

# Balancing Objectives: Ethical Dilemmas and AI's Temptation for Immediate Gains in Team Environments

**Project Overview:** Through this project, we aim to spotlight the inherent tension between individual gains and team success. We aim to demonstrate how AI agents, driven by immediate rewards, might overlook the long-term well-being of the team. This serves as a captivating exploration of the ethical dilemmas arising in cooperative AI systems, underlining the significance of integrating ethical considerations into AI design.

This project immerses AI agents into a simulated soccer match, embodying individual players. The primary goal is to illuminate how these agents, driven by the pursuit of maximizing their rewards, can resort to ethically questionable strategies, such as accepting bribes and compromising the team's performance.

**Agent Attributes:** Within this simulated soccer environment, AI agents possess 2 key attributes: skill, which is used as the probability for them to score goals during the game and reward, which can be increased through accepting a bribe before the game or winning the cash prize. Each time an agent accepts a bribe, their skill level decreases by 0.1. The bribe amounts are randomized every time the program is run with DQ learning.

**Ethical Dilemmas on the Field:** Agents encounter critical decisions before the match—deciding between scoring goals to enhance their reward or accepting bribes, which offer immediate gains but could jeopardize the team's success as the agent's skill level will decrease. Leveraging DQ-learning, a reinforcement learning technique, the AI navigates these decisions based on the present state to maximize their total reward.

## Project Findings:

Generally, there has been a mix of results. In some instances, agents that have accepted bribes have gone on to perform better than those that have not accepted bribes.. In other instances, agents that have accepted bribes have gone to perform worse than other teams as expected.

```
Bribe of $80.42927783712881 accepted by Agent 1.
New skill level: 0
Bribe of $80.42927783712881 accepted by Agent 2.
New skill level: 0
Bribe of $80.42927783712881 accepted by Agent 3.
New skill level: 0
Bribe of $80.42927783712881 accepted by Agent 4.
Team 3 vs Team 2: 5 4
Agent 2 rejected the bribe offer of $80.42927783712881.
Agent 2 rejected the bribe offer of $80.42927783712881.
New skill level: 0
Bribe of $80.42927783712881 accepted by Agent 3.
New skill level: 0
Bribe of $80.42927783712881 accepted by Agent 4.
Team 3 vs Team 2: 5 4
New skill level: 0
Bribe of $80.42927783712881 accepted by Agent 1.
New skill level: 0
Bribe of $80.42927783712881 accepted by Agent 2.
New skill level: 0
Bribe of $80.42927783712881 accepted by Agent 3.
Bribe of $80.42927783712881 accepted by Agent 4.
Team 1 vs Team 2: 5 4
New skill level: 0
Bribe of $80.42927783712881 accepted by Agent 1.
New skill level: 0
Bribe of $80.42927783712881 accepted by Agent 2.
Agent 3 rejected the bribe offer of $80.42927783712881.
Agent 4 rejected the bribe offer of $80.42927783712881.
Team 1 vs Team 2: 5 4
Training DQN - Episode 1
Training DQN - Episode 2
Training DQN - Episode 3
Training DQN - Episode 4
Training DQN - Episode 5
Training DQN - Episode 6
Training DQN - Episode 7
Training DQN - Episode 8
Training DQN - Episode 9
Training DQN - Episode 10
Training DQN - Episode 11
Training DQN - Episode 12
Training DQN - Episode 13
Training DQN - Episode 14
Training DQN - Episode 15
Training DQN - Episode 16
Training DQN - Episode 17
Training DQN - Episode 18
Training DQN - Episode 19
Training DQN - Episode 20
F1: Team 3 vs Team 2: 5 4
Total Rewards
Agent 1: 626.439185506812
Agent 2: 652.297786318053
Agent 3: 1286.8686325939325
Agent 4: 1286.8686325939325
Bribes accepted:
Agent 1: 1286.439185506812
Agent 2: 1552.297786318053
Agent 3: 1286.8686325939325
Agent 4: 1286.8686325939325
(base) dhruvkaul@Dhruvs-MacBook-Pro-2 r1_soccer %
```

Figure 1: An example execution of the code

# Appendix

## A1

The `main()` function orchestrates the initialization of four agents, constituting two teams, and a `DQNAgent` designed for training. A `SoccerGame` instance is instantiated with predefined parameters, including match duration, bribe amount, and batch size for training the `DQNAgent`. A loop iterates over 20 episodes, overseeing the simulation of matches and `DQNAgent` training through replaying experiences. Outcomes of each match determine rewards and bribes for individual agents. The data pertaining to rewards and accepted bribes by each agent is aggregated for subsequent analysis.. Finally, the `main()` function is invoked to execute the simulation and training procedures.