Planning with Quantized Opponent Models

Xiaopeng Yu, Kefan Su, Zongqing Lu

School of Computer Science, Peking University

NeurIPS 2025

Introduction

Planning with Quantized Opponent Models (**QOM**)

- Multi-agent settings with unknown opponent strategies
- Challenge: balance adaptability and computational tractability
- Existing methods:
  - Type-based (handcrafted, limited scalability)
  - Model-free (sample inefficient)

Introduction

Planning with Quantized Opponent Models (QOM)

- Compress opponent policy space via a quantized autoencoder
- Maintain Bayesian belief over latent opponent types
- Integrate belief directly into Monte-Carlo Planning

Introduction

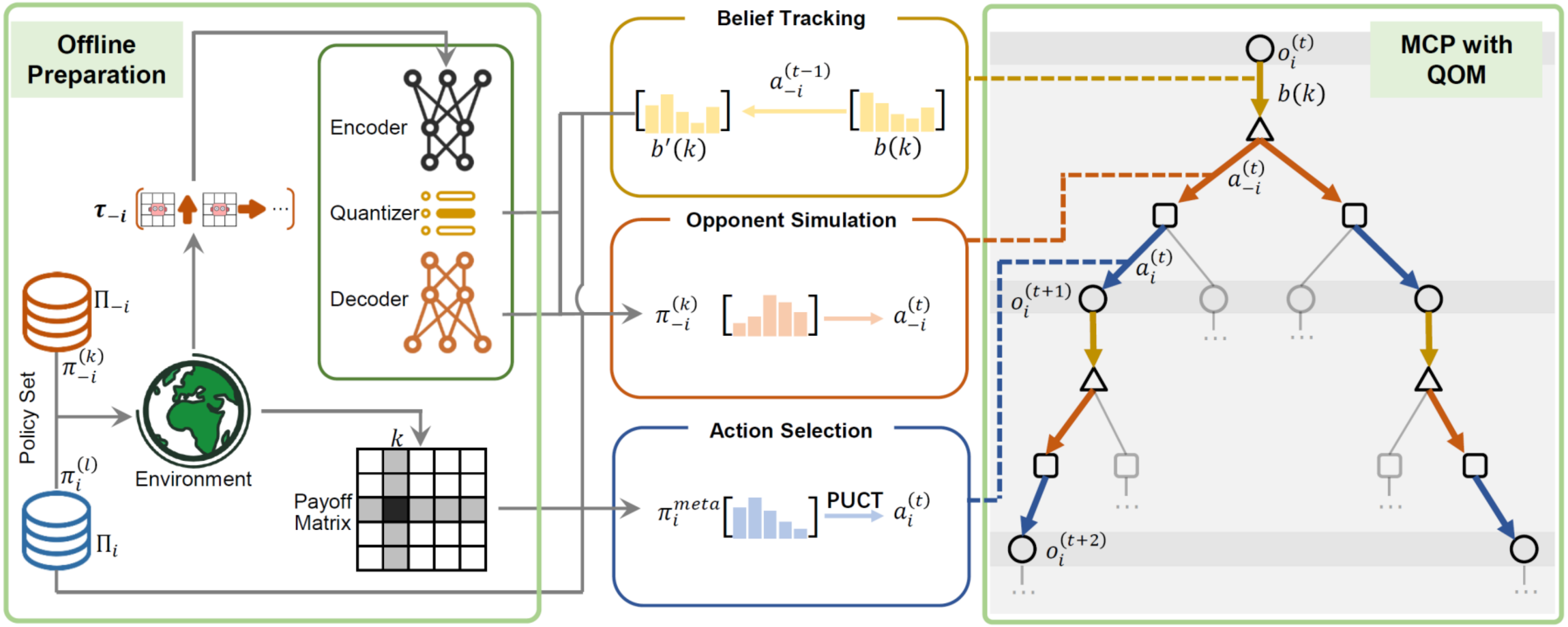
Intuition

- Opponent policy space is huge
- Most opponents share a few behavioral “types”
- Quantization turns continuous uncertainty into discrete belief

Method

Framework Components

- Quantized Autoencoder (offline type discovery)
- Belief Tracking (Bayesian inference)
- Belief-Aware MCP (online decision)

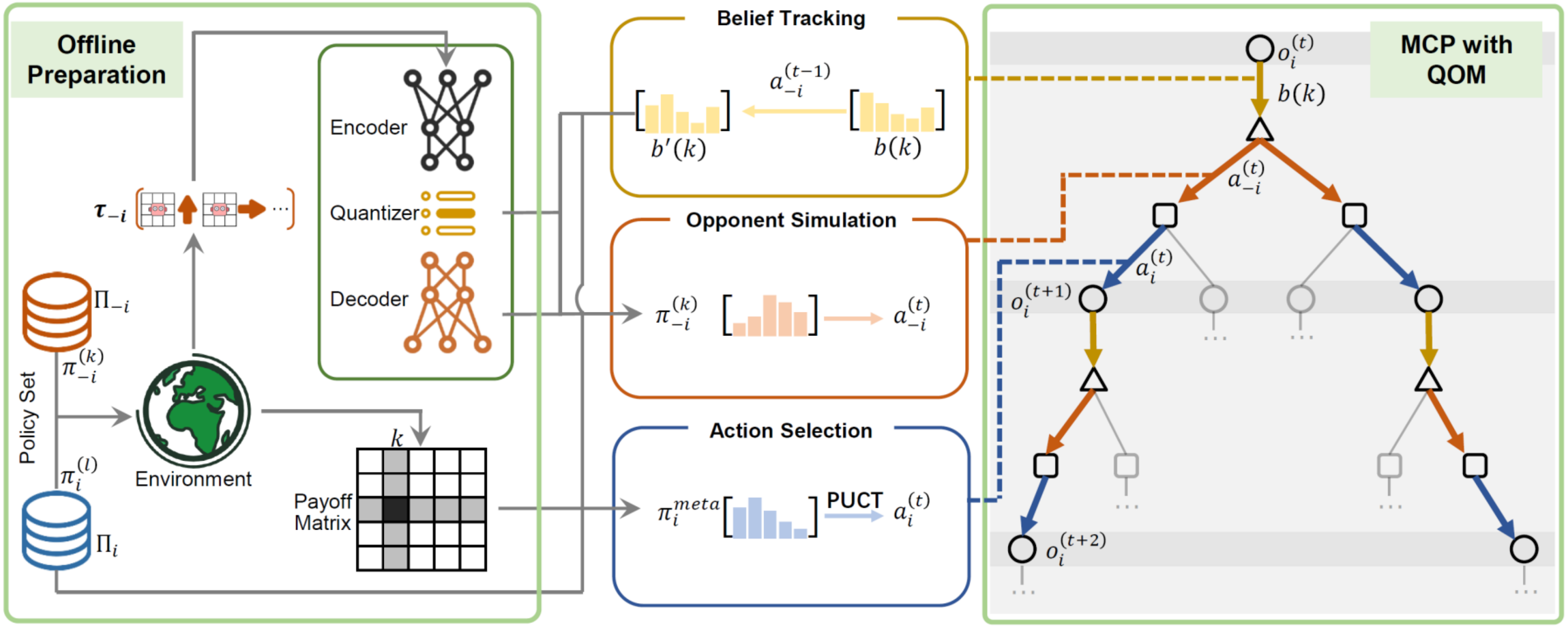


Method

Quantized Autoencoder

Offline Type Discovery

- Encode trajectories into latent embeddings
- Quantize embeddings into K discrete types
- Decode to reconstruct opponent policy likelihood

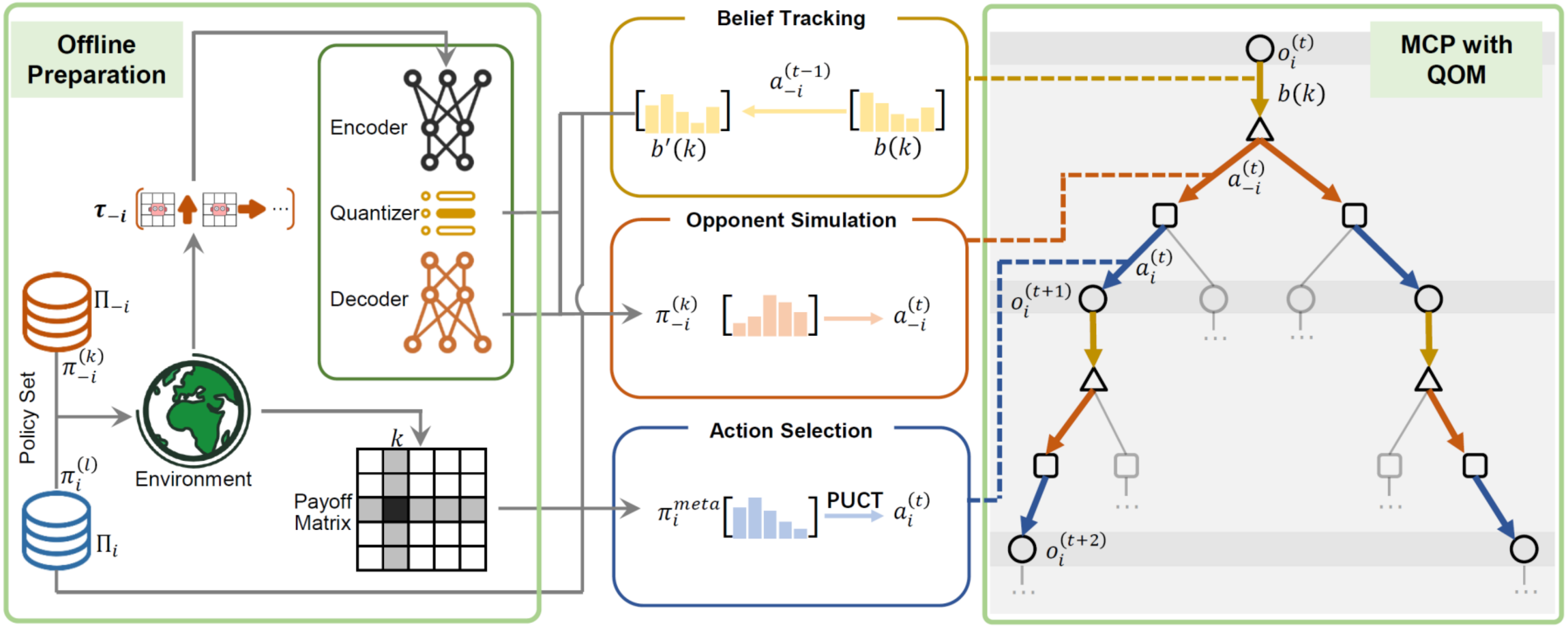


Method

Belief Tracking

Online Bayesian Inference

- Maintain belief $b_t(k)$ over opponent types
- Update using likelihood from decoder
- Smooth updates to avoid overconfidence

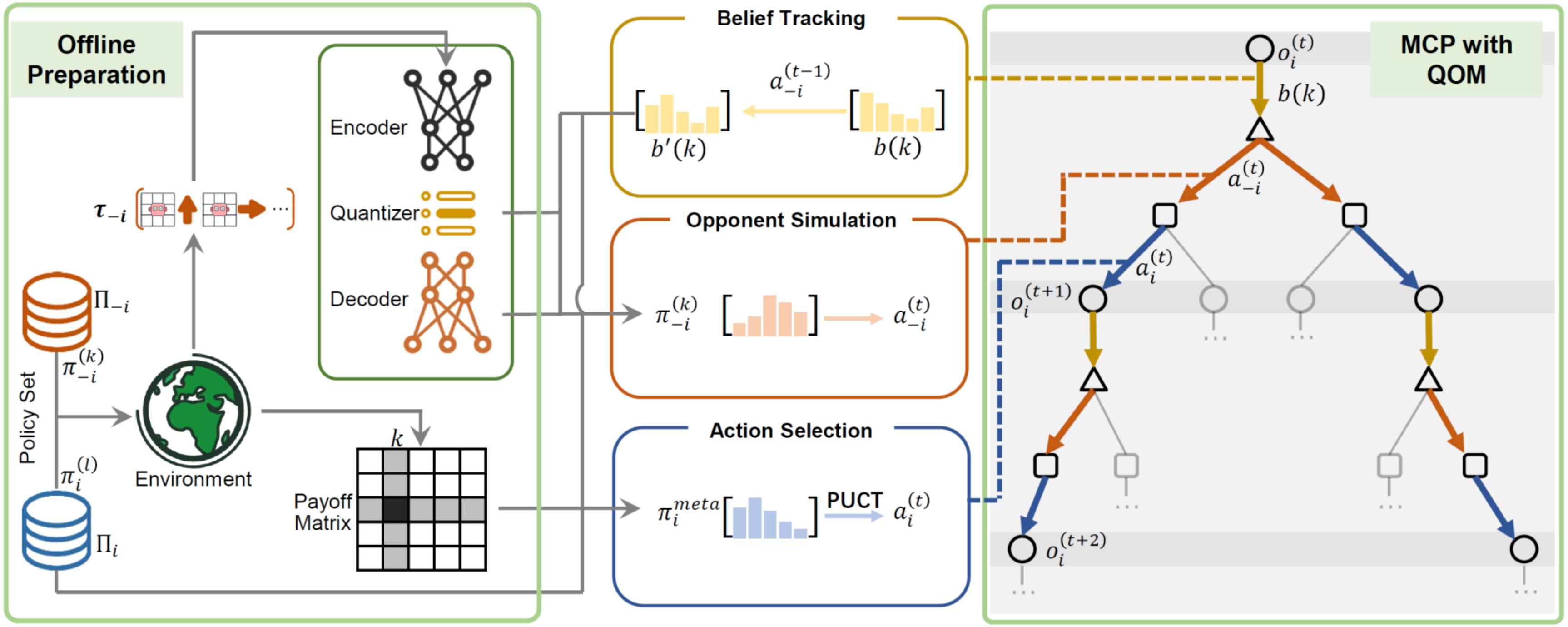


Method

Meta-Policy Construction

Belief-Weighted Meta-Policy

- Compute soft best-responses for each type
- Mix responses using current belief
- Enables adaptive and robust action selection



Method

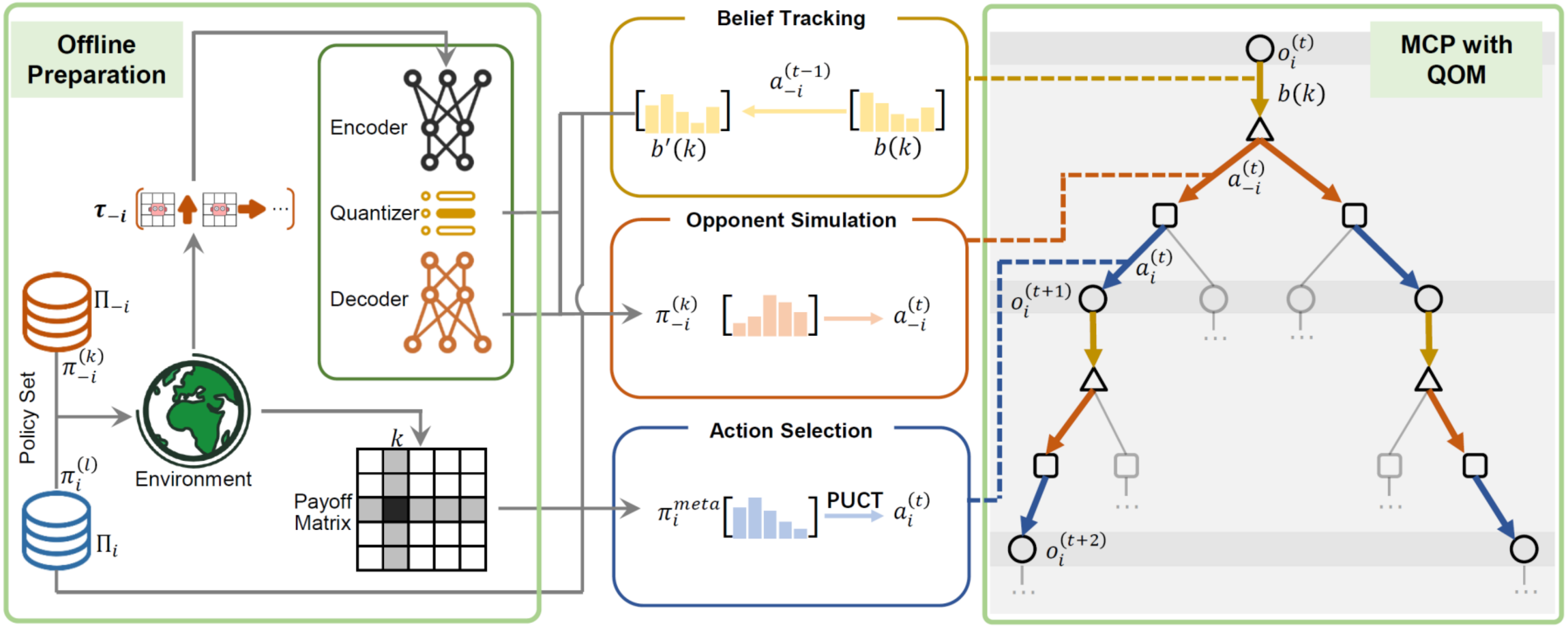
Planning with QOM

Belief-Aware MCP

- Integrate belief into PUCT-based tree search

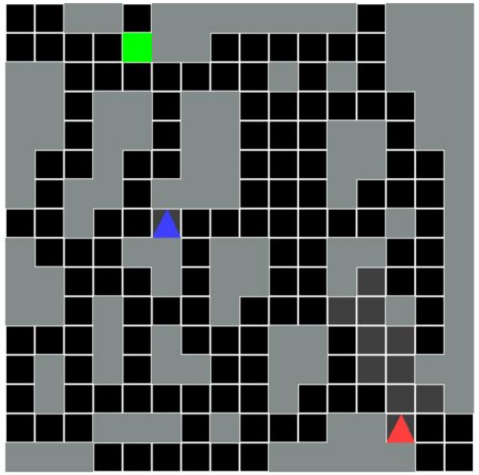
- Integrate belief into PUCT-based tree search

- Update belief along simulated rollouts

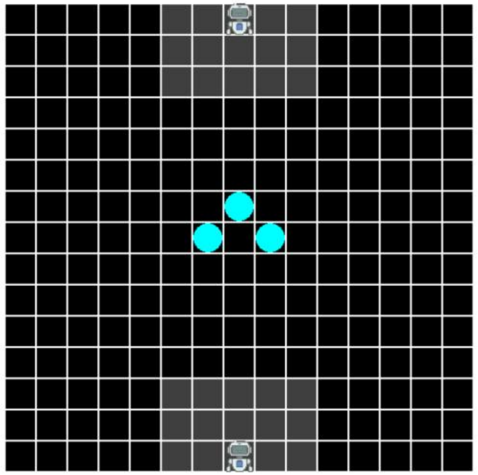


Experiment

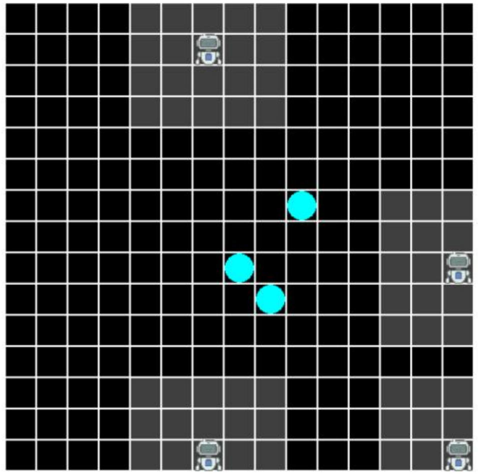
Test environments



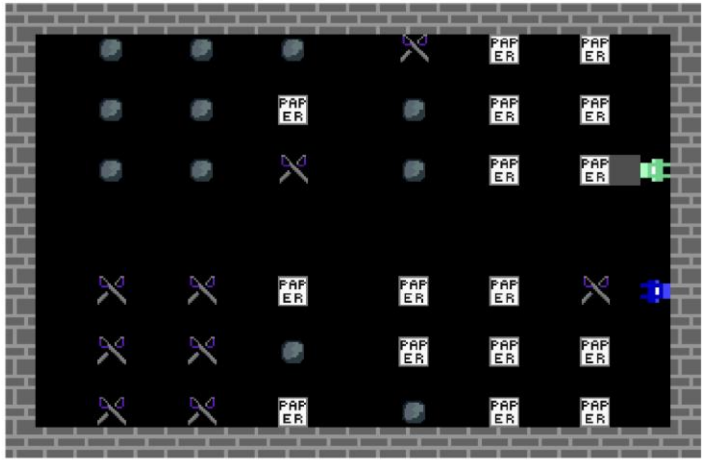
(a) PursuitEvasion



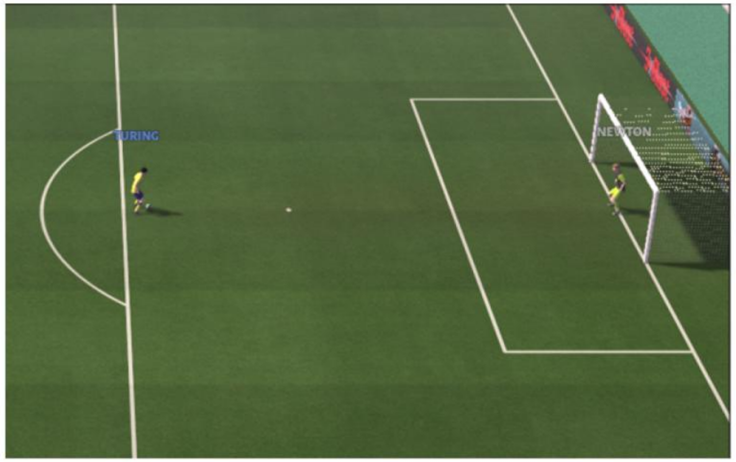
(b) PredatorPrey(two-agents)



(c) PredatorPrey(four-agents)



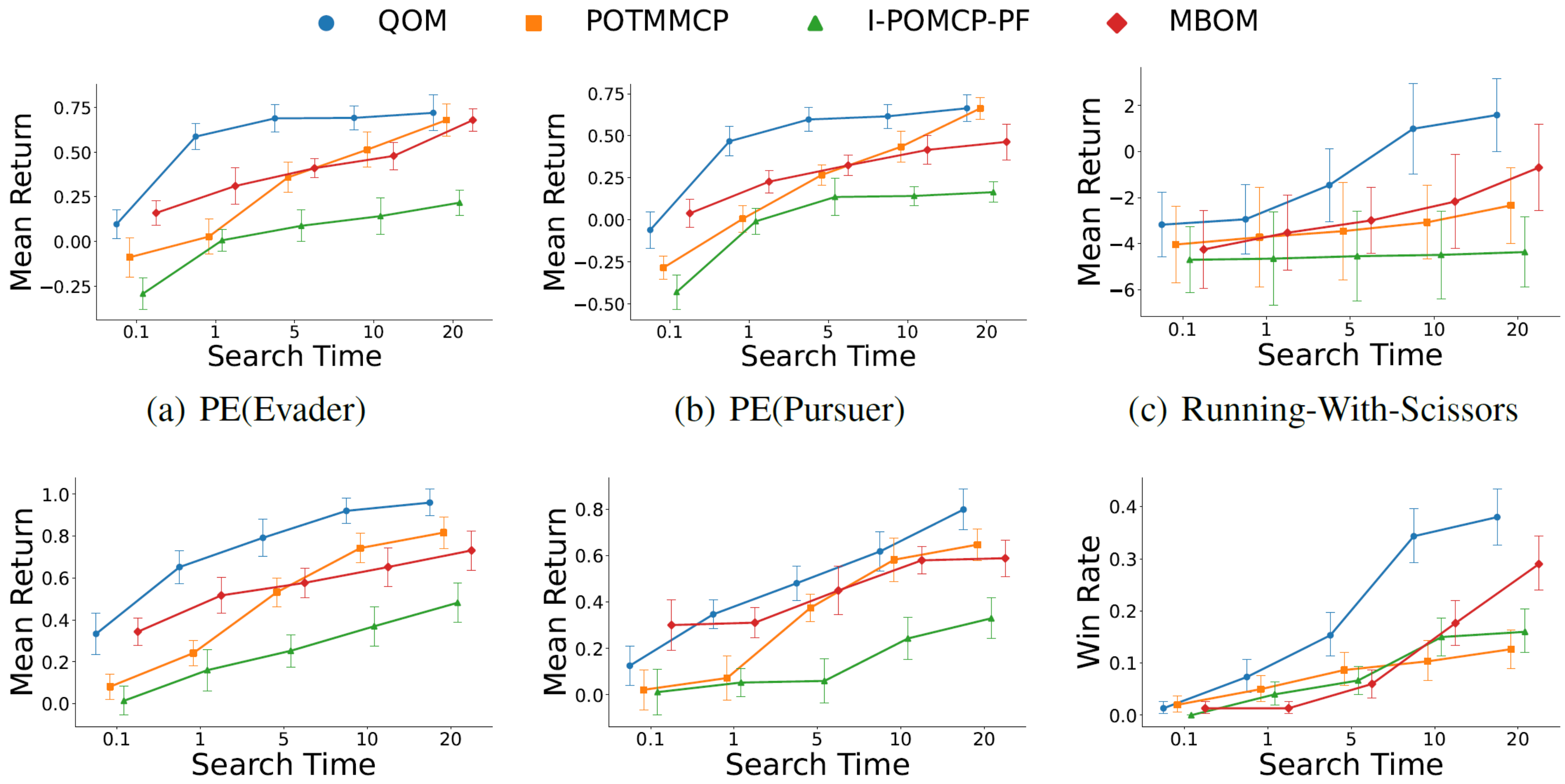
(d) Running-with-Scissors



(e) One-on-One

Experiment

Results (Static Opponents)

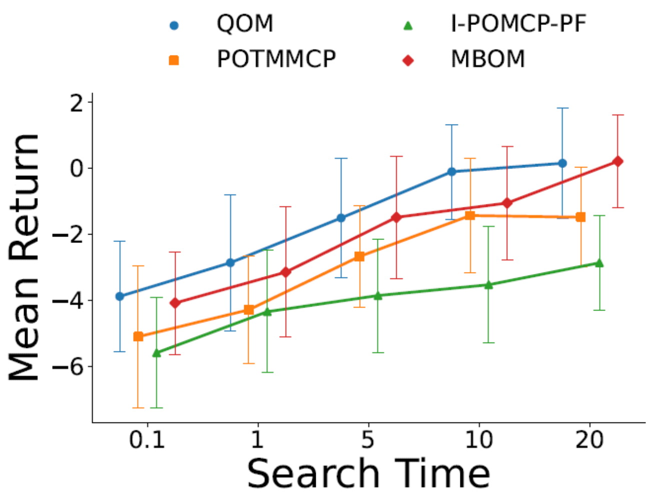


(a) PE(Evader) (b) PE(Pursuer) (c) Running-With-Scissors

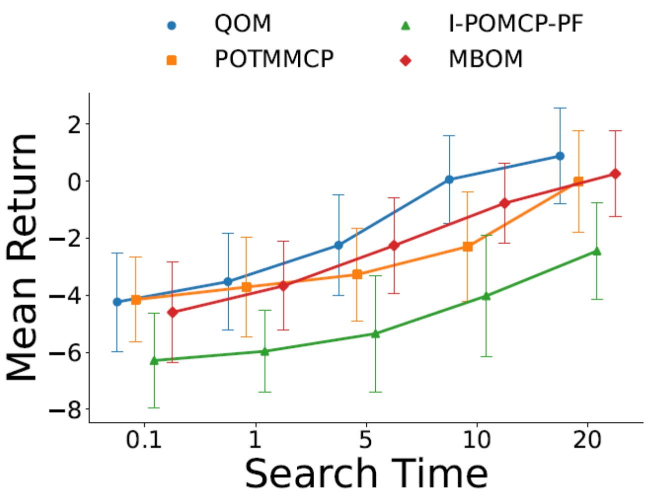
(d) PP(two-agents) (e) PP(four-agents) (f) One-on-One

Experiment

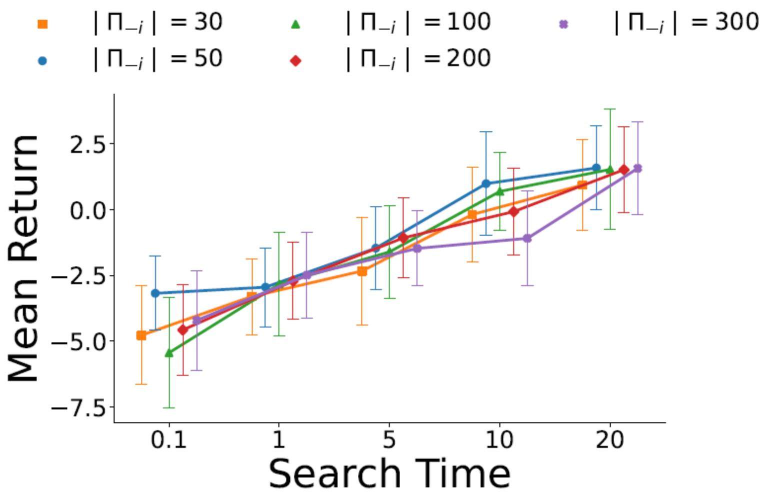
Results (Dynamic & Unseen Opponents)



(a) Adaptation



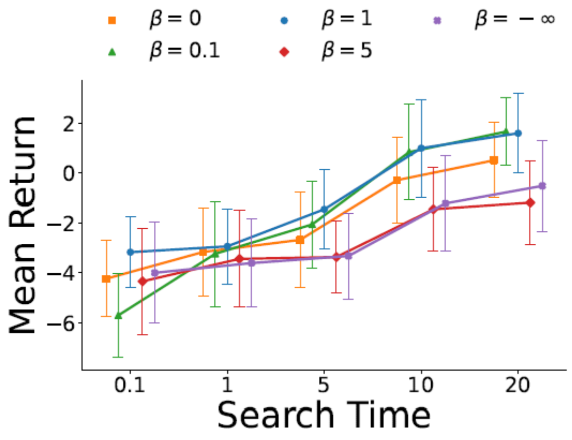
(b) Generalization



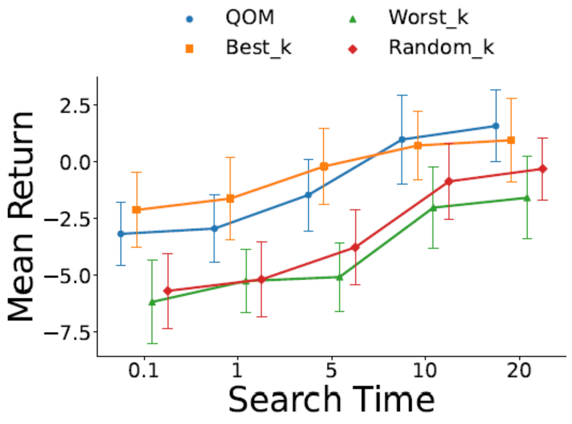
(c) Scalability

Experiment

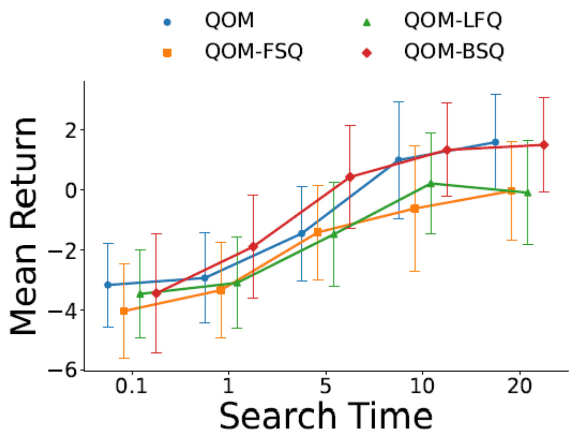
Belief Dynamics and Ablation



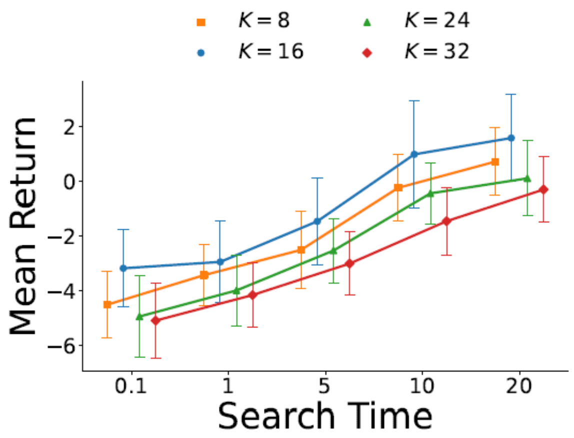
(a) β



(b) $\tilde{\pi}_{-i}^{(k)}$



(c) Quantization



(d) K

Summary

01 Compact latent opponent representation via quantization.

02 Bayesian belief enables uncertainty-aware planning.

03 Unified framework for scalable, interpretable opponent modeling.

Thank you for listening.