Modeling Human Behavior

An Introduction to game theory – Part 3: Nash equilibrium

Christian Hilbe, christian.hilbe@it-u.at
Version May 27, 2026

Remark 1.22 (A recapitulation of the previous classes)
In our last class, we introduced two concepts:

1. We defined the arguably simplest class of non-trivial games – called normal-form games. Those are games between *two players*, who both have *finitely many actions*, players move *simultaneously*, and they have *complete information* about all aspects of the game (except for the co-player's decision). Such games can be represented by a payoff matrix. Rows represent the possible actions of player 1, whereas columns represent the possible actions of player 2.
2. We learned what it means for an action to be dominated: *Action a is (strictly) dominated if there is some other action a' that always yields a better payoff, no matter what the co-player does.* We argued that rational players should never use dominated actions; so those actions should be *eliminated*. Moreover, if both players are rational (and know of each other's rationality, etc), then we can iteratively eliminate dominated actions. We showed that for some games, (iterated) elimination of dominated strategies yields a unique solution. Such games are called *dominance solvable*. The two examples we had were (a slightly modified) *Split or Steal* and *Split or Steal with Remorse*,

	Split	Steal
Split	$20k, 20k$	$-1k, 40k$
Steal	$40k, -1k$	$0, 0$

	Split	Steal
Split	$20k, 20k$	$-1k, 40k$
Steal	$5k, -1k$	$0, 0$

In both cases, the unique solution is (Steal,Steal). However, the cognitive requirements are different. In the first game, we only needed to assume that both players are rational. In the second game, we needed that both players are rational *and* that the first player knows the second player is rational.

1.4 Nash equilibrium

Remark 1.23 (Optimal behavior in the stag-hunt game)
Unfortunately, it is fairly easy to come up with games that are not dominance-solvable. One example is the stag-hunt game with payoff matrix

	Stag	Hare
Stag	$10, 10$	$0, 4$
Hare	$4, 0$	$4, 4$

Opting for Stag is not dominated – after all, it is the best thing I can do if I expect my co-player to hunt stag as well. Similarly, hunting Hare is not dominated either. Still, some outcomes seem less likely than others. For example, it would be odd to consider (Stag, Hare) a 'solution' (i.e., for the row-player to go for a stag-hunt, while the column-player hunts a hare). Here, at least one of the players would immediately want to deviate from that outcome. In that regard, an outcome like (Hare, Hare) or (Stag, Stag) seems much more reasonable.

Definition 1.24 (Nash equilibrium)
A pair of actions (a_R, a_C) for the two players (one action for the row-player and one for the column-player) is a Nash equilibrium if neither player can improve its payoff by unilaterally deviating.

In-Class Exercise 1.25 (Three examples)

1. What are the Nash equilibria of the stag-hunt game?
2. What are the Nash equilibria of the snowdrift game (assuming the benefit of cooperation b is larger than the cost c of cooperating),

	Cooperate	Defect
Cooperate	$b-c/2, b-c/2$	$b-c, b$
Defect	$b, b-c$	$0, 0$

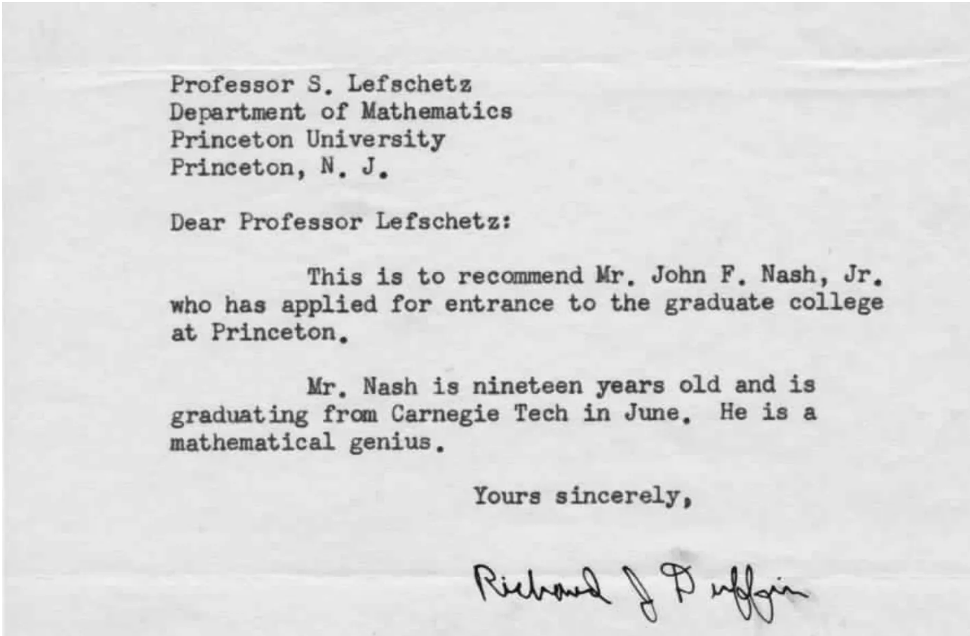
3. Penalty kicks. Consider a game between a striker and a goalkeeper. The striker needs to choose whether to shoot left or right, the goalkeeper chooses in which direction to jump. To simplify matters, suppose the striker scores if and only if the goalkeeper jumps into the wrong direction. Moreover, if we assume the striker is the row-player, we can represent this game as follows,

	Left	Right
Left	$0, 1$	$1, 0$
Right	$1, 0$	$0, 1$

Can we find a pair of actions that is a Nash equilibrium in that game?

Remark 1.26 (Mixed strategies.)

- The game of penalty kicks suggests that sometimes, players may wish to be unpredictable (there are many other such examples, such as in Rock-Scissors-Paper, or in warfare). We can incorporate this by saying that players do not choose a single action with certainty; rather they play their actions according to some probability distribution (x_L, x_R). In that case, we say that the player adopts a 'mixed strategy'. [In the special case that the mixed strategy says one action should be played with certainty, e.g., $x_L = 1$ and $x_R = 0$, we also speak of a 'pure strategy'].
- One can define a payoff for mixed strategies by considering their expected value. For example, in the penalty kicks game, suppose the striker chooses 'Left' with 60% probability; moreover, suppose the goalkeeper jumps left with 90% probability. Then the striker's payoff is $0.6 \cdot 0.1 + 0.4 \cdot 0.9 = 0.42$.



Professor S. Lefschetz
Department of Mathematics
Princeton University
Princeton, N. J.

Dear Professor Lefschetz:

This is to recommend Mr. John F. Nash, Jr. who has applied for entrance to the graduate college at Princeton.

Mr. Nash is nineteen years old and is graduating from Carnegie Tech in June. He is a mathematical genius.

Yours sincerely,

Richard J Duffin

Figure 1: A recommendation letter for John Nash.

- One can extend the previous definition of Nash equilibria to mixed strategies: a pair of strategies is an equilibrium if neither player has an incentive to deviate. For penalty kicks one can show: if both players set $x_L = x_R = 1/2$, that is a Nash equilibrium.

Theorem 1.27 (**Nash (1950)**)
Any normal-form game has at least one equilibrium point (possibly in mixed strategies).

Remark 1.28 (About John Nash and his equilibrium result)

- John Nash was a mathematical genius (you don't need to believe me; you can just trust his undergraduate supervisor, see **Figure 1**). Nash derived the above theorem as a part of his PhD thesis, published in 1950 (32 pages, written on a typewriter). He got the Nobel prize in Economics in 1995. He also got the Abel prize in 2015 (for something else entirely). Apart from these successes, his life was rather tragic (see the movie 'A beautiful mind').
- John Nash's proof is an example of a *non-constructive* proof. That is, the proof does not tell you how an equilibrium point for a particular game would look like. It merely tells you that such an equilibrium point *must* exist. Technically speaking, the proof is based on a *fixed point argument*. This technique roughly uses the fact that if you have a sufficiently nice function $f(x)$ that maps the points of some nice space S into itself, then there must be at least one x^* that is mapped onto itself, $f(x^*) = x^*$ (compare to the fact that if you have a continuous function f that takes numbers in [0,1] and maps them onto [0,1]; then there must be at least one fixed point x^*).

Remark 1.29 **(On the status of the Nash equilibrium concept)**

- It is fair to say that the Nash equilibrium is the standard concept to 'solve' normal-form games.
- In some situations, human behavior is also fairly well predicted by the Nash equilibrium. For example, empirical papers suggest quite a good agreement for penalty kicks in soccer (Chiappori et al., 2002) or for tennis serves (Walker and Wooders, 2001). [Bonus question: Why do researchers use sports to compare the Nash equilibrium concept to data?]
- However, it is also fairly easy to construct artificial games in which human subjects considerably deviate from the Nash equilibrium predictions. For example, consider the following two examples by Goeree and Holt (2001):

	Left	Right
Top	$80, 40$	$40, 80$
Bottom	$40, 80$	$80, 40$

	Left	Right
Top	$320, 40$	$40, 80$
Bottom	$40, 80$	$80, 40$

In both games, there is a unique Nash equilibrium. According to that equilibrium, the row-player should choose both options 50-50 in each of the two games. This prediction is experimentally confirmed for the left hand game. However, it is refuted in the right hand game; there, 96% of participants choose Top.

Remark 1.30 **(A summary)**

In this first part of this course, you should have learned

- what game theory is about (about strategic decision-making when individuals need to take others' possible decisions into account)
- how to formalize a game in a mathematical model (by identifying the players, their actions, the order of moves, the information they have, and the payoffs)
- how to 'solve' that model (we considered two solution concepts: iterated elimination of dominated strategies, and the Nash equilibrium).

Administrative Remark 1.31 **(A few administrative points we should discuss)**

- Now there is a one-week break. On June 8, we will talk about *evolutionary game theory*. On June 9, we will talk a bit about game-theoretic concepts we haven't covered in this course; I will also talk a bit about the research we do in my research group. Finally, I would suggest to have the interviews/conversations on that day. Could we start a bit earlier (e.g., 10:00 instead of 10:45)?
- *Interviews:* As part of your evaluation, I would like to have a 10 minute conversation on game theory with each of you. During that conversation, you should demonstrate that you understood the basic concepts of game theory. To make it easier, I created a list of possible questions; you can find this list online (either on my homepage, or on Canvas).
- Everyone should have access to the course on Canvas by now. How to upload assignments?
- Does everyone know how to program a simple simulation (e.g., in Python)?

References

Chiappori, P.-A., Levitt, S., and Groseclose, T. (2002). Testing mixed-strategy equilibria when players are heterogeneous: The case of penalty kicks in soccer. *American Economic Review*, 92(4):1138–1151.

Goeree, J. K. and Holt, C. A. (2001). Ten little treasures of game theory and ten intuitive contradictions. *American Economic Review*, 91(5):1402–1422.

Nash, J. F. (1950). Equilibrium points in n-person games. *Proceedings of the National Academy of Sciences USA*, 36:48–49.

Walker, M. and Wooders, J. (2001). Minimax play at Wimbledon. *American Economic Review*, 91:1521–1538.