Optimal Distribution of Tasks in Human-Autonomy Teams: Decentralized Reinforcement Learning

Haochen Wu1, Amin Ghadam1, A. Emrah Bayrak2, Bogdan I. Epureanu1, Victor Paul1, John Brabbs1, Jillyn Alban2, Mert Eglímez2
1University of Michigan, Ann Arbor 2Stevens Institute of Technology 3Ground Vehicle Systems Center 4Veoneer-Nissin

Introduction and Motivation

• Current approaches focus on homogenous team coordination methods or on qualitative human factors with autonomy being subordinate members.
• Human-autonomy teams can achieve superior effectiveness in complex operations by combining efficiency of autonomy and reliability of humans.

Goal

Fundamental Research Questions

• How to synergistically distribute tasks in a complex mission among humans and autonomous agents to maximize mission effectiveness?
• How do the unique benefits and limitations of human and autonomy affect the solution approach and performance?

Objectives

• Develop an artificial intelligence framework for human-autonomy teams to learn collaborative decisions through reinforcement learning
• Create real-time decentralized multiagent coordination strategies under uncertainties in adversarial environments

Approach – Problem Formulation

Overview

• Human and autonomous agents: individual decision-making processes
• Benefits & limitations of heterogeneous teams quantitatively investigated

Extended Decentralized Partially-Observable Markov Decision Process

• Goal: maximize team performance described by expected future reward
• Task allocation: decisions based on partial observations & communication
• Agents: same types of attributes and heterogeneity in capability levels
• State: factored into tasks described by severity level
• Environment: dynamic and stochastic

Noisy Rational Model for Human Agents

• Humans take risks and make suboptimal decisions.

\[ \Pr(a|s) = \frac{\exp(\theta \cdot Q(s,a))}{\sum_{a' \in A} \exp(\theta \cdot Q(s,a'))}, \theta \in [0, \infty) \]

Approach – Learning with Beliefs

Belief Update – Bayesian Approach

• Agents maintain probabilities over the task state (beliefs).
• Belief summarizes past action-observation experiences.
• Each agent takes action given the belief under the optimal policy.

Decentralized Deep Q-Learning for Beliefs

• Approximate the optimal Q-Value $Q^*(b,a)$ for continuous belief state
• Optimal policy $\pi(b) = \arg\max_{a \in A} Q^*(b,a)$
• Minimize loss between predicted and target Q-Values: Bellman equation

\[ L = \mathbb{E}(f_{\theta}(b) + \gamma \max_{a',\pi(b)} Q(b',a') - Q(b,a))^2 \]

Results

• A fire-fighting and rescue operation with 4 tasks and 3 agents
• Human agents have lower capabilities & noisy rational decision process.
• Humans bring more uncertainties to the team.
• Human-autonomy teaming ($n_h = 1$) outperforms human-only ($n_h = 3$) and autonomy-only ($n_a = 0$) teams.
• Human creativity occasionally helps learning process.

Conclusions

• Modeled team coordination as extended Dec-POMDP with capabilities and team-level communication explicitly defined
• Approached optimal task allocation by belief-oriented deep Q-Learning
• Heterogeneous teams outperform homogenous teams with trade-offs.

Future Work

• Perform sensitivity analysis on characteristics of the decision-making models with focus on human factors
• Increase operation complexity and create a hierarchical AI framework
• Develop a game-based platform to quantitatively analyze human-autonomy teams with actual human decision-makers