Variable-Player Learning for Simulation-Based Games

Madelyn Gatchel

Advisor: Bryce Wiedenbeck

Outline

- **Introduction and Motivation**
- Background
- Model: Approximating robust symmetric Nash equilibria
- Analysis: Equilibrium robustness metrics
- Experiments
- Ongoing Work

Introduction

- **Game theory**: branch of economics that aims to model how people or "agents" interact and make decisions

- **Machine learning**: branch of computer science which uses mathematical techniques to learn functions from data

- **Our work**: Uses machine learning to analyze large, symmetric, variable-player simulation-based games

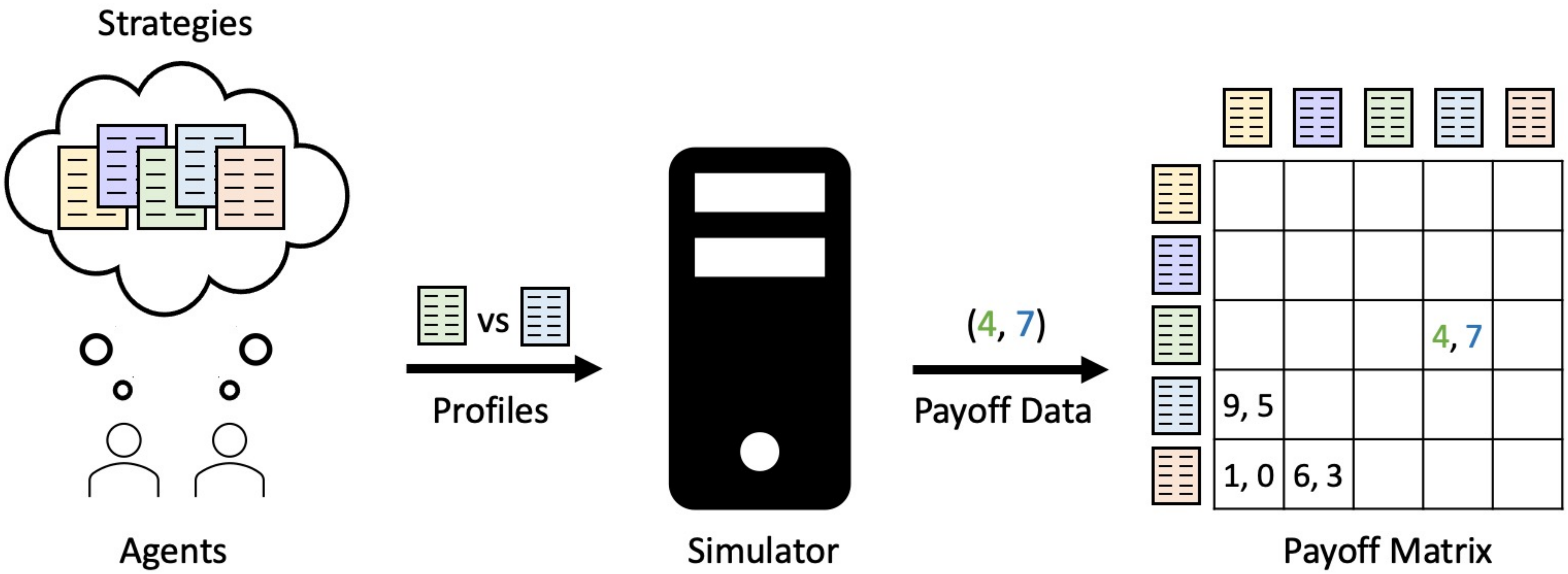
Normal-Form Games

- Type of simultaneous-move game
- Fixed set of players
- For each player
  - Set of strategies
  - Utility function
- Represented using payoff matrix

Rock-Paper-Scissors

		Player 2		
		R	**P**	**S**
Player 1	**R**	0, 0	-1, 1	1, -1
	P	1, -1	0, 0	-1, 1
	S	-1, 1	1, -1	0, 0

Simulation-Based Games



Motivation

- Applications of SBGs
  - Stock market
  - Cybersecurity
  - Credit networks
  - Trading agent competitions
- Likely that the number of players is **large** and **unknown**

Our Work

- How do we construct a variable-player game-theoretic model?

- How do we analyze a variable-player game-theoretic model?

Outline

- Introduction and Motivation
- **Game Theory Background**
- Model: Approximating robust symmetric Nash equilibria
- Analysis: Equilibrium robustness metrics
- Experiments
- Ongoing Work

Background

- **Symmetric game**: game in which any permutation of players produces same game
- **Pure strategy:** any action available to player
- Examples: **R**, **P**, **S**
- **Mixed strategy:** probability distribution over actions
- Example: (1/3, 2/3, 0)

Rock-Paper-Scissors

1

Player ~~2~~

Player ~~1~~ 2	**R**	**P**	**S**
R	0, 0	-1, 1	1, -1
P	1, -1	0, 0	-1, 1
S	-1, 1	1, -1	0, 0

Background (cont.)

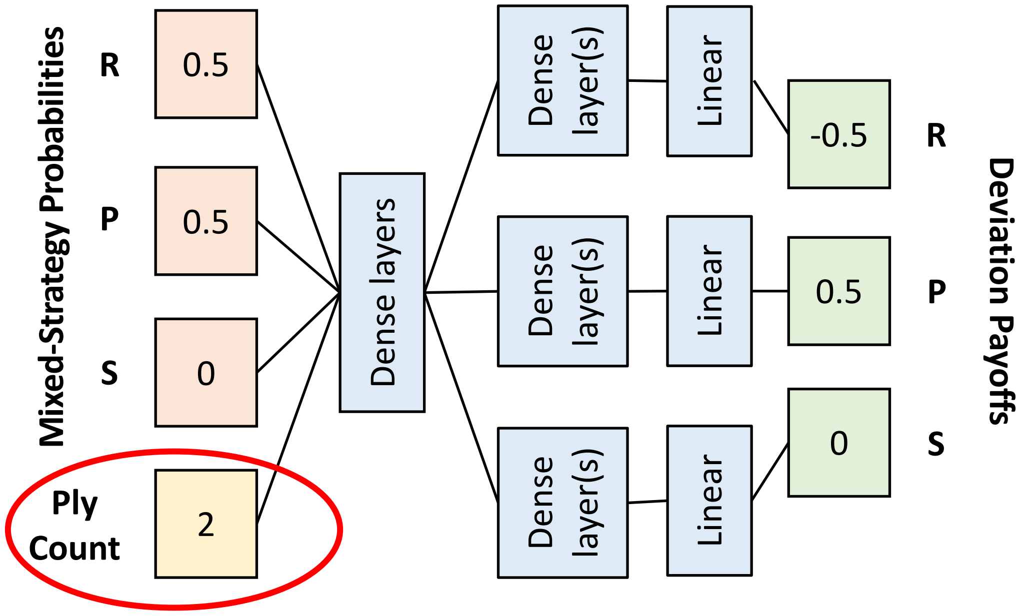
- **Deviation payoff:** the expected payoff a player would receive by deviating or changing strategies, given the mixed strategies everyone else is playing
- **Regret:** the maximum payoff amount any player can gain by deviating to any other strategy
- **Nash equilibrium:** a set of strategies such that no player can gain by deviating
- **ε-Nash equilibrium:** a set of strategies such that no player can gain more than ε by deviating

Outline

- Introduction and Motivation
- Game Theory Background
- **Model**
  - **Approximating deviation payoffs**
  - **Approximating robust symmetric Nash equilibria**
- Analysis: Equilibrium robustness metrics
- Experiments
- Ongoing Work

Model: Approximating deviation payoffs

- **Hypothesis:** the payoffs in a game with x players are similar or related to the payoffs in the same game with $x \pm 1$ players, given a large value of x
- We use a **multi-headed neural network** to learn a mapping from mixed strategy profiles and number of players to deviation payoffs

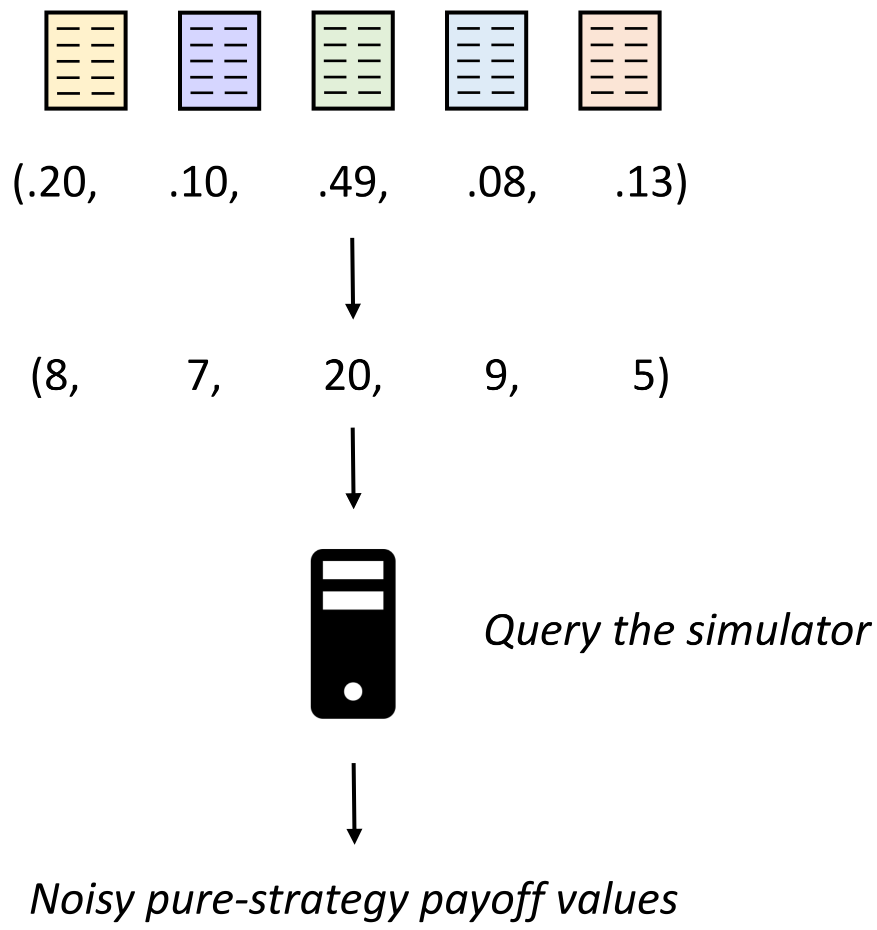


Neural Network Architecture for Variable-Player Learning

Model: Approximating deviation payoffs

Generating training data:

- Random mixed-strategy profile (Dirichlet distribution)
- Sample a pure-strategy profile according to mixed strategy profile for each opponent
- Query simulator for PS payoffs



Model: Approximating robust symmetric NE

- Want to focus learning on areas of simplex where we think there might be approximate Nash equilibria
- Algorithm overview
  - Train network
  - For *i* iterations
    - Run Nash-finding algorithm
    - Sample in neighborhood of candidate Nash and corresponding player counts
    - Retrain network, adding resamples to training data
  - Run Nash-finding algorithm
  - Apply robustness metrics

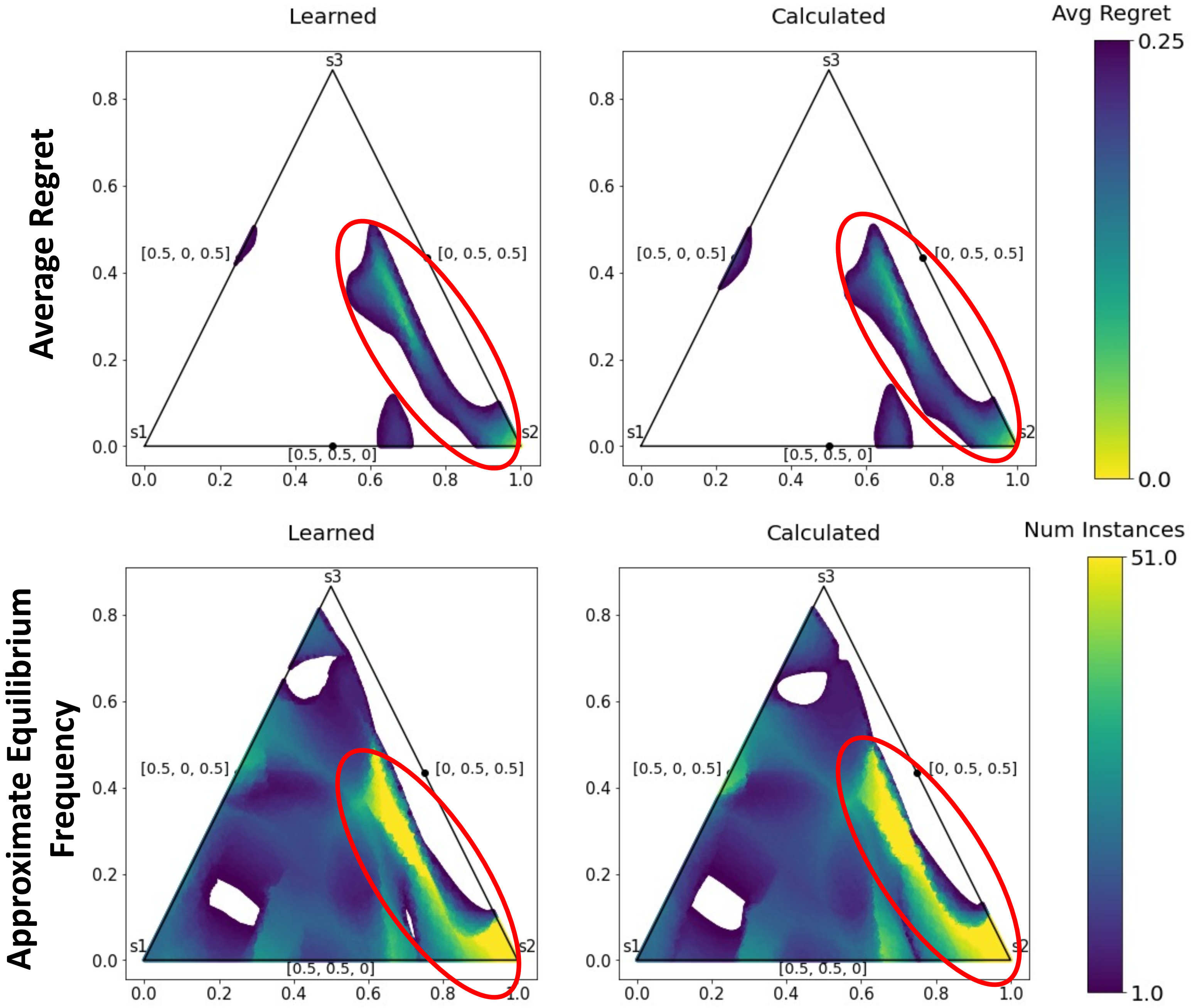
Outline

- Introduction and Motivation
- Game Theory Background
- Model: Approximating robust symmetric Nash equilibria
- **Analysis**
  - **Equilibrium robustness metrics**
  - **Comparison of robustness metrics**
- Experiments
- Ongoing Work

Analysis: Equilibrium robustness metrics

- Typical game-theoretic analysis: find approximate NE in games with fixed number of players
- Finding approximate NE in game with variable number of players is not as straightforward
- Robustness: measure of how well an equilibrium generalizes across all instances of game
- Several proposed robustness metrics
  - Average regret
  - Median regret
  - Max regret
  - Approximate equilibrium frequency

Analysis: Comparison of robustness metrics



Outline

- Introduction and Motivation
- Game Theory Background
- Related Work
- Model: Approximating robust symmetric Nash equilibria
- Analysis: Equilibrium robustness metrics
- **Experiments**
  - **Experimental specification**
  - **Comparison to existing work**
- Ongoing Work

Experiments: Experimental specification

Random Games

- 250 random symmetric games
- Range: 50 to 100 players
- 5 strategies

Evaluation

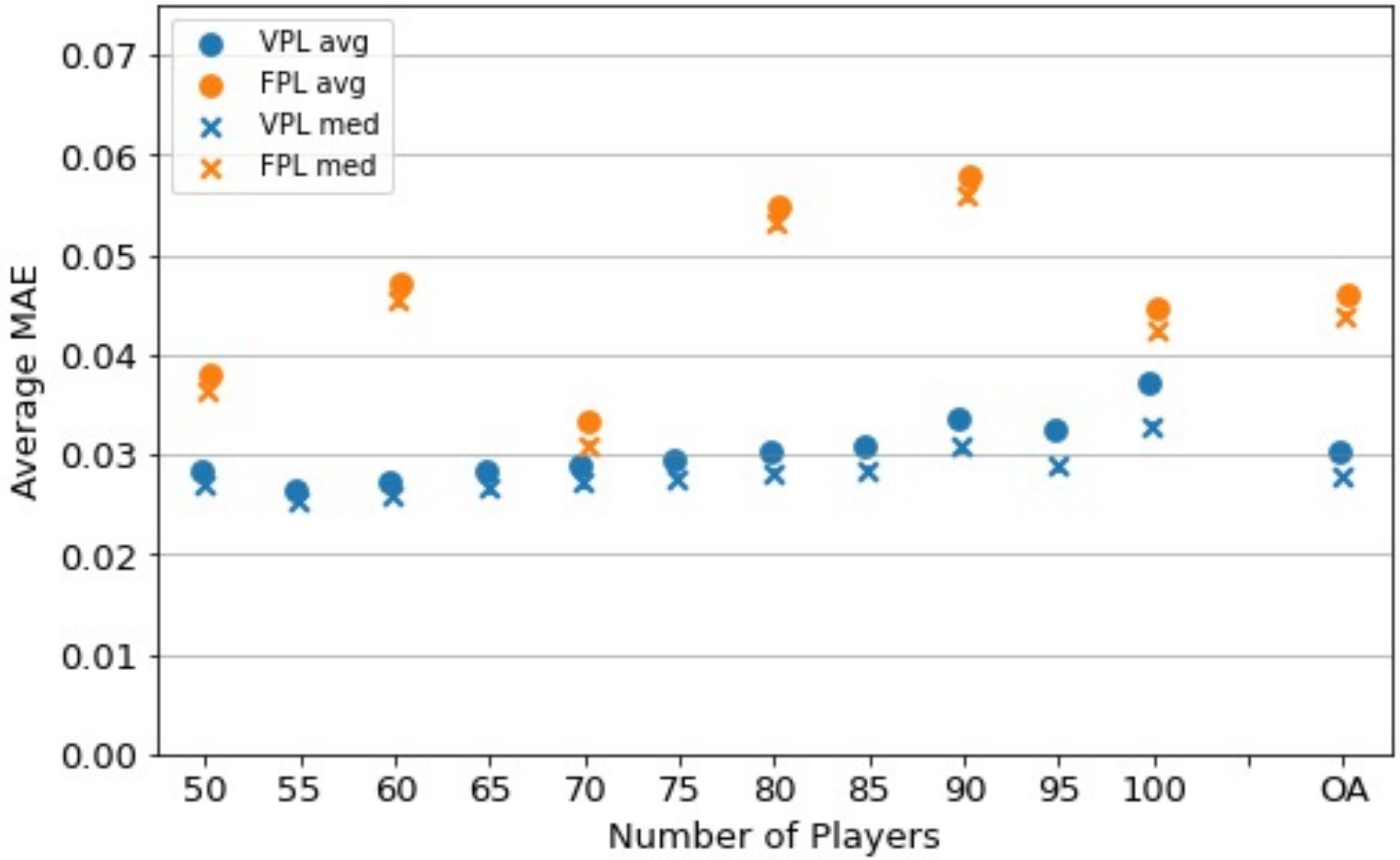
- Average deviation payoff MAE across all strategies

Learning Models

- Fixed-Player Learning (FPL)
  - Train 6 NNs with player counts: 50, 60, 70, 80, 90, 100
- Variable-Player Learning (VPL)
  - Train 1 NN which learns across range of player counts
- Hyperparameters optimized separately

Experiments: Comparison to Existing work

- 60,000 total training examples
  - 10,000 per FPL NN
- 95% confidence intervals
- VPL: evaluated on 495 mixtures * 12 instances * 250 games
- FPL: evaluated on 495 mixtures * 6 instances * 250 games



Outline

- Introduction and Motivation
- Game Theory Background
- Model: Approximating robust symmetric Nash equilibria
- Analysis: Equilibrium robustness metrics
- Experiments
- **Ongoing Work**

Ongoing Work

In the Thesis*

- Evaluate robust approximate NE algorithm
- Scalability experiments
    - Size of range
    - Magnitude of player counts
    - Number of strategies

Future Work

- Evaluate performance on a wider range of games
- Extend this technique to analyze variable-player role-symmetric games
- Extend this technique to vary parameters of simulation environment
- Theoretical guarantees?

*hopefully