ME 292B Class Project – A Game-Theoretic Analysis of the "Bor-Bor Zan" Game

Yatong Bai

1. Introduction and Problem Formulation

In the lectures of ME 292B and the used textbook [1], we characterized the Nash Equilibria (NE) of the "rock-paper-scissors" game. As it is an instance of circulant games, certain evolutionary dynamics (such as the pairwise proportional imitation protocol) fail to converge to equilibrium. In this project, we study another relatively simple but exciting game – "Bor-Bor Zan". This game was (and perhaps still is) extremely popular when I was in elementary school. Similar to rock-paper-scissors, Bor-Bor Zan also adopts the two-player scenario as its basic form, and each player must deploy one of the three strategies – "build", "defend", and "attack" for each round. Consider two players, A and B. The rules are the following:

- If player A attacks player B and player B fails to defend, then B dies, and vice versa.
- Each attack consumes one unit of energy. The only way to gain energy is to deploy "build", which makes the player vulnerable for that round. A player cannot choose to attack without sufficient energy (i.e., the probability of the "attack" strategy must be zero).
- If both players attack simultaneously, then the attacks cancel out, and both players consume one unit of energy.

Since this game has the notion of the energy state, it is much more complicated than "rock-paper-scissors", and requires breaking down into several cases to analyze. In particular, one interesting question is whether the trivial never-losing strategy of deploying "defense" forever is optimal.

There have been discussions about Bor-Bor Zan from game-theoretic perspectives [2]. However, the author of [2] considers a different variant of the game, incorporating an additional fourth strategy. Furthermore, we use different reward designs inspired by the preference of maximizing future rewards. This report also offers more technical rigor compared with [2].

2. Game-Theoretic Analyses

We start with designing the rewards. One straightforward formulation is that for either player, at each round, the reward is $r_{\text{win}} > 0$ if the player wins at this round, $r_{\text{lose}} < 0$ if the player dies, and 0 if the game continues to the next round. We thus start from this design. Note that due to the dynamic nature of this game, different energy states can lead to different optimal strategies at each round. Therefore, we must break down the Bor-Bor Zan game into cases (i.e., a series of bi-matrix games) to make the analysis tractable. Fortunately, the game satisfies the Markov property, meaning that the current energy state can completely define the (probabilistic) future outcomes, and there is no need to analyze past states.

To facilitate the analysis, we introduce the notations $e_{\text{A}} \in \mathbb{N}$ and $e_{\text{B}} \in \mathbb{N}$ to denote the energy state, where $\mathbb{N}$ is the set of natural numbers. We also use $\mathbb{N}_+$ to denote the set of positive integers. We also denote the player A's strategies as $x_{\text{A}} =: (x_{\text{A}}^{\text{build}}, x_{\text{A}}^{\text{defend}}, x_{\text{A}}^{\text{attack}}) \in \mathcal{S}^3$, where $\mathcal{S}^d$ denotes the d-dimensional simplex, defined as the set $\{x \in \mathbb{R}^d : \sum_{i=1}^d x_i = 1\}$. Player B's strategies x_{B} are defined in the same manner.

We make the assumption that for each player, if the game continues to the next round, a higher energy state is strictly preferred against a lower energy state, independent from the energy state of the opponent. Note that winning/losing is still the most/least desirable outcome, overriding the preference for energy state. This assumption makes sense because a higher energy state will not forbid a player from attacking, whereas a lower energy state may disallow attacking. Since attacking is the only way of winning, not being able to attack at certain rounds strictly diminishes the chance of winning. We further assume that the strengths of such a preference are equal for the two players. We leave the discussion about whether this preference for a high-energy state is truly optimal for future work. With this assumption, we can now analyze the game case by case.

2.1. Case 0: $e_{\mathrm{A}} = e_{\mathrm{B}} = 0$

In this setting, the considered game is symmetric, meaning that the optimal action should be the same for A and B. Thus, the payoff for player A is Ax_{A} for some matrix A. If we only consider the winning/losing reward, a naive choice for A is then $A = \begin{bmatrix} 0 & 0 & r_{\mathrm{lose}} \\ 0 & 0 & 0 \\ r_{\mathrm{win}} & 0 & 0 \end{bmatrix} \begin{pmatrix} -x_{\mathrm{A}}^{\mathrm{attack}}, 0, x_{\mathrm{A}}^{\mathrm{build}} \end{pmatrix}$. However, by construction, $x_{\mathrm{A}}^{\mathrm{attack}} = x_{\mathrm{B}}^{\mathrm{attack}} = 0$, and thus the payoffs associated with the two remaining strategies are always zero, meaning that actions "defend" and "build" are equally preferred according to our model. This is incorrect per our assumption, as "build" enables the player to attack at the next state and thus should be preferred. Therefore, we must embed the preference for a high-energy state into the payoff matrices.

Now, we regard the name as a two-strategy game for this case for simplicity (exploiting the fact that $x_{\mathrm{A}}^{\mathrm{attack}} = 0$). Due to the symmetry and the fairness of the Bor-Bor Zan game between players A and B, the mathematical modeling of the payoff for this scenario would be $\bar{A} \begin{bmatrix} x_{\mathrm{A}}^{\mathrm{build}} \\ x_{\mathrm{A}}^{\mathrm{defend}} \end{bmatrix}$ at optimality, where $\bar{A} = \begin{bmatrix} \alpha_{00}^0 & \alpha_{01}^0 \\ \alpha_{10}^0 & \alpha_{11}^0 \end{bmatrix}$ and α_{ij}^0 represents the expected payoff for player A when the game enters the state resulted from the A's action $i \in \{0, 1\}$ and B's action $j \in \{0, 1\}$ starting from Case 0.

While it is hard to obtain the accurate values of α_{ij}^0, it is straightforward to see that $\alpha_{01}^0 > \alpha_{11}^0$ and $\alpha_{00}^0 > \alpha_{10}^0$ (since a higher future energy state is always preferred), and therefore the only NE is $x_{\mathrm{A}} = (1, 0, 0)$. I.e., both A and B should always deploy "build" in this state, and we enter the energy state $e_{\mathrm{A}} = e_{\mathrm{B}} = 1$. Furthermore, since $\begin{bmatrix} \alpha_{00}^0 & \alpha_{01}^0 \end{bmatrix} x > \begin{bmatrix} \alpha_{10}^0 & \alpha_{11}^0 \end{bmatrix} x$ for all $x \geq 0$, it holds that $x^\top \bar{A} x < \begin{bmatrix} 1 & 0 \end{bmatrix} Ax$ for all $x \in \mathcal{S}^2 \backslash \{\begin{bmatrix} 1 & 0 \end{bmatrix}\}$, and thus $\begin{bmatrix} 1 & 0 \end{bmatrix}$ is an evolutionarily stable state (ESS).

2.2. Case 1: $e_{\mathrm{A}} = e_{\mathrm{B}} = \epsilon$ for some $\epsilon \in \mathbb{N}_+$

In this setting, the considered bi-matrix game is again symmetric, meaning that the optimal action should be the same for A and B. Now, we similarly define the quantities α_{ij}^1 for $i, j \in \{0, 1, 2\}$ as the expected payoff resulted from the (i, j) action pair starting from Case 1. The payoff for player A is then $\begin{bmatrix} \alpha_{00}^1 & \alpha_{01}^1 & r_{\mathrm{lose}} \\ \alpha_{10}^1 & \alpha_{11}^1 & \alpha_{12}^1 \\ r_{\mathrm{win}} & \alpha_{21}^1 & \alpha_{22}^1 \end{bmatrix} x_{\mathrm{A}}$.

Note that $r_{\mathrm{win}} > \alpha_{00}^1 > \alpha_{10}^1$, $\alpha_{01}^1 > \alpha_{11}^1 > \alpha_{21}^1$, and $\alpha_{12}^1 > \alpha_{22}^1 > r_{\mathrm{lose}}$, because a higher future energy state is better than a lower energy state by our assumption, but winning at the current round is the best outcome whereas losing is the worst. We now consider the following scenarios:

1. $x_{\mathrm{A}} = (1, 0, 0)$. The payoff is then $(\alpha_{00}^1, \alpha_{10}^1, r_{\mathrm{win}})$, and the best response is $(0, 0, 1)$. Thus, this case cannot admit an NE.

2. $x_{\mathrm{A}} = (0, 1, 0)$. The payoff is then $(\alpha_{01}^1, \alpha_{11}^1, \alpha_{21}^1)$, and thus the best response is $(1, 0, 0)$. Thus, this case cannot admit an NE.

3. $x_{\mathrm{A}} = (0, 0, 1)$. The payoff is then $(r_{\mathrm{lose}}, \alpha_{12}^1, \alpha_{22}^1)$, and thus the best response is $(0, 1, 0)$. Thus, this case cannot admit an NE. Thus, there is no deterministic NE for this state.

4. $x_{\mathrm{A}} = (a, 1-a, 0)$ for some $a \in (0,1)$. The payoff is then $\begin{bmatrix} a \cdot \alpha^1_{00} + (1-a) \cdot \alpha^1_{01} \\ a \cdot \alpha^1_{10} + (1-a) \cdot \alpha^1_{11} \\ a \cdot r_{\text{win}} + (1-a) \cdot \alpha^1_{21} \end{bmatrix}$. Note that since $\alpha^1_{00} > \alpha^1_{10}$ and $\alpha^1_{01} > \alpha^1_{11}$, the action "build" is strictly better than "defend". This makes sense because we assume the probability of an attack is zero, so there is no need to defend. Thus, having $x^{\text{defend}}_{\mathrm{A}} > 0$ is suboptimal, and this case cannot admit an NE.

5. $x_{\mathrm{A}} = (a, 0, 1-a)$ for some $a \in (0,1)$. The payoff is then $\begin{bmatrix} a \cdot \alpha^1_{00} + (1-a) \cdot r_{\text{lose}} \\ a \cdot \alpha^1_{10} + (1-a) \cdot \alpha^1_{12} \\ a \cdot r_{\text{win}} + (1-a) \cdot \alpha^1_{22} \end{bmatrix}$. Note that since $\alpha^1_{00} < r_{\text{win}}$ and $r_{\text{lose}} < \alpha^1_{12}$, the action "attack" is strictly better than "build". Note that this setting is equivalent to the hawk-dove game, where "hawk" is always the preferred action even though simultaneously "doving" results in a high combined payoff than both "hawking", unless the price of both "hawking" is too high. Thus, having $x^{\text{build}}_{\mathrm{A}} > 0$ is suboptimal, and this case cannot admit an NE.

6. $x_{\mathrm{A}} = (0, a, 1-a)$ for some $a \in (0,1)$. The payoff is then $\begin{bmatrix} a \cdot \alpha^1_{01} + (1-a) \cdot r_{\text{lose}} \\ a \cdot \alpha^1_{11} + (1-a) \cdot \alpha^1_{12} \\ a \cdot \alpha^1_{21} + (1-a) \cdot \alpha^1_{22} \end{bmatrix}$. Note that since $\alpha^1_{11} > \alpha^1_{21}$ and $\alpha^1_{12} > \alpha^1_{22}$, the action "defend" is strictly better than "attack". This makes sense because we assume the probability of building is zero, so attacks can never be successful. Since the "attack" response is nothing but a waste of energy, having $x^{\text{attack}}_{\mathrm{A}} > 0$ is suboptimal, and this case cannot admit an NE.

7. Since this game must admit an NE, the NE must lie in the interior of $\mathcal{S}^3$. I.e., the NE has non-zero probabilities for all three possible actions.

Now, we can draw some preliminary conclusions about the Bor-Bor Zan game. As illustrated above, the game can enter states where $e_{\mathrm{A}} \neq e_{\mathrm{B}}$, and we must analyze those cases. Moreover, at the NE, it is still possible to lose the game. Finally, we have proven that the trivial strategy of always defending is suboptimal, even though it guarantees that the player will not lose the game.

2.3. Case 2: $e_{\mathrm{A}} > e_{\mathrm{B}}$ or $e_{\mathrm{A}} < e_{\mathrm{B}}$

Without loss of generality, consider the case of $e_{\mathrm{A}} > e_{\mathrm{B}}$. Note that e_{B} may or may not be zero. From A's perspective, consider the following three cases:

1. Player B decides to deploy "build". In this case, A should attack and win the game.

2. Player B chooses to defend. In this case, the best response of A is to build. This is because A cannot win the game at this round regardless, and thus it should try to enlarge its energy state advantage.

3. Player B chooses to attack. In this case, the best response of A is to defend. This is because A must choose between "defend" and "attack" in this round, with the attack action leaving A with less remaining energy.

Therefore, a deterministic response is suboptimal for A unless B is deterministic. On the other hand, from B's perspective, it also faces three scenarios:

1. Player A decides to build. In this case, if $e_{\mathrm{B}} > 0$, then B should attack and win. However, since A's optimal response to B's attack is to defend rather than to build, this is not an NE. If $e_{\mathrm{B}} = 0$, then B should "build" to improve its energy state. However, since A's optimal response to B's "build" is to attack rather than to "build", this is again not an NE.

2. Player A decides to attack. In this case, B should defend (due to the same reasoning on A's side) regardless of its energy state. However, since A's optimal response to the defending B is to build rather than to attack, this is not an NE.

3. Player A decides to defend. In this case, B should build (due to the same reasoning on A's side) regardless of its energy state. However, since A's optimal response to the defending B is to attack rather than to defend, this is also not an NE.

Therefore, being deterministic is suboptimal for B for both $e_{\mathrm{B}} = 0$ and $e_{\mathrm{B}} \neq 0$. As a result, the optimal strategy is again stochastic for both players. The same conclusion applies to the case of $e_{\mathrm{A}} < e_{\mathrm{B}}$.

3. Conclusion

This report analyzes the two-player three-strategy Bor-Bor Zan game from a game-theoretic perspective. Under the reasonable assumption that each player prefers having a high-energy state if the game continues to the next round, we make the following findings:

1. While always deploying the "defend" action guarantees that the player will not lose, this strategy is not an NE.

2. If both players have no energy at a given round (one example is the first round of the game), then both players deploying "build" is the only NE.

3. If both players have equal non-zero energy states, then the NE is a stochastic strategy with non-zero probabilities for all three actions. Therefore, the game will not get stuck in this state.

4. If the two players have different energy states, the optimal strategies are once again stochastic for both players.

The above results demonstrate that the "Bor-Bor Zan" game is fair, exciting, and well-designed, with no trivially optimal strategies. Moreover, the application of dynamic programming as an attempt to solve this problem is non-trivial, if not impossible. The specific location of the NE depends on how strongly each player prefers high-energy states. There are also several directions for future discussion:

1. Understand the specific location of the NEs.

2. Perform Monte-Carlo simulations to understand the dynamics of the game better.

3. Prove whether having a preference for a high-energy state is optimal. If so, investigate the optimal preference strength associated with each energy state.

4. Learn an optimal strategy via reinforcement learning.

References

[1] William H. Sandholm. *Population Games And Evolutionary Dynamics*. Economic learning and social evolution. MIT Press, 2010.

[2] ShaDanXiaoMoWang. Does the "Bor-Bor Zan" game have optimal solutions, hybrid strategies, or Nash equilibria? `https://www.zhihu.com/question/275344377/answer/380364234`, 2020.