AA228 Final Project Proposal

Autonomous Decision-Making for Lunar Regolith Sample Collection under Energy and Environmental Uncertainty

Martin Yitao bpalalto@stanford.edu
Jason Gunn jwgunn@stanford.edu
Loïc Poisson lpoisson@stanford.edu

Goal

Future lunar rovers will need to autonomously search the lunar surface for high-value regolith samples that can be utilized as in-situ construction materials. We formulate this challenge as a decision-making problem, where the objective is to maximize the total value of collected samples while operating under constraints imposed by limited energy resources and uncertain environmental conditions. Solving this problem would enable more autonomous and efficient resource utilization on the Moon, reducing reliance on Earth-based control and paving the way for sustainable long-term lunar habitation and construction.

Decision Making

We model the problem as a Partially Observable Markov Decision Process (POMDP), since the rover has limited and uncertain knowledge of the terrain and resource distribution. The rover would be affected by the environment during its mission process. For example, the onboard solar panels could charge the rover in the daytime. A more specific modeling of the problem is stated below:

- **State**: rover position, remaining energy, remaining payload capacity, terrain type, estimated resource richness, and time.
- **Actions**: move to a new grid cell, collect a sample, recharge, or terminate the mission.
- **Transitions**: probabilistic, affected by energy consumption rate and terrain slope.
- **Rewards**: positive for valuable sample collection, negative for energy use/time spent, as well as operating within hazardous gridspaces.
- **Observations**: noisy sensor readings of terrain type and regolith quality.

The decisions taken influence future opportunities by altering available energy and the rover's knowledge of the environment, making the problem sequential.

Sources of Uncertainty

We identify the uncertainty to be as the following:

- Stochastic elements focused around layout of hazardous cells and associated penalties for them (crater, rocks, etc.) as well as the materials and their value.
- Random chance of rover undergoing an accident (lunar events) and making a movement different from what it determined was optimal at its current state
- Without knowledge of the entire existing state of the M x N grid (grid state can change over time?) but rather can see X cells away from itself in any direction

Further Improvements as time allows:

- If we can find relative cost comparisons between payloads can execute a cost analysis on configurations of a rover (more exquisite sensor payloads), as well as quantity of rovers if there is a benefit
    - Add a power budget where more exquisite payloads require more power but allow for further optical sensing of the gridspace.
    - Upfront cost of an additional rover may not increase value gained by a sufficient amount to warrant its implementation, but maybe two cheaper (less exquisite) rovers would.
- Power budget could also be mutually exclusive to the cost comparisons, and add another time-based requirement for the rover. Could also add areas further away from the value regions where the rover can recharge.

Some potential methods to solve this problem would be to implement Q-learning or Deep Q-learning. We could first implement at the simplest case to build the model and explore how introducing more uncertainty impacts the model.