R with 60% times.

For the example above, suppose the initial states are:
$\eta_1^0 = (3, 0) \quad \eta_2^0 = (1, 2.5)$
$\mu_1^0 = (1, 0) \quad \mu_2^0 = (\frac{1}{3.5}, \frac{2.5}{3.5})$

First iteration:
$Player_1$ believes $Player_2$ will play L with probability 1, and thus the best response will be playing D.
For $Player_2$, the pay off of playing L: $\frac{1}{3.5} \times 3 + 0$, the pay off of playing R: $0 + \frac{2.5}{3.5} \times 1$. Then $Player_2$ will pick the larger one and play L.
Then update the counting and corresponding belief:
$\eta_1^1 = (4, 0) \quad \eta_2^1 (= 1, 3.5)$
$\mu_1^1 = (1, 0) \quad \mu_2^1 = (\frac{1}{4.5}, \frac{3.5}{4.5})$

Second iteration:
$Player_1$ believes $Player_2$ will play L, and thus the best response will be playing D.
For $Player_2$, the pay off of playing L: $\frac{1}{4.5} \times 3 + 0$, the pay off of playing R: $0 + \frac{3.5}{4.5} \times 1$. Then $Player_2$ will play R.
Then update the counting and corresponding belief:
$\eta_1^2 = (4, 1) \quad \eta_2^2 = (1, 4.5)$
$\mu_1^2 = (0.8, 0.2) \quad \mu_2^2 = (\frac{1}{5.5}, \frac{4.5}{5.5})$

...
and we will see that they play the same strategy after that. Since no player has the incentive to change, if this iteration converges, then the limit point is a NE. In a more general sense, the fictitious play may converge to a distribution of strategies (i.e. the Matching Pennies Game). In this case, the distribution of the belief function follows a periodical pattern, and the long term average of the frequency distribution

forms a NE. Formally,if

$$\lim_{T\to\infty} \frac{\sum_{t=1}^{T} I_{\{s_i^t = s_i\}}}{T} = \sigma(S_i),$$

for $s_i \in S_i$ for all i exists then σ is a NE, where I is the 0 or 1 indicator function. We note that to implement fictitious play, only payoff of one player is needed (for best response). The players do not need to know payoff of the opponent.

Fictitious play is guaranteed to converge to a Nash Equilibrium if the game is any one of the following:

1. Two players Zero Sum Game
2. Two players Non-zero Sum Game where each player has at most two strategies
3. Any game that is dominant solvable
4. Identical interest for all players
5. Potential games

4 Dynamic Games with Complete Information and Perfect Monitoring

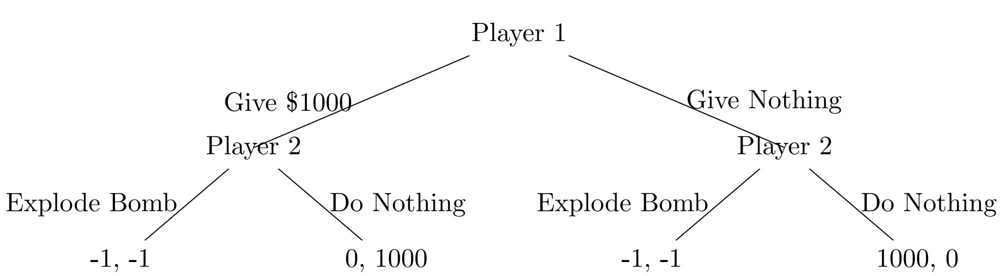
So far, we have studied static games with complete information. Now we move on to dynamic games with complete information, starting with those which also have *perfect monitoring*. By this we mean that agents move one at a time in *stages*, with the ability to view the entire history of moves. Examples of such games include chess and checkers. By contrast, games with *imperfect monitoring* may not reveal all of an opponent's moves, such as when players move simultaneously instead of one at a time.

4.1 Example 1: A 2-Player, 2-Stage Game

Consider a game with two stages. In the first stage, Player 1 can choose between giving Player 2 \$1000, or giving Player 2 nothing. Player 2 then chooses between exploding a bomb that kills both players and doing nothing. The normal form representation of the game is shown below, as though it were a static game.

	Explode Bomb	Do Nothing
Give \$1000	-1, -1	0, 1000
Give Nothing	-1, -1	1000, 0

This game has a Nash Equilibrium at (Give Nothing, Do Nothing). However, the game is not static, and Player 2 may develop a *strategy*, or a plan of action given each possible history of actions. An example strategy is as follows: if Player 1 gives Player 2 \$1000, he or she will not explode the bomb, but if Player 1 fails to pay, Player 2 will blow them both up. To represent a sequential game like this, we may think of the game as a tree, where the first number at each leaf corresponds to the payoff of player 1 (conventionally, the first mover) and the second number is for player 2 (the player which moves second).



We then find four distinct strategies for Player 2, defined by an ordered pair representing his or her respective actions if Player 1 does or does not pay the $1000, (Action if Paid, Action if Not Paid).

	(EB, EB)	(DN, EB)	(EB, DN)	(DN, DN)
Give $1000	-1, -1	*0, 1000*	-1, -1	0, 1000
Give Nothing	-1, -1	-1, -1	*1000, 0*	*1000, 0*

We can then see that there are three equilibria, whose payoffs are represented by italics; however, because the game is sequential rather than static, we must acknowledge that Player 2 never has incentive to explode the bomb after Player 1 has denied them the money, since it worsens Player 2's utility to do so (0 is better than -1). As a result, we call a strategy like (Do Nothing, Explode Bomb) *sequentially irrational* and refer to Player 2's ultimatum as a *non-credible threat*, since we know a rational, self-interested agent would never act on it.

When analyzing sequential games, therefore, we have the concept of a *sub-game perfect Nash Equilibrium*, often abbreviated as SPNE or SPE. This means that for any history of play, we require an agent's strategy to be an equilibrium on the remaining subgame. The definition is recursive and lasts until we reach the bottom most subtrees. To find a SPNE, we start with the bottom subtrees and work upwards. A strategy profile of a finite extensive-form game is a subgame perfect equilibrium (SPE) if and only if there exist no profitable one-stage deviations for each subgame and every player. This means that in order to verify that something is an SPE, we need to check both along equilibrium and off equilibrium path and make sure in all possible outcomes/subtrees, there is no profitable deviation. This process is known as *backward induction*. In backward induction, we start with the last player to make a move and work backwards up the tree.

References

[1] A. Ozdaglar, "6.254 : Game Theory with Engineering Applications Lecture 15: Infinitely Repeated Games", MIT, 2010.

[2] R. Berry and R. Johari, "Economic Modeling in Networking: A Primer", in Foundations and Trends in Networking, 6.3, 2013, pp. 165-286.