

Verifying Truthfulness in Auctions Using Reinforcement Learning

Blaise Swartwood | Preksha Sarda

CSSE490: Algorithmic Game Theory

Ideal Auctions

- ▶ Truthful
 - ▶ Agents have a dominant strategy to bid their true valuation for an object
- ▶ In auction design, we typically assumed truthfulness to begin, and design around it



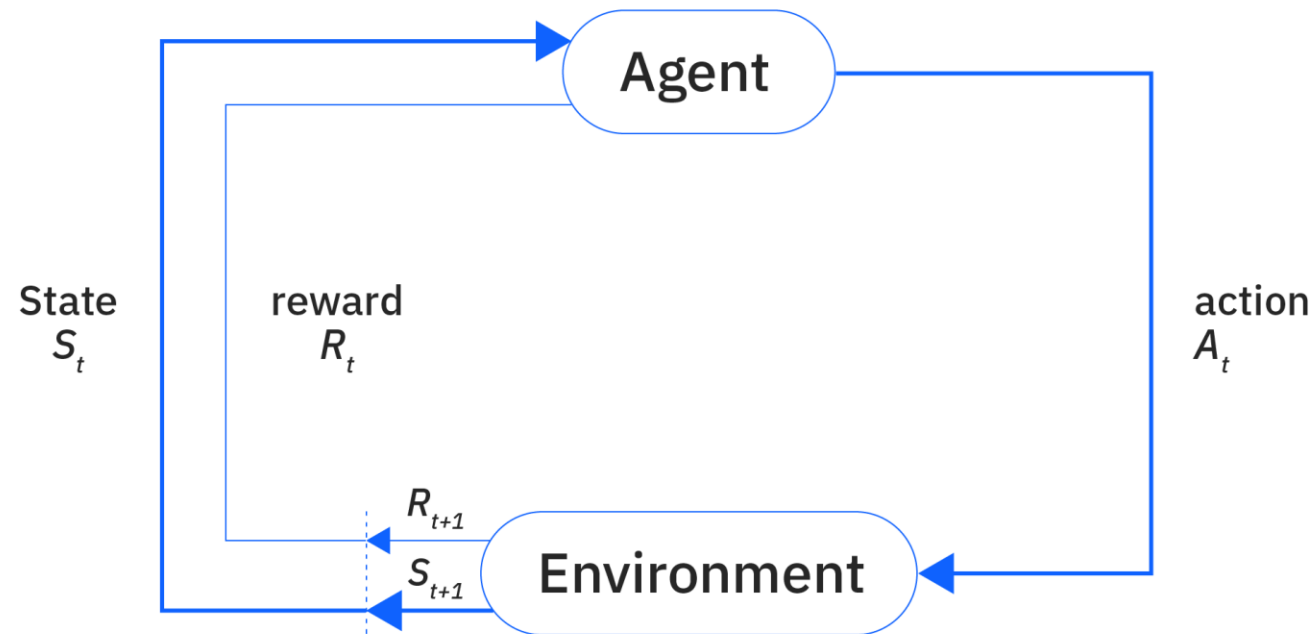
Project Goal

Verify truthfulness in these auctions using reinforcement learning

1. Start with agents bidding random initial values
2. Agent learns how to bid to maximize rewards
3. Show that the learned strategy is truthfulness

Reinforcement Learning

- ▶ Agents explore an environment by taking actions and **receiving rewards**
- ▶ Goal is to **maximize long term accumulation of rewards**



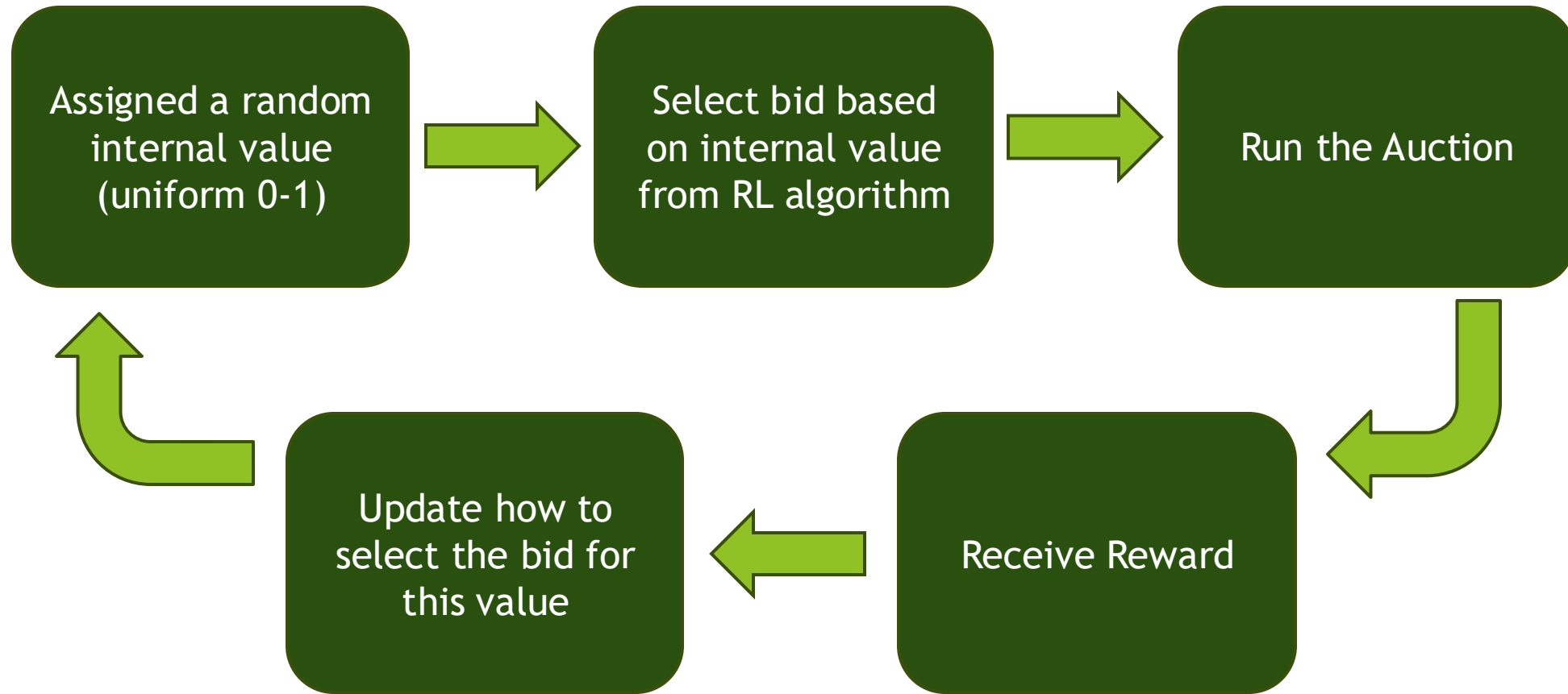
RL vs Best Response Dynamics

Aspect	Best Response Dynamics	Reinforcement Learning
Assumptions	Knows payoff structure and others' strategies	Doesn't know environment or other agents
Adaptation	Computes best response analytically	Adjusts from reward feedback
Learning	No learning; purely reactive	Yes; learns over time through reward signals
Exploration	None (calculated)	Required to discover optimal actions

Design

Diagram | Methods | Auctions

Diagram

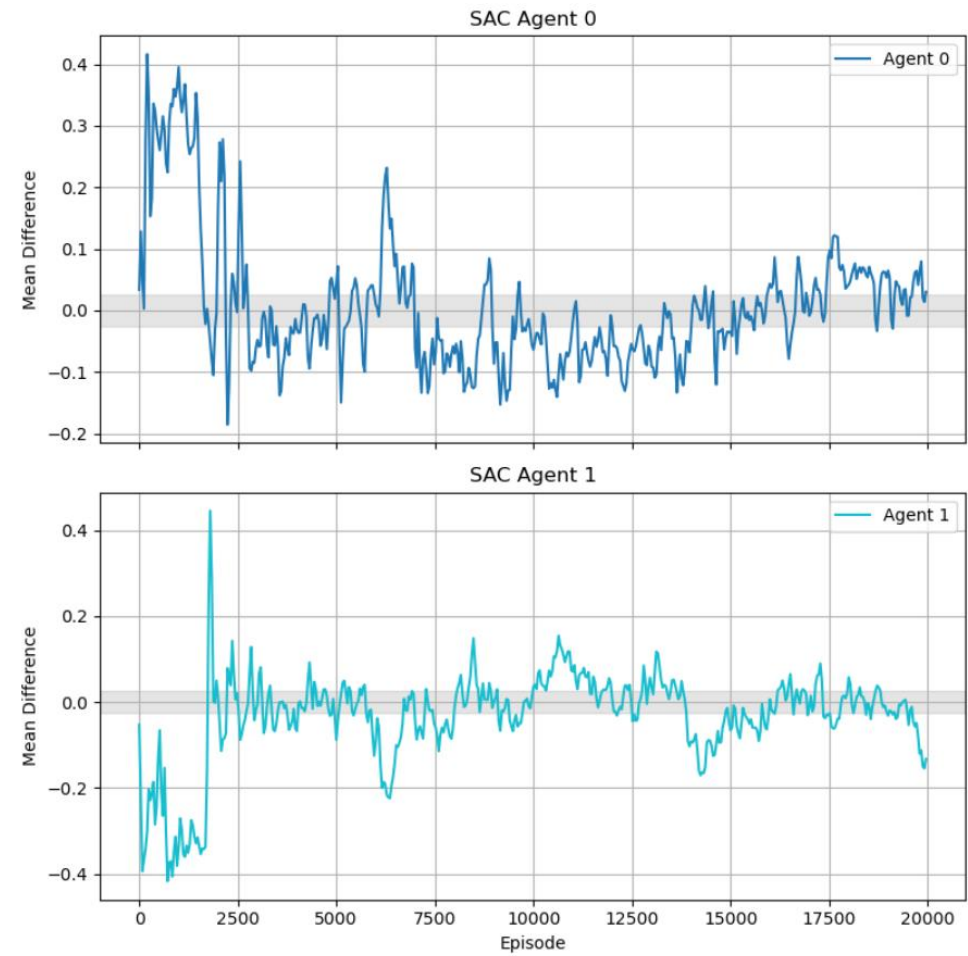
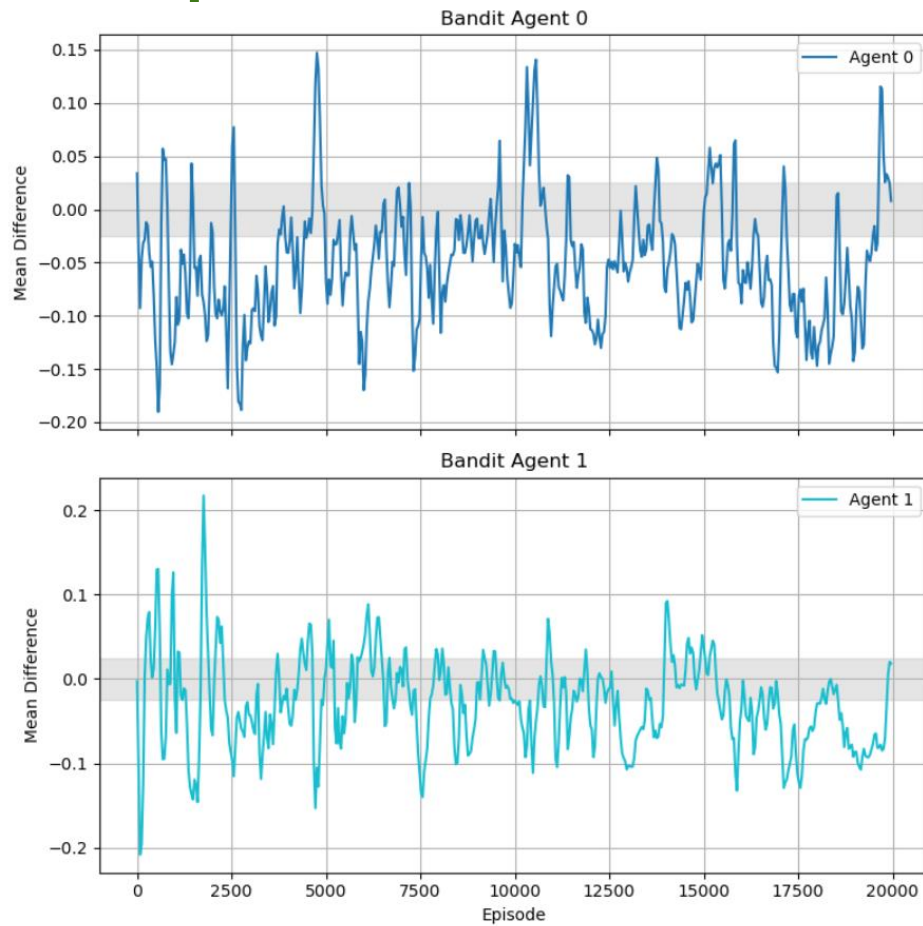


Reinforcement Learning Methods

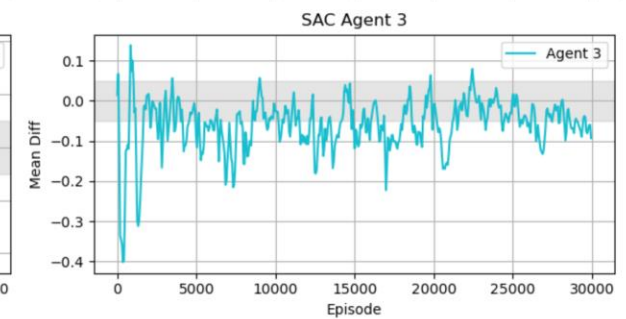
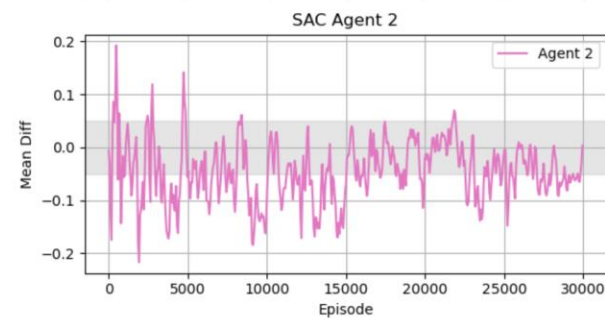
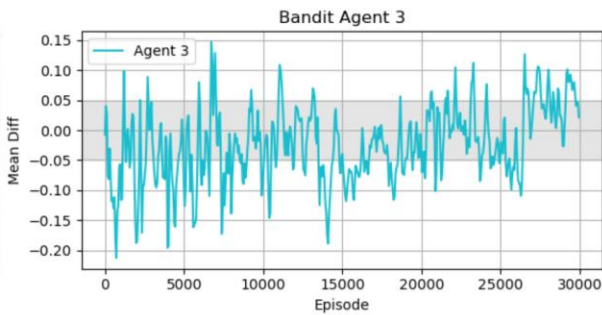
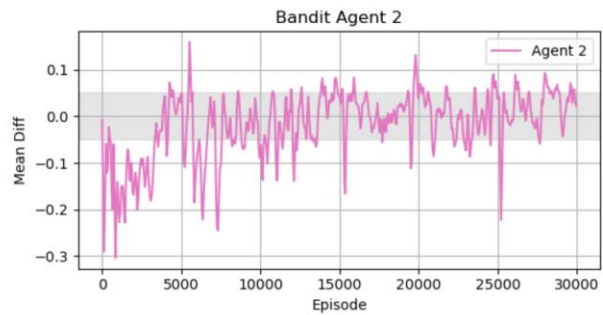
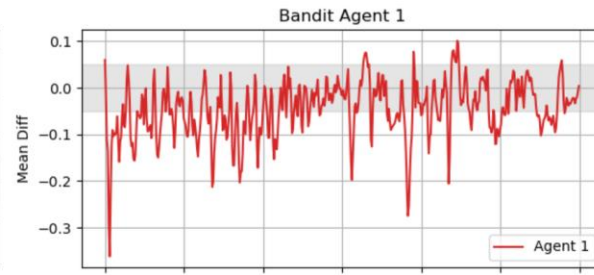
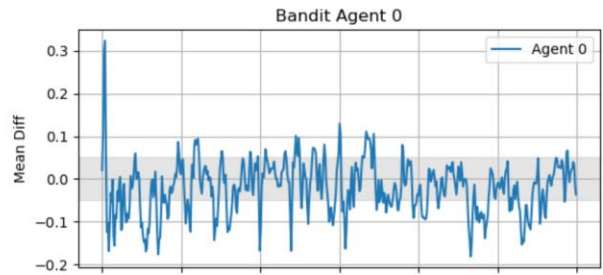
Method	Q-Learning	Thomson Bandit (Supervised)	SAC
Type	Dynamic Programming	Contextual Bandit	Deep RL (actor-critic)
Exploration	ϵ -greedy	Bayesian (dropout)	Entropy-based (stochastic policy)
Action Space	Discrete Bids	Continuous Bids	Continuous Bids

Results

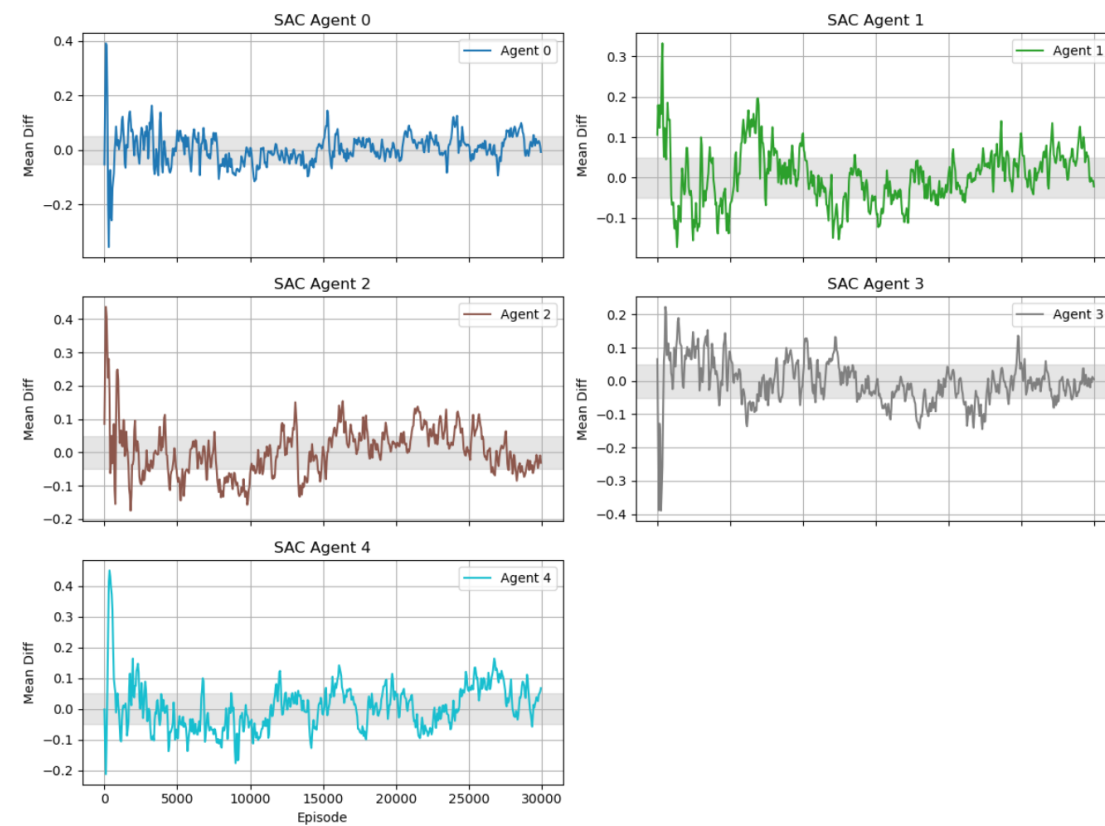
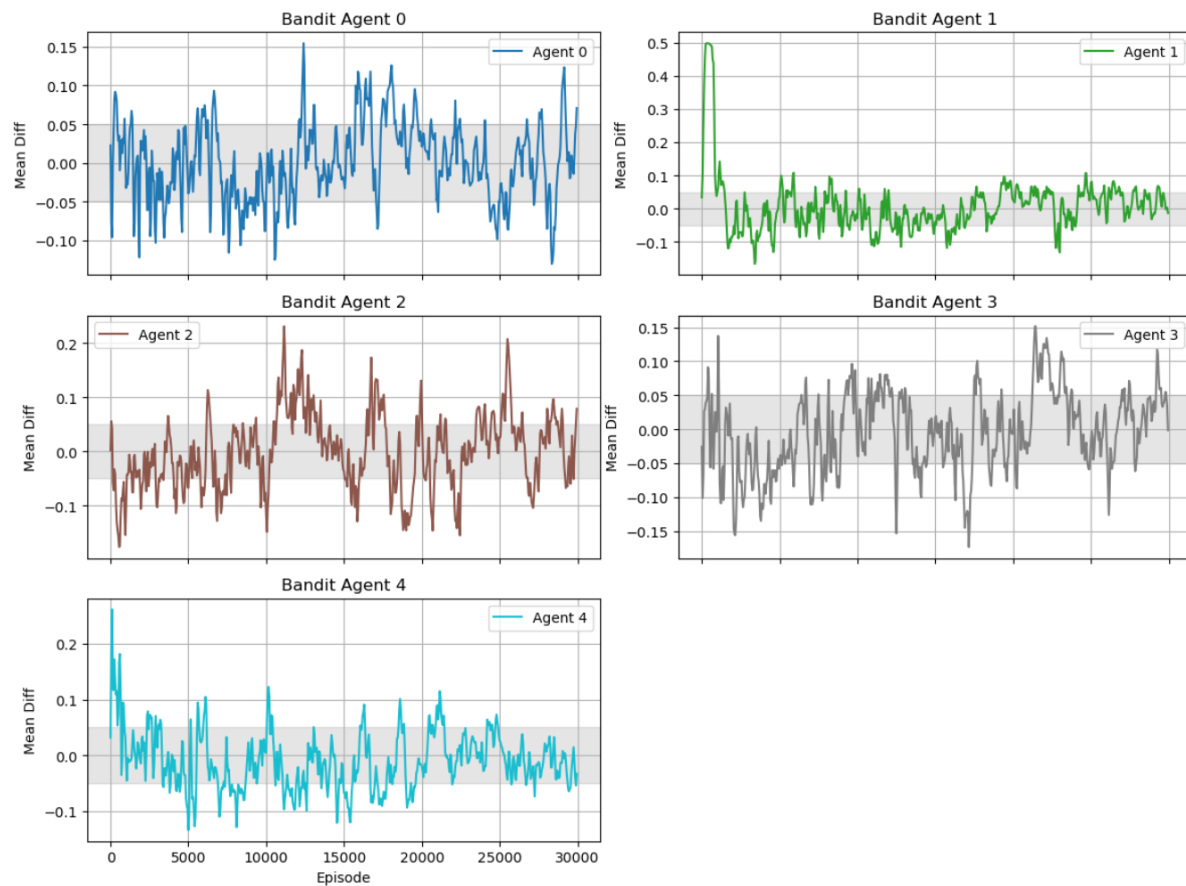
Simple



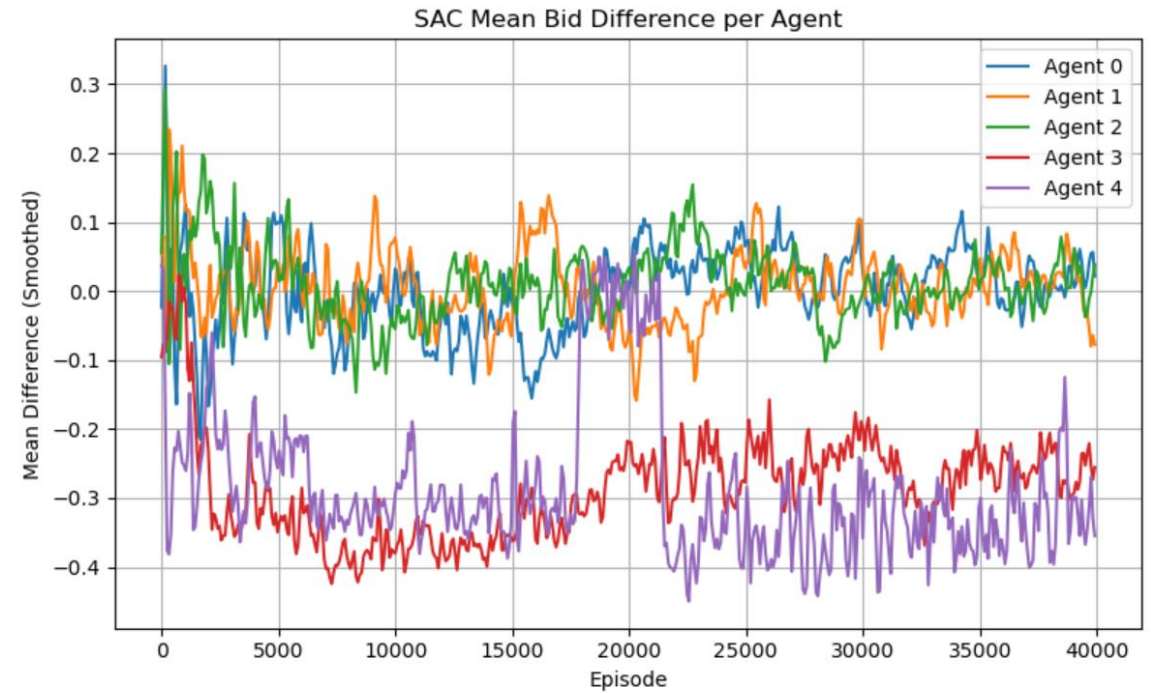
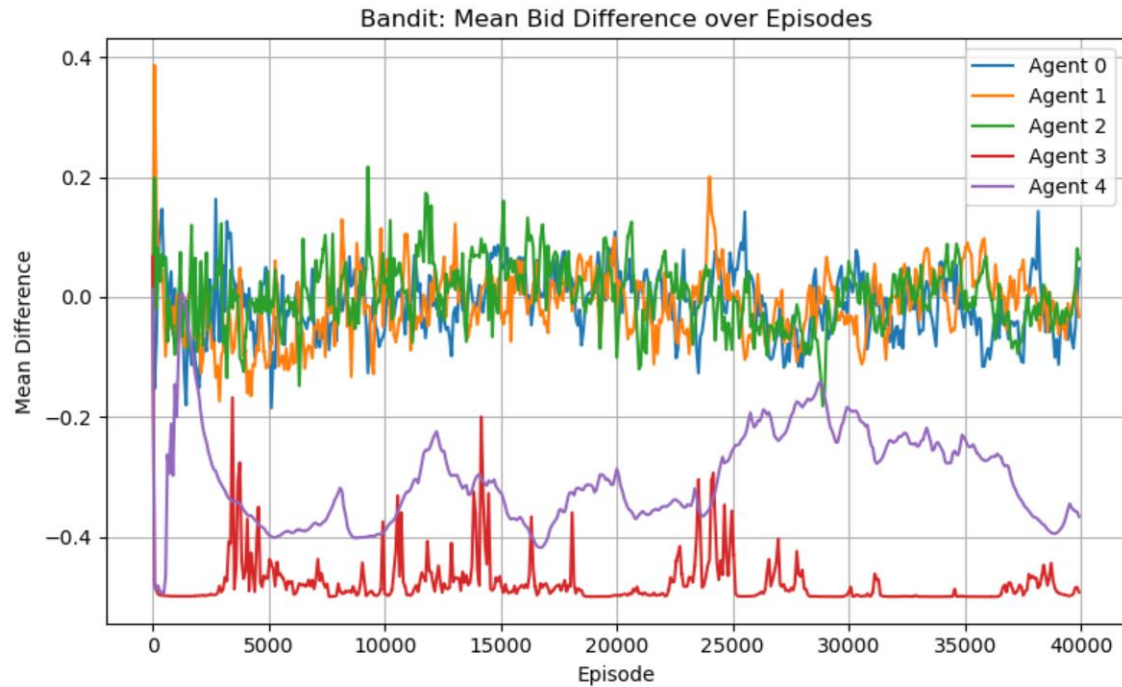
Multi-Agent



Multi-k-Item



Adversarial Agents



Results

- ▶ Agent bids did converge towards internal values
- ▶ **Imperfect convergence** is a symptom of RL algorithms
 - ▶ Sensitive to initial conditions and random selection
- ▶ Shows truthful bidding is the dominant strategy
 - ▶ Regardless of how other agents bid
 - ▶ Most apparent in adversarial auction design

Future Work

- ▶ **Took a truthfulness approach:** Could explore maximizing social welfare, maximizing revenue by ‘learning’ payment structure or optimal reserve price
- ▶ **Move towards best response dynamics:** Use RL to learn how one agent responds to another
- ▶ **Look at combinatorial auctions:** implement combinatorial auctions (with and without additive valuations) in both a discrete and continuous space



Thank you