Lecture 20: Nash Equilibrium

Lecturer: *Sahil Singla*

Last updated: March 28, 2022

1 Introduction

To clarify, we won't learn how to play Starcraft, Poker, or Settlers of Catan.

Game theory is a field where we want to study how "rational" (payoff maximizing) decision making agents interact with each other. The difference between "games" and just ordinary "decisions" is the following: when you decide what to eat for breakfast, your payoff doesn't depend on what anyone else in the world does. If you like apples a lot, you might get payoff 10. If you don't like bananas as much, you might get payoff 5. So you should eat apples.

When you decide how to vote, your payoff depends on how you vote, but *also on how others vote*. This is what makes voting a "game" instead of just a "decision." The goal of this section is to think a little bit about how to reason about playing games.

Game Theory is an enormous field, and we'll only touch on the basics. We'll start with settings where it is "easier" to predict *rational* behavior. But keep in mind that *rational* is not synonymous with *smart* or *what happens in practice*. We'll discuss this informally through class, on Piazza, etc.

2 General Games

2.1 Setup

A *game* has n players, who each have m strategies. Each player has a *payoff function*, $p_i(\cdot)$ which takes as input strategies from each player and outputs a payoff. So $p_i : [m]^n \to \mathbb{R}$. The collection of payoff functions is sometimes called the *payoff matrix*. All the examples we'll see this week involve $n = 2$ players, but you should try to process the definitions and notation more generally.

2.2 Example: Rock-Paper-Scissors

You and a friend are playing a game where there are 3 actions: (a) rock, (b) paper, or (c) scissor. You both simultaneously select an action. If both play the same action then it is a tie, and you both end with a 0 payoff. Otherwise, we have rock beats scissor, scissor beats paper, and paper beats rock. So whoever plays the beating action gets 1 payoff and the other gets -1 payoff. This payoff can be represented using the following matrix.

Definition 1 (Simultaneous Game)**.** *There are n players where each has m strategies (actions). Each player has a* payoff *function $p_i : \{1, 2, \ldots, m\}^n \to \mathbb{R}$ that takes as input the* *strategies of all the m players and returns their payoff/utility. We usually let s_i denote the strategy of player i and let s_{-i} denote the strategies of all players except i.*

This week we will focus on $n = 2$ players case, where the payoffs can be written as a $m \times m$ matrix where the (s, t) entry contains $(p_1(s_1, s_2), p_2(s_1, s_2))$, i.e., the payoff of the first and second player for actions s_1 and s_2, respectively.

	Rock	Paper	Scissor
Rock	$(0, 0)$	$(-1, 1)$	$(1, -1)$
Paper	$(1, -1)$	$(0, 0)$	$(-1, 1)$
Scissor	$(-1, 1)$	$(1, -1)$	$(0, 0)$

3 Dominant Strategies

Example (Prisoner's Dilemma). You and your partner in crime are suspected of robbing a bank. Fortunately for you, there's not much in the way of concrete evidence. Unfortunately for you, the police understand Game Theory, put you in separate rooms and give you the opportunity to confess.

- If both deny, you each go to jail for a year.
- If one confesses but the other denies, the confessor goes free and the other gets twenty years.
- If you both confess, you both get ten years.

The table below describes the payoff that you and your partner will get, as a function of both of your decisions. The number on the left denotes the payoff to the "row player" (who chooses Confess or Deny row), and the number on the right denotes the payoff to the "column player" (who chooses Confess or Deny column).

	Confess	Deny
Confess	$(-10, -10)$	$(0, -20)$
Deny	$(-20, 0)$	$(-1, -1)$

So now let's think: if your partner is going to confess, what decision gives you the highest payoff? You should also confess and get ten versus twenty years. What if your partner's going to stay silent? Then you should confess and get zero versus one year. So no matter what, it's better for you to confess. We say that confessing *dominates* denying.

Definition 2 (Dominant Strategy). *We say that strategy a* strongly dominates *strategy b for player i if* no matter what *strategies the other players use, i receives* strictly higher *payoff using a than using b. We say that it* weakly dominates *b if i receives higher payoff using a than using b always, and there is at least one strategy profile s_{-i} such that i receives strictly higher payoff using a against s_{-i} than using b. We say that a is a (strongly/weakly)* dominant strategy *if a (strongly/weakly) dominates all other strategies. Note that if a strongly dominates b, then a also weakly dominates b (and that if a is strongly dominant, then a is also weakly dominant).*

If every player in the game has a dominant strategy, it seems reasonable to expect them to play it — no matter what the other players do, you're better off playing the dominant strategy. Note the connection to incentive compatibility from previous lectures: incentive compatibility is essentially saying that telling the truth is a dominant strategy.

Finding dominant strategies is straight-forward (but not always "easy") - just take any two strategies and compare their performance against every possibility. Recall every time we proved that a voting-rule/matching-algorithm was incentive compatible: we needed to show that telling the truth outperformed any lie no matter what the other players did.

Note: to prove that strategy a weakly dominates strategy b, one needs to prove both (i) the payoff of using a is always at least as good as using b and (ii) there exists a strategy such that the payoff from a is strictly better than using b. In most lectures after the "pure game theory" sections, we will focus on just (i) as (ii) is often tedious. Despite the fact that we'll often omit this portion of the proof, property (ii) is important as otherwise games could have multiple dominant strategies.

4 Nash Equilibrium

Last section, we saw games with dominated strategies, where it was straight-forward to define "rational" behavior (Prisoner's Dilemma). We also saw games where dominated strategies could be iteratively deleted (Traveler's Dilemma). In general, games need not have this property. There was also a lot of discussion about whether being rational was "smart."

Consider the following "parking game:" You park your car on campus, and can pay or not pay the meter. The university can hire inspectors or not. The idea is that the university suffers if you park without paying the meter because there will be no available parking. At least if they catch you, the cost is marginally offset by the ticket.

- If you pay the meter, then your payoff is 0 no matter what.
- If you don't pay, your payoff is 10 if the university doesn't inspect, and -90 if they do (heavy fine).
- If the university doesn't inspect, their payoff is 0 if you pay the meter, and -10 if you don't (because no one will be able to find parking).
- If they do inspect, their payoff is -1 if you pay the meter (waste of everyone's time), and -6 if you don't (-1 for paying the inspector, -10 for you parking excessively long, but +5 for your ticket revenue).

	Inspect	Not Inspect
Pay	$(0, -1)$	$(0, 0)$
Not Pay	$(-90, -6)$	$(10, -10)$

Let's try to reason through how the players might behave. If the driver always pays the meter, what should the university do? They should not inspect because it will just be a waste of time. But if the university never inspects, what should the driver do? They should not pay, because they'll never get caught. Let's try to make this formal.

Note: In lecture, we'll give this definition without referencing mixed strategies first. Then motivate mixed strategies, then update the definition. We also only gave the definition for two players for easier notation.

Definition 3 (Best Response)**.** *We say that a (possibly mixed) strategy s is a* best response *for player i to the (possibly mixed) strategy profile $\vec{s}_{-i}$ if s maximizes $p_i(s; \vec{s}_{-i})$ over all possible (possibly randomized) strategies s.*

Definition 4 (Nash Equilibrium)**.** *A strategy profile is a* Nash Equilibrium *if for every i, s_i is a best response to $\vec{s}_{-i}$.*

This seems like a sensible way to predict behavior: if I'm best responding to you, and you're best responding to me, neither of us has any incentive to change our behavior.[1] Unfortunately, the parking game doesn't yet admit a Nash equilibrium: the best response to pay is don't inspect, but the best response to don't inspect is not pay. The best response to don't pay is inspect, but the best response to inspect is pay. So we need to update our definitions slightly to allow players to sometimes play randomly.

Definition 5 (Mixed Strategy)**.** *A* mixed strategy *is a distribution over pure strategies. For example, player i could choose to play each of their strategies s with probability $x_i(s)$. We can refer now to a strategy for player i as a vector $\vec{x}_i$ (which has a coordinate $x_i(s)$ for all strategies s). We also need to define what player i's payoff is when players are using mixed strategies.*

Player i's payoff is just the expectation of $p_i(\vec{s})$, when each player j independently samples s_j according to $\vec{x}_j$. Expanding the math, we also get that player i's payoff when players use the strategy vectors $\vec{x}_j$ is:

$$p_i(\vec{x}) := \mathbb{E}_{s_j \leftarrow \vec{x}_j,\ \forall j}[p_i(\vec{s})] = \sum_{\vec{s}} \prod_j x_j(s_j) \cdot p_i(\vec{s}).$$

The above formula is just computing the expected payoff of $p_i(\vec{s})$ when s_j is drawn independently from $\vec{x}_j$, by summing over all $\vec{s}$, of the probability that we wind up with exactly $\vec{s}$, times the payoff that i gets in this case.

Observation 1. *Observe that the expected payoff for strategy $\vec{x}_i$ for player i when all other players are using $\vec{x}_{-i}$ is exactly:*

$$\sum_{\ell} x_i(\ell) \cdot p_i(\ell; \vec{x}_{-i}).$$

[1] But this is certainly not a university prediction. In fact, there are many issues with Nash equilibria that we won't cover in class, but feel free to ask about them in class, on Piazza, office hours, etc.

That is, the payoff for player i for using $\vec{x}_i$ when all other players use $\vec{x}_{-i}$ just sums over all strategies ℓ for player i, of $x_i(\ell)$ times the payoff that player i gets for using s_ℓ against $\vec{x}_{-i}$.

Proof. To see this, we just expand:

$$\begin{aligned}\sum_\ell x_i(\ell) \cdot p_i(\ell; \vec{x}_{-i}) &= \sum_\ell x_i(\ell) \cdot \sum_{\vec{s}_{-i}} \prod_{j \neq i} x_j(s_j) \cdot p_i(\ell; \vec{s}_{-i}) \\ &= \sum_{\vec{s}} \prod_j x_j(s_j) \cdot p_i(\vec{s}).\end{aligned}$$

The first line follows by just expanding the definition of $p_i(\ell; \vec{x}_{-i})$ (the payoff of using ℓ against $\vec{x}_{-i}$). The second line just observes that first summing over all ℓ, and then taking a product over the remaining $\vec{s}_{-i}$ is the same as just directly summing over all $\vec{s}$. $\square$

Observation 2. *Note that mixed strategies are* also *strategies, so the definitions regarding dominant strategies still apply. For instance, it could be the case that on a large game no pure strategy dominates any other, but a mixture of two may dominate some strategies. This allows rational players to discard them and make progress towards finding rational strategies.*

The idea is that if the parker always pays the meter, then the university can just not inspect. If the parker never pays the meter, the university will surely inspect. If instead the parker *sometimes* pays the meter, the university may be happy to either inspect or not inspect. Let's now update our definition of best response to accommodate mixed strategies, and observe the following:

Observation 3. *For all s_{-i}, there is always a pure strategy that is a best response. A mixed strategy is a best response if and only if it plays only pure best responses with non-zero probability.*

Proof. To see this, recall from Observation 1 that the payoff i gets by playing $\vec{x}_i$ is a weighted average of the payoff that i gets by playing each of their pure strategies. Because a weighted average is clearly never more than the maximum, we see that no mixed strategy can be strictly better than the best pure strategy. So there is always a pure best response. Also, if the mixed strategy is ever not playing a pure best response, it must yield expected payoff strictly lower than that pure best response (and be suboptimal), because it will be a non-zero weighted average of values all of which are at most the best pure strategy, and at least one of which is strictly less. $\square$

In the parking game, let's consider the following strategies: you pay the meter with probability 0.8, and the university inspects with probability 0.1. Then given that you pay the meter with probability 0.8, the university achieves expected payoff -2 by not inspecting ($-10 \cdot 0.2$), or expected payoff -2 by inspecting ($-1 \cdot 0.8 + -6 \cdot 0.2$). So both strategies are a best response. You get payoff 0 for paying the meter, and 0 for not ($10 \cdot 0.9 + -90 \cdot 0.1$). So both strategies are a best response. Therefore, both players are best responding and we've found a Nash equilibrium. What's interesting about Nash equilibria is that they always exist (but we won't prove it in this class).

Theorem 1 (Nash [2]). *For all finite n, m, every game with n players and m strategies has at least one (possibly mixed) Nash equilibrium.*

We won't prove the theorem here, since it requires several non-trivial steps. If you are interested then [5] presents a clear proof based on combinatorial Sperner's Lemma and Brouwer's Fixed Point Theorem.

There's not really consensus within any community as to when you should expect players to behave according to a Nash equilibrium. One barrier is that no one knows any natural dynamics that are known to converge to Nash equilibria. In particular, it's widely believed that finding a Nash equilibrium requires time exponential in m, even for two-player games. There's a formal statement for this, but it involves a complexity class besides just P and NP so we won't get into it (Formally, finding a Nash in a two-player m-action game is PPAD-complete).

For this class, we won't stress about whether Nash is predictive or not of behavior - it's still an extremely useful analytical concept and has been the driving force behind Game Theory since Nash's discovery - he won the 1994 Nobel prize for his theorem.

5 Extended Form Games - Subgame Perfect Equilibria

Extended form games are a special case of general games that have a natural time/sequential aspect to them. Consider for instance any board game: you could lock in your entire strategy at the very start and refuse to change it (this would be the game in its "normal-form," but the nature of the game allows for you to refine decisions later on).

Consider the following example, which is actually one of the early uses of Game Theory in policy (especially during e.g. the cold war). Say two countries, A and B both have nukes. A is considering an aggressive action. If A attacks, B can choose either to retaliate or back down (B has no action if A maintains peace).

- If A attacks and B retaliates, both receive payoff $-L$ (think of $L = \infty$) because the world is destroyed.
- If A attacks and B backs down, A receives payoff $+1$ and B receives payoff -1.
- If A stays peaceful, then no matter what B does each receives payoff 0.

To be extra clear, A has two strategies in this game "attack" or "maintain peace." B has two strategies in this game, which we will call "retaliate" and "back down." The semantic meaning of "retaliate" is "if A chooses to attack, B will retaliate. If A maintains peace, then nothing." The semantic meaning of "back down" is "if A chooses to attack, B will back down. If A maintains peace, then nothing."

This is a two-player two action game, and it has two pure Nash equilibria: A can remain peaceful and B can "retaliate," and both are best responding (if A deviates, then the world explodes. B's action doesn't matter because A is peaceful). Also, A can escalate and B can back down. Which one seems more reasonable to you, taking the sequential nature of the game into account?

Well, if you're B, and A has already attacked, you're faced with a pretty clear decision to either just accept it or blow up the world. So while it really sucks, you're most likely going to accept it. So A can reason that the "threat to retaliate" isn't credible because it doesn't benefit B. So the first Nash equilibrium is perhaps unreasonable, and the latter one is more likely to arise. This is called *subgame perfect*.

Definition 6 (Subgame Perfect Equilibrium)**.** *An extensive form game consists of a tree. Every non-leaf node of the tree is associated with a single player, who chooses actions at that node. Each action points to a different child. Each leaf node contains a payoff for each player.*

A pure strategy *for a player i selects a pure action for* every *node where player i might act. A (mixed) complete strategy is a distribution over pure strategies. A Nash equilibrium is then defined as before.*

A Nash equilibrium is subgame perfect *if for every possible subtree, the induced strategies on that subtree are a Nash equilibrium.*

So for example, we might draw the War Game as having a root labeled with player A, and two edges coming out (Attack and Peace). The Peace edge points to a leaf with payoffs (0,0). The Attack edge points to a node labeled with player B, which itself has two edges coming out labeled Retaliate and Back Down. The Retaliate edge points to a leaf with payoffs (-L, -L). The Back Down edge points to a leaf with payoffs (1,-1). The Nash equilibrium where B threatens to Retaliate and A maintains Peace is not subgame perfect, because if we restrict attention to the game which starts at the node labeled B, it is not a best response for B to Retaliate. That is, the game which starts at B just requires B to choose whether to blow up the world or accept minor defeat, and it is not a best response for B to blow up the world. The Nash equilibrium where B Backs Down and A Attacks is subgame perfect, because no matter which node we start the game at, this is a Nash equilibrium.

Important note: Recall above that the definition of a strategy for a player requires a list of what action to perform at every node where they act. This is necessary even for paths which aren't actually played in order for players to reason about what "would have" happened if they acted differently.

5.1 Centipede Game

Let's consider another game, called the Centipede game: The game will last 100 rounds. Initially, there is 4 dollars in the pot in round 0, when player 1 acts. Players alternate taking turns between *push* and *take*. If a player *takes* from the pot, they get half of the pot (rounded down) plus two. If they *push*, they increase the size of the pot by one, but now the other player gets to act. On turn 100, player 1 must take.

So the payoff structure looks like this: If the first turn that a player takes is i, then that player gets payoff $\lfloor i/2 \rfloor + 4$. The other player gets payoff $\lfloor i/2 \rfloor$. Let's try to reason about a possible subgame perfect equilibrium.

Starting in the last round (100), player 1 has to take, so they would get payoff 54, and player 2 would get payoff 50. In round 99, player 2 could then push and get payoff 50, or take and get payoff 53. So any subgame perfect equilibrium must have player 2 take at round 99. Now back to 98, player 1 could take and get payoff 53, or push and get payoff 50 when player 2 takes. So any subgame perfect equilibrium must have player 1 take at round 98.

Reasoning all the way backwards, we see that at round i, the acting player may take and get payoff $\lfloor i/2 \rfloor + 4$, or push and get payoff $\lfloor (i+1)/2 \rfloor$ when the opposing player takes in round $i+1$. So going all the way back to round 0 we see that the only subgame perfect equilibrium is for every player to take at every round, and player 1 gets payoff 4, player 2 gets payoff 0, even though clearly both would be happier at essentially any later round.

Lemma 2. *The only subgame perfect equilibrium in the Centipede game is where each player takes in every round in which they act.*

Proof. We prove this by backwards induction (and in fact, the SPNE of any game can be found by backwards induction). We claim, starting from $i = 100$, that any subgame-perfect equilibrium must have the acting player take at round j for all $j \geq i$.

The base case is $i = 100$, where the acting player must take by definition.

Now assume for inductive hypothesis that the claim holds for $i + 1$, and we will prove it for i.

As the acting player has switched in between i and $i + 1$, and the other player will take during round $i+1$ in any SPNE, this means that player i will get payoff $\lfloor (i+1)/2 \rfloor$ if they push, or payoff $\lfloor i/2 \rfloor + 4 > \lfloor (i+1)/2 \rfloor$ if they take. Therefore, in any potential equilibrium where the active player takes in round $i + 1$, the active player must also take in round i to possibly be best responding in the subgame which starts at round i. By the inductive hypothesis, all SPNE must have the active player take in round $i + 1$, and therefore all SPNE must also have the active player take in round i.

Note that in order to complete this reasoning, we needed to reference *the subgame starting at round i* — we couldn't have done this reasoning exactly for Nash which are not subgame perfect. □

Notes

This lecture is based on COS 445 notes from Princeton University [4].

References

[1] Osborne, Martin. J.. "An Introduction to Game Theory". 2004.

[2] Nash, John. "Non-cooperative games." Annals of mathematics (1951): 286-295.

[3] http://timroughgarden.org/f13/l/l20.pdf

[4] Algorithmic Game Theory. Nisan, Roughgarden, Tardos, Vazirani (eds.), Cambridge University Press 2007.

[5] https://www.cs.princeton.edu/∼smattw/Teaching/cos445sp19.htm