Open-World Game Theory

John Lim,[1] Mayank Kejriwal[2]

[1] Northwestern University, [2] University of Southern California Information Sciences Institute

USC Viterbi School of Engineering Information Sciences Institute

NSF

Work performed under REU Site program supported by NSF grant #2051101

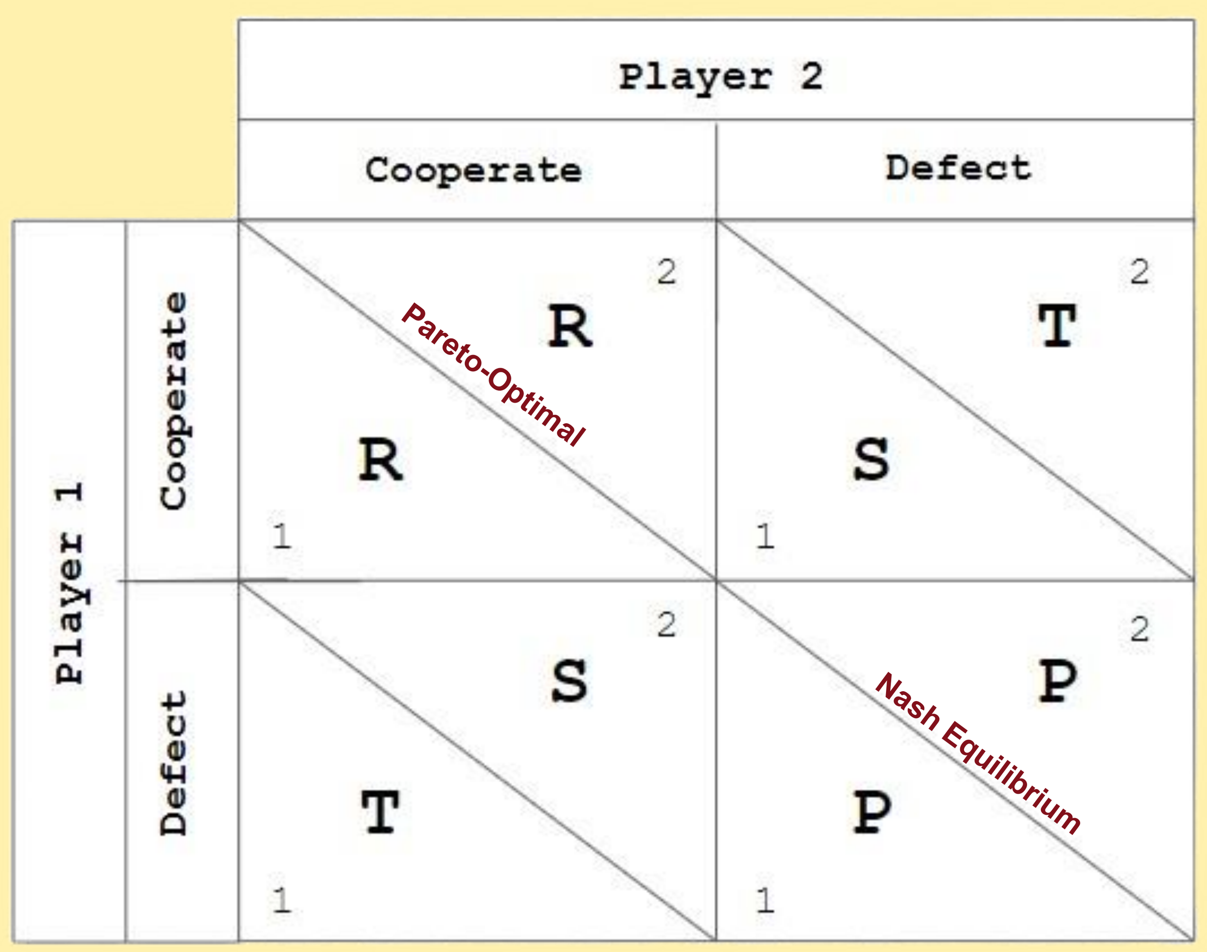
Introduction

Game theory is an influential study designed to understand strategic interactions among rational players. While Prisoner's Dilemma, a classic model of game theory, has been extensively studied for various agent interactions, its applications in **open-world** settings, where unexpected events can, and do, occur, remains relatively shallow due to the heightened complexity involved.

Developing both theoretical and empirical methodologies in support of the open-world game theory have the potential for broader impact as AI systems continue to be applied to real-world, or open-world settings.

Prisoner's Dilemma

Figure 1: Pay-off Matrix in the Prisoner's Dilemma Game



- Two Decisions: Cooperate or Defect
- 4 Total Outcomes: Different payoff for each player
- $T > R > P > S$ → Dominant Strategy: Defect
- Non-cooperative and Non-Zero-Sum
- Single or Iterative: Players can play the game consecutively, giving them the chance to learn about their counterpart and act accordingly

Pure vs. Mixed Strategy

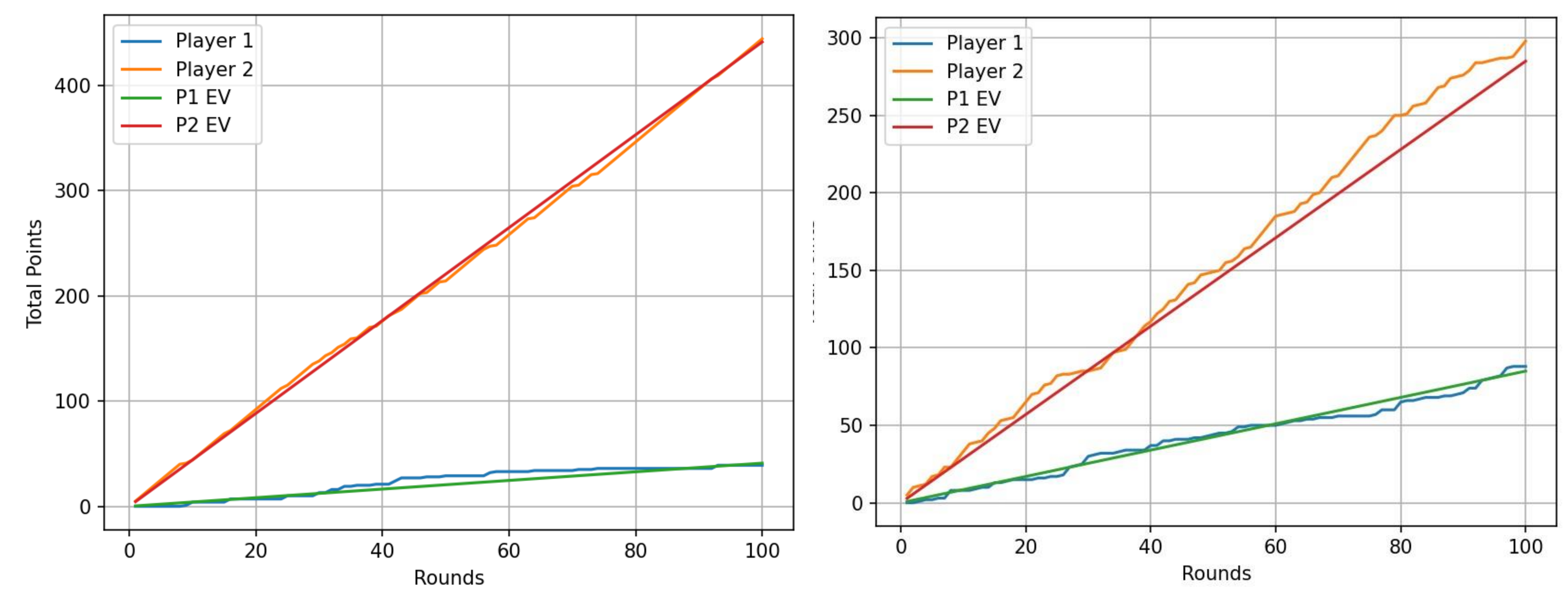
What if players make their decisions based on some probability?

$q = probability\ of\ player\ 1\ defecting$
$p = probability\ of\ player\ 2\ defecting$

Expected Value:

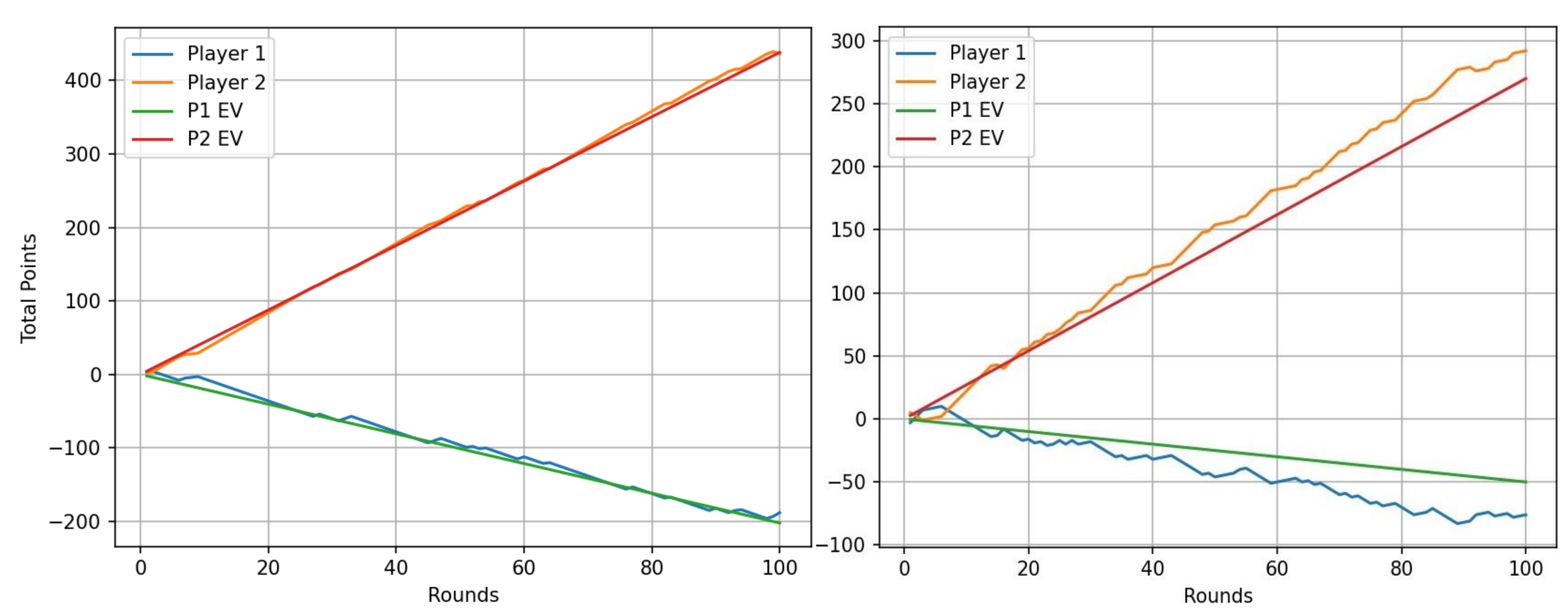
$$EV_{P1} = R(1-q)(1-p) + S(1-q)(p) + T(q)(1-p) + P(q)(p)$$
$$EV_{P2} = R(1-q)(1-p) + T(1-q)(p) + S(q)(1-p) + P(q)(p)$$

Figure 2: Probabilistic Approach To Iterative Prisoner's Dilemma Game
T = 5, R = 3, P = 1, S = 0



$q = 0.1 \quad p = 0.9$ $q = 0.5 \quad p = 0.9$

Figure 3: Penalties Introduced to Iterative Prisoner's Dilemma Game
T = 5, R = 3, P = 1, S = -3



$q = 0.1 \quad p = 0.9$ $q = 0.5 \quad p = 0.9$

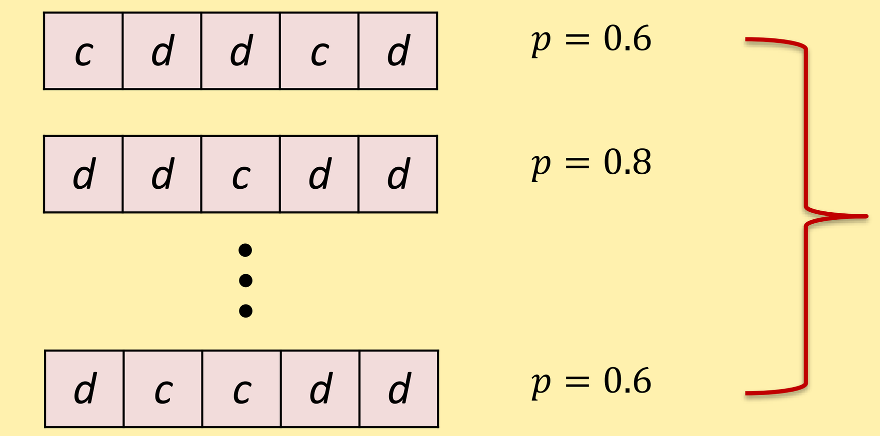
- Player with higher probability of defecting → Greater total points
- Both players defecting more → Less Total Points
- Increasing Variance→ Disparity between EV and total points
- Penalties → Higher Variance

$$Var(PD) = (R-EV)^2x_1 + (S-EV)^2x_2 + (T-EV)^2x_3 + (P-EV)^2x_4$$
$$Var(PD_1 + PD_2 \ldots PD_{100}) = 100 \times Var(PD)$$

Change of Strategies

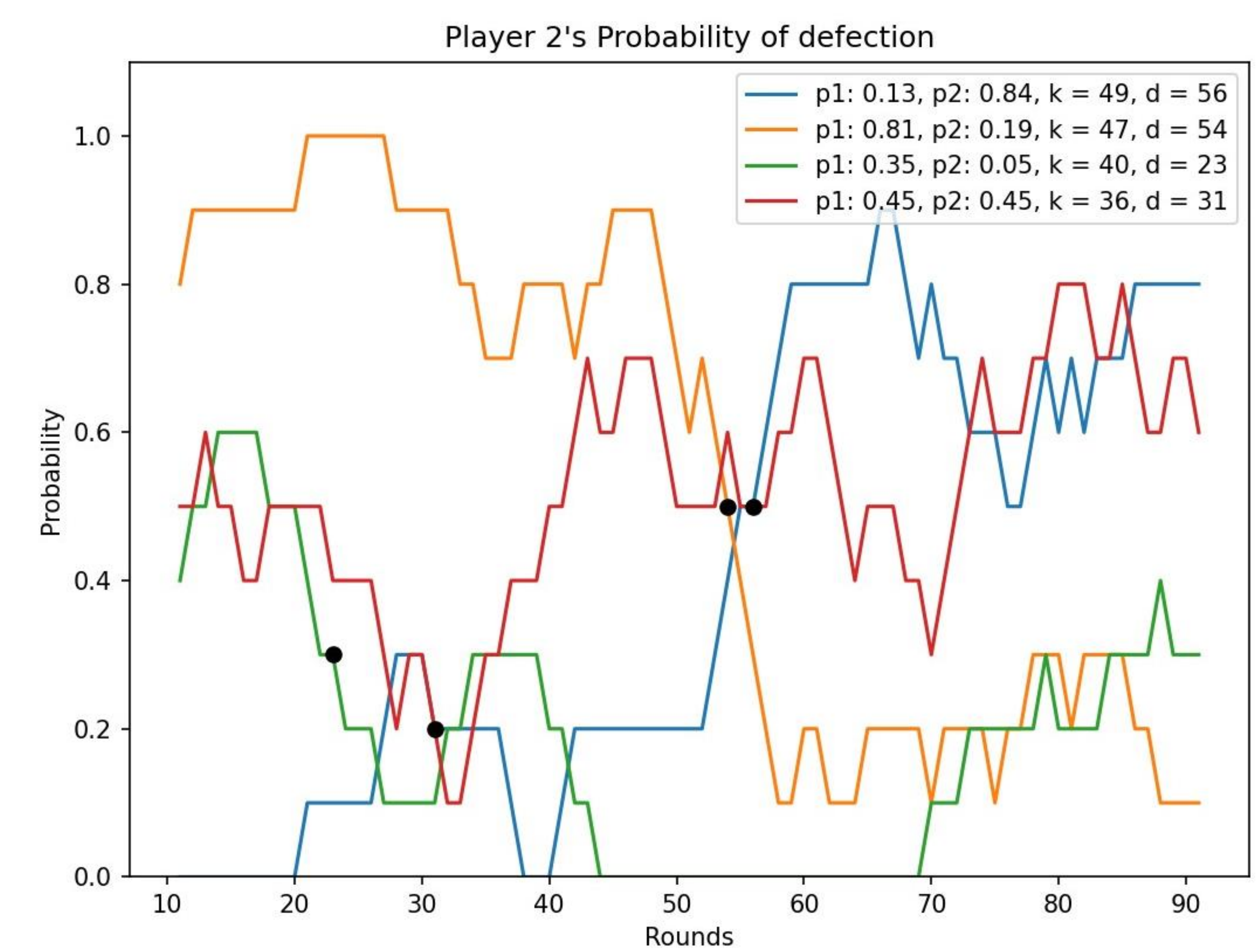
How do I detect a change in high confidence?

Sliding Window:



$|p_x - p_y| > threshold$

Figure 4: Change Detection of Player 2's Probability of Defection
k = Round of Probability Change, d = Round of Detection



Next Steps

- Develop various detection methods
- Assess and compare the methods
  1. **False Positive Rate**
  2. **Prediction Error**
- How to act after the detection?
- Introduce other players and elements into the game