MIT 6.S890 — Topics in Multiagent Learning Tue, Sep 10th 2024

Lecture 2

Setting and equilibria: the Nash equilibrium

Instructor: Prof. Gabriele Farina (gfarina@mit.edu)*

Normal-form games model simultaneous-move interactions with a single move (think about rock-paper-scissors). Despite their simplicity, normal-form games will provide a natural ground for looking into important concepts in multiagent settings, such as notions of equilibria (Nash, maxmin, correlated, ...), and learning from repeated play. In the second part, we will move on to notions of games that explicitly capture more complex phenomena, such as sequential moves and imperfect information.

1 Normal-form games and the Nash equilibrium

When introducing a (finite) normal-form game, we need to specify the following quantities:

- The set of players $[n] = \{1, ..., n\}$.
- For each player $i \in [n]$, a finite set of actions A_i.
- For each player $i \in [n]$, the payoff function $u_i : A_1 \times \cdots \times A_n \to \mathbb{R}$.

To represent a normal-form game, it is customary to use a matrix representation. For example, in the 2×2 game on the right, called "prisoner's dilemma", the rows correspond to the actions of Player 1, and the columns correspond to the actions of Player 2. The entries at row i and column j are the payoffs of the two players when Player 1 plays action i and Player 2 plays action j.

	deny	confess
deny	$-1, -1$	$-3, 0$
confess	$0, -3$	$-2, -2$

■ **Strategies**. A *randomized strategy* (also known as *mixed strategy*) for a generic player $i \in [n]$ is a distribution over the set of actions. We can represent such an object as a vector $x_i \in \Delta(A_i)$, that is, such that $x_i \geq 0$ and $\sum_{a_i \in A_i} x_{i,a_i} = 1$. To lighten the notational burden, we will write the expected utility when all players play according to strategies $x_1, ..., x_n$ reusing the same letter u_i as the payoff, that is

$$u_i(x_1, ..., x_n) := \mathbb{E}_{\substack{a_1 \sim x_1 \\ \cdots \\ a_n \sim x_n}}[u_i(a_1, ..., a_n)] = \sum_{a_1 \in A_1} \cdots \sum_{a_n \in A_n} x_{1,a_1} \cdots x_{n,a_n} \cdot u_i(a_1, ..., a_n).$$

We will sometimes intersperse deterministic actions and mixed strategies freely and write expressions such as $u_i(a_1, x_2, ..., x_n)$ to mean the expected utility when player 1 plays action a_1 and the other players play according to the strategies $x_2, ..., x_n$.

1.1 Dominant-strategy equilibrium

The question of what constitutes rational play for players can get complicated depending on the game. But, in some lucky cases, like the prisoner's dilemma game above, it turns out that some actions are just *better* than others, *no matter what the other players do*. In such cases, we say that a player has a *dominant*

*These notes are class material that has not undergone formal peer review. The TAs and I are grateful for any reports of typos.

strategy. In the case above, both Player 1 and Player 2 have a dominant strategy to confess. In this case, we expect that the players will play their dominant strategy, and this is called a *dominant-strategy equilibrium*.

1.2 Maxmin strategies

The benefit of dominant-strategy equilibria is that they require no counterspeculation: some strategies just are better no matter what anyone else does. However, in many games, no player has a dominant strategy. Consider, for example, rock-paper-scissor: all actions are symmetric, and no action is strictly better than the others. How can we find a good strategy for that?

One way to think about this is to consider the worst-case scenario: what is the best strategy for a player if they assume the other players are trying to minimize their payoff? This is the idea behind *maxmin strategies*. A maxmin strategy for Player i is a strategy x_i that maximizes the minimum payoff that Player i can get, that is,

$$x_i \in \arg\max_{x_i \in \Delta(A_i)} \min_{\substack{x_j \in \Delta(A_j) \\ \text{for all } j \neq i}} u_i(x_i', x_{-i}),$$

where the notation x_{-i} is popular syntectic sugar to denote the tuple $\left(x_j\right)_{j \neq i}$.[1] Thinking back about rock-paper-scissors, it is clear that the maxmin strategy is to play uniformly at random: the opponent could exploit anything else by playing the counteraction more often.

The above idea has some merits, especially in two-player zero-sum games, that is, those two-player games where $u_1(a_1, a_2) + u_2(a_1, a_2) = 0$ for all combinations of actions. In those games, players are in direct competition, so it makes sense to assume that players will be "out to get us". But in more general games, the maxmin strategy can be too conservative, since it assumes that all other players have nothing better going on than to minimize our payoff, even if that hurts them.

1.3 The Nash equilibrium

In general, defining what constitutes "optimal play" is tricky. But we can start from what is convincingly *not* optimal play: if we predict that the players should play according to some strategies $x_1, ..., x_n$, then it is not optimal if it turned out that any player would be better off by switching to something else. This is the idea behind the *Nash equilibrium*.

Definition 1.1 (Nash equilibrium). A strategy profile $(x_1, ..., x_n) \in \Delta(A_1) \times \cdots \times \Delta(A_n)$ is a *Nash equilibrium* if no player benefits from unilaterally deviating from their strategy. In symbols,

$$\forall i \in [n], x_i' \in \Delta(A_i), \qquad u_i(x_i', x_{-i}) \leq u_i(x_1, ..., x_n).$$

Remark 1.1. Without loss of generality, when verifying if a profile $(x_1, ..., x_n)$ is a Nash equilibrium, it is sufficient to consider only *deterministic* deviations $x_i' \in A_i$. Indeed, if a player has a profitable randomized deviation, this must mean that at least one of the actions they are randomizing over is profitable.

It is clear that a dominant-strategy equilibrium is a special case of a Nash equilibrium, since in a dominant-strategy equilibrium, by definition,

$$\forall i \in [n], x_i' \in \Delta(A_i), x_{-i}' \in \Delta(A_i), \qquad u_i(x_i', x_{-i}') \leq u_i(x_1, x_{-i}'). \qquad \text{(Dominant-strategy eq.)}$$

(note the stronger quantifiers.) As we will discuss more in depth shortly, in two-player zero-sum games, it turns out that Nash equilibrium and maxmin equilibrium are equivalent.

[1]This notation appears often in game theory, since we are often in studying the effect of changing a *single* player i's strategy, while keeping all "the other" strategies x_{-i} fixed.

Before continuing, we consider a couple of examples that help illustrate a couple of important properties of the Nash equilibrium.

Example 1.1. The only Nash equilibrium in the game of rock-paper-scissors is for all players to play the uniform strategy. This shows that in some games, no Nash equilibrium exists in pure (*i.e.*, non-randomizing) strategies

Example 1.2 (Theater or football). Consider the following small game. Player 1 can either insist on going to the theater or accept attending the football match. Player 2 can either insist on going to the theater or accept attending the football match. The payoffs are as follows:

	insist	accept
insist	0, 0	5, 1
accept	1, 5	0, 0

This game has two obvious Nash equilibria: Player 1 insisting and Player 2 accepting, or vice versa (top right and bottom left corners). However, there is a third equilibrium as well: both players accept a probability of $1/6$ and insist with a probability of $5/6$.

This is not a coincidence: in two-player nondegenerate games, there is always an *odd* number of Nash equilibria. This fact comes from more profound connections with some combinatorial objects that we will uncover toward the end of the course.

2 Existence of mixed-strategy Nash equilibrium

In 1950, John Nash established one of the most celebrated results in game theory:[2] mixed-strategies in Nash equilibria exist in all games, no matter the number of players or number of actions. The proof of Nash is nonconstructive, and fundamentally boils down to showing that one can think of Nash equilibria as fixed points. Two remarks are in order:

- The idea that Nash equilibria can be thought of as fixed points should feel extremely natural. By definition, an equilibrium is a situation where nobody has an incentive to deviate—that is, no player has a more profitable strategy given the strategies of the other players. Hence, if we are able to introduce a function that maps a strategy profile to the profile of "most profitable strategies for each player keeping everyone else fixed", then a fixed point of such a function would be a Nash equilibrium. We could then use one of the many theorems in analysis that guarantee existence of fixed points.

  Of course, the devil is in the details: *how to handle multiple profitable responses? how to ensure continuity of the deviation function?* These are the questions that Nash had to answer.

- The idea of viewing Nash equilibria as fixed points is not just natural, but also *the only possible*. We will make this formal towards the end of this course, when we discuss the *computational complexity* of Nash equilibria. In particular, we will show that the computation of fixed points of continuous functions and computation of Nash equilibria are computationally equivalent, in the sense that each problem can be reduced to the other in polynomial time.

In the remainder of the section, we will give a proof of the existence of Nash equilibria. While the first proof of Nash, J. F., Jr. [Nas50] invokes Kakutani's fixed point theorem, a year later Nash noticed that a much more elementary proof can be given [Nas51]. We present a variation of the latter today.

[2]John Nash went on to win the Nobel prize in economics for his fundamental contributions to game theory.

2.1 The Nash improvement function

As mentioned above, one can think about Nash equilibria as fixed points of a "profitable response" function from the set to mixed strategy to itself. Intuitively, this function must calculate a profitable response for each player. Furthermore, to invoke fixed point theorems, this function must be continuous. The key insight of Nash was to find a simple continuous function that, given a strategy profile, calculates a "profitable response" for each player. For lack of a better term, I will refer to this function with the term *"Nash improvement function"*.

■ **Regret.** To formally define the Nash improvement function, we first introduce a simple quantity called *regret*, which will be a staple of this course. The *regret* that Player i experiences with respect to action $a_i \in A_i$ is the difference between the payoff that Player i would have obtained by playing a_i, and the payoff that Player i actually obtained:

$$r_{i,a_i}(x_1, ..., x_n) \coloneqq u_i(a_i, x_{-i}) - u_i(x_1, ..., x_n).$$

■ **The Nash improvement function.** The idea now is simple: if an action a_i has very large regret, then the current strategy profile cannot be an equilibrium, because Player i would want to increase the probability of playing a_i. Thus, an "improved" strategy for Player i should move more probability mass to a_i. We need to handle two complications: (1) if multiple actions have positive regret, how should we prioritize adding mass to those? and (2) for those actions whose regret is negative (that is, "bad" actions), should we forcefully decrease the mass?

Nash's answers to the above questions are as follows: (1) add mass to all actions with positive regret, and the amount of mass added should be proportional to the regret; (2) do not decrease the mass for actions with negative regret. To retain the fact that the output of the improvement function must be a valid strategy, the step is renormalized so that the sum of the mass across all actions of any player is 1. We can formalize this process by using the following definition.

Definition 2.1 (Nash improvement function [Nas51]). Let $x_1 \in \Delta(A_1), ..., x_n \in \Delta(A_n)$ be arbitrary strategies. The *Nash improvement function* $\varphi : \Delta(A_1) \times \cdots \times \Delta(A_n) \to \Delta(A_1) \times \cdots \times \Delta(A_n)$ is the map

$$\varphi_{i,a_i}(x_1, ..., x_n) \coloneqq \frac{x_{i,a_i} + \left[r_{i,a_i}(x_1, ..., x_n)\right]^+}{1 + \sum_{a_i' \in A_i} \left[r_{i,a_i'}(x_1, ..., x_n)\right]^+} \tag{1}$$

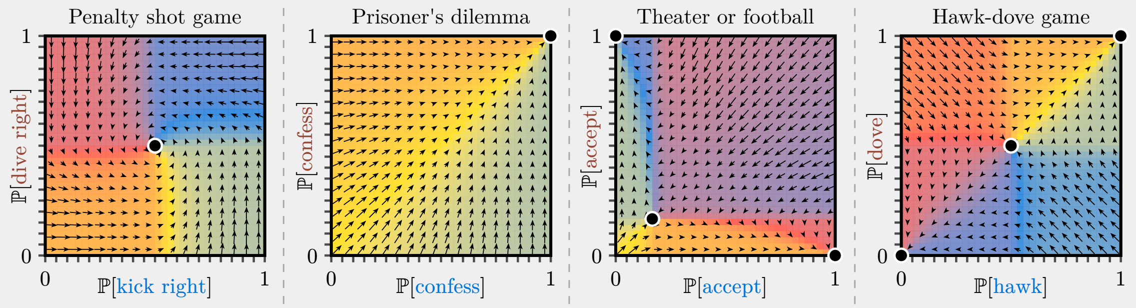
for all player $i \in [n]$ and action $a_i \in A_i$. Here, $[r]^+ \coloneqq \max\{0, r\}$ denotes the positive part of r.

It is straightforward to verify that φ is well-defined and maps strategy profiles into strategy profiles. Indeed, the numerator in 1 is always nonnegative, and the denominator is always at least 1, implying that $\varphi_{i,a_i} \geq 0$ for all $a_i \in A_i$ and player $i \in [n]$. Furthermore,

$$\forall i \in [n], \qquad \sum_{a_i \in A_i} \varphi_{i,a_i} = \frac{\sum_{a_i \in A_i} \left(x_{i,a_i} + \left[r_{i,a_i}(x_1, ..., x_n)\right]^+\right)}{1 + \sum_{a_i' \in A_i} \left[r_{i,a_i'}(x_1, ..., x_n)\right]^+} = 1,$$

where we used the fact that $\sum_{a_i \in A_i} x_{i,a_i} = 1$ since x_i is a valid strategy. Finally, observe that φ is a continuous function. The following example visualizes the Nash improvement function in the small games we have seen so far.

Example 2.1. The plots below visualize the displacement $\varphi(x_1, x_2) - (x_1, x_2)$ induced by the Nash improvement function for four games, whose payoff matrices are noted below each plot, after projecting away the probability of the first action of each player (and keeping around only the probability of the second action, which is sufficient to uniquely recover the strategy of the player since each player only has two actions). The black dots denote the fixed points of the Nash improvement function. These correspond exactly to the Nash equilibria of the game, as we make formal below.



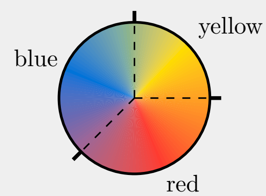
	dive L	dive R
kick L	−1, 1	1, −1
kick R	1, −1	−1, 1

	deny	confess
deny	−1, −1	−3, 0
confess	0, −3	−2, −2

	insist	accept
insist	0, 0	5, 1
accept	1, 5	0, 0

	hawk	dove
dove	0, 4	2, 2
hawk	−2, −2	4, 0

The background of the plots highlights the angle of displacement induced by the Nash improvement function, according to the gradient wheel shown on the right. [If the color scheme seems arbitrary, as a small spoiler it will play a fundamental role in the proof of the *computational complexity* of Nash equilibria, which we will discuss later on in this course. In particular, the three regions of the coloring scheme will be key in defining an important *combinatorial* construction called *Sperner coloring*.]



2.2 Quantifying the increase in utility of the improvement step

We validate our intuition that the Nash improvement function is a "profitable response" function. The following result shows that if a player has positive regret for any action, then the Nash improvement function unilaterally increases that player's utility. This is a key property that will allow us to show that the fixed points of the Nash improvement functions must be Nash equilibria.

Theorem 2.1. For any strategy profile $(x_1, \ldots, x_n)$, and any player $i \in [n]$, the Nash improvement function φ satisfies

$$u_i(\varphi_i(x_1, \ldots, x_n), x_{-i}) - u_i(x_1, \ldots, x_n) = \frac{\sum_{a_i \in A_i} \left(\left[r_{i,a_i}(x_1, \ldots, x_n)\right]^+\right)^2}{1 + \sum_{a_i \in A_i} \left[r_{i,a_i}(x_1, \ldots, x_n)\right]^+}.$$

So, if even one action of a player i has positive regret, then the Nash improvement function unilaterally *strictly* increases the utility of that player.

Proof. Since we are focusing on a generic player i and keeping all the other ones fixed (and playing strategies x_{-i}), we will reduce the notational burden by using the following shorthands:

$$r_{i,a_i} := r_{i,a_i}(x_1, \ldots, x_n), \qquad \text{(regret of action } a_i \text{ of Player } i \text{ fixing } x_{-i})$$

$$u_{i,a_i} := u_i(a_i, x_{-i}), \qquad \text{(utility of action } a_i \text{ of Player } i \text{ fixing } x_{-i})$$

$$x'_{i,a_i} := \varphi_{i,a_i}(x_1, \ldots, x_n) = \frac{x_{i,a_i} + r^+_{i,a_i}}{1 + \sum_{a'_i \in A_i} r^+_{i,a'_i}} \qquad \text{(prob. of action } a_i \text{ in the improved strategy)},$$

where the notation z^+ is a shorthand for the positive part of z, that is, $z^+ := [z]^+ := \max\{0, z\}$.

The increase in utility is then computed as

$$\begin{aligned}
\sum_{a_i \in A_i} u_{i,a_i} \cdot \left(x'_{i,a_i} - x_{i,a_i}\right) &= \sum_{a_i \in A_i} u_{i,a_i} \cdot \left(\frac{x_{i,a_i} + r^+_{i,a_i}}{1 + \sum_{a'_i \in A_i} r^+_{i,a'_i}} - x_{i,a_i} \right) \\
&= \sum_{a_i \in A_i} u_{i,a_i} \cdot \frac{r^+_{i,a_i} - \sum_{a'_i \in A_i} r^+_{i,a'_i} \cdot x_{i,a_i}}{1 + \sum_{a'_i \in A_i} r^+_{i,a'_i}} \\
&= \frac{1}{1 + \sum_{a'_i \in A_i} r^+_{i,a'_i}} \left(\sum_{a_i \in A_i} r^+_{i,a_i} \cdot u_{i,a_i} - \sum_{a'_i \in A_i} \left(r^+_{i,a'_i} \sum_{a_i \in A_i} x_{i,a_i} \cdot u_{i,a_i} \right) \right) \\
&= \frac{1}{1 + \sum_{a'_i \in A_i} r^+_{i,a'_i}} \left(\sum_{a_i \in A_i} r^+_{i,a_i} \cdot \left(u_{i,a_i} - \sum_{a'_i \in A_i} u_{i,a'_i} \cdot x_{i,a'_i} \right) \right) \\
&= \frac{1}{1 + \sum_{a'_i \in A_i} r^+_{i,a'_i}} \left(\sum_{a_i \in A_i} r^+_{i,a_i} \cdot r_{i,a_i} \right).
\end{aligned}$$

Using the fact that $z^+ \cdot z = (z^+)^2$ for all $z \in \mathbb{R}$, we obtain the statement. □

At this point, the following is a simple corollary.

Theorem 2.2. A strategy profile $(x_1, ..., x_n)$ is a Nash equilibrium if and only if it is a fixed point of the Nash improvement function φ.

Proof. ($\Longrightarrow$) If $(x_1, ..., x_n)$ is a Nash equilibrium, then by definition for all $i \in [n]$ and $a_i \in A_i$, we have $r_{i,a_i}(x_1, ..., x_n) \leq 0$. Hence, for all $i \in [n]$ and $a_i \in A_i$, we have

$$\varphi_{i,a_i}(x_1, ..., x_n) = \frac{x_{i,a_i} + \left[r_{i,a_i}(x_1, ..., x_n)\right]^+}{1 + \sum_{a'_i \in A_i} \left[r_{i,a'_i}(x_1, ..., x_n)\right]^+} = x_{i,a_i},$$

that is, $(x_1, ..., x_n)$ is a fixed point of φ.

($\Longleftarrow$) Conversely, suppose that $(x_1, ..., x_n)$ is a fixed point of φ. Then, for all $i \in [n]$, from Theorem 2.1 we have

$$\frac{\sum_{a_i \in A_i} \left(\left[r_{i,a_i}(x_1, ..., x_n)\right]^+\right)^2}{1 + \sum_{a_i \in A_i} \left[r_{i,a_i}(x_1, ..., x_n)\right]^+} = u_i(\varphi_i(x_1, ..., x_n), x_{-i}) - u_i(x_1, ..., x_n) = 0.$$

Hence, it must be $r_{i,a_i}(x_1, ..., x_n) \leq 0$ for all $i \in [n]$ and $a_i \in A_i$ (or the left-hand side would be strictly positive), and therefore $(x_1, ..., x_n)$ is a Nash equilibrium. □

By invoking Brouwer's fixed-point theorem, we recover the central result of this lecture: Nash equilibria always exist.

Corollary 2.1. Since φ is continuous and maps the nonempty compact convex set $\Delta(A_1) \times \cdots \times \Delta(A_n)$ into itself, by Brouwer's fixed point theorem, it has a fixed point. By Theorem 2.2, this implies that every game has (at least) one Nash equilibrium in mixed strategies.

Bibliography

[Nas50] J. F. Nash Jr., “Equilibrium Points in N-Person Games,” *Proc. Natl. Acad. Sci. U.S.A.*, vol. 36, no. 1, p. 48, Jan. 1950, doi: `10.1073/pnas.36.1.48`.

[Nas51] J. Nash, “Non-Cooperative Games,” *Annals of Mathematics*, vol. 54, no. 2, pp. 286–295, 1951, [Online]. Available: `http://www.jstor.org/stable/1969529`

Changelog

- Sep 10: fixed location of NE dots in the rightmost plot of Example 2.1.