Multi-Arm Bandits for Automated Decision Making

Steve McGuire, Nisar Ahmed: CU Boulder
P.M. Furlong, Terry Fong: NASA Ames Research Center
Research Vision: Autonomy Failure Recovery

Scenario: Autonomy isn’t perfect and will eventually fail. Then what?
Problem: How does a robot in need of help decide who to ask?

Model of an Ideal Learning Robot

- Planner
- Task Executive
- Machine Learning
- Assistant Selector
- Embedded Agent
- Actor

Key Research Challenges:
- Individual Human Modeling
- Signals of Opportunity
- Long-Term Operations

Mission: Joint Exploration

S. McGuire, P.M. Furlong, T. Fong, N. Ahmed
2018 NASA Space Autonomy Workshop
Plan of Action

Method: Reinforcement Learning via Contextual Multi-Arm Bandits

Implementation Levels:

▶ Assessment in simulation: Failure is Not an Option: Policy Learning for Adaptive Recovery in Space Operations, IEEE RA-L, Jan 2018
  Signals of opportunity are useful!

  Works with real humans!

▶ Practical Problems: In progress
  ▶ Initialization
  ▶ Dynamic Human Models

Future Impact:

▶ More agile missions
▶ Collaboration with less structure
▶ Enabling for multi-robot ops
  ▶ Better use of human assets
  ▶ Leverage onboard autonomy

Human-in-the-loop experiments showing improvement when contextual information’s impact is learned (red, lower is better)

S. McGuire, P.M. Furlong, T. Fong, N. Ahmed
2018 NASA Space Autonomy Workshop