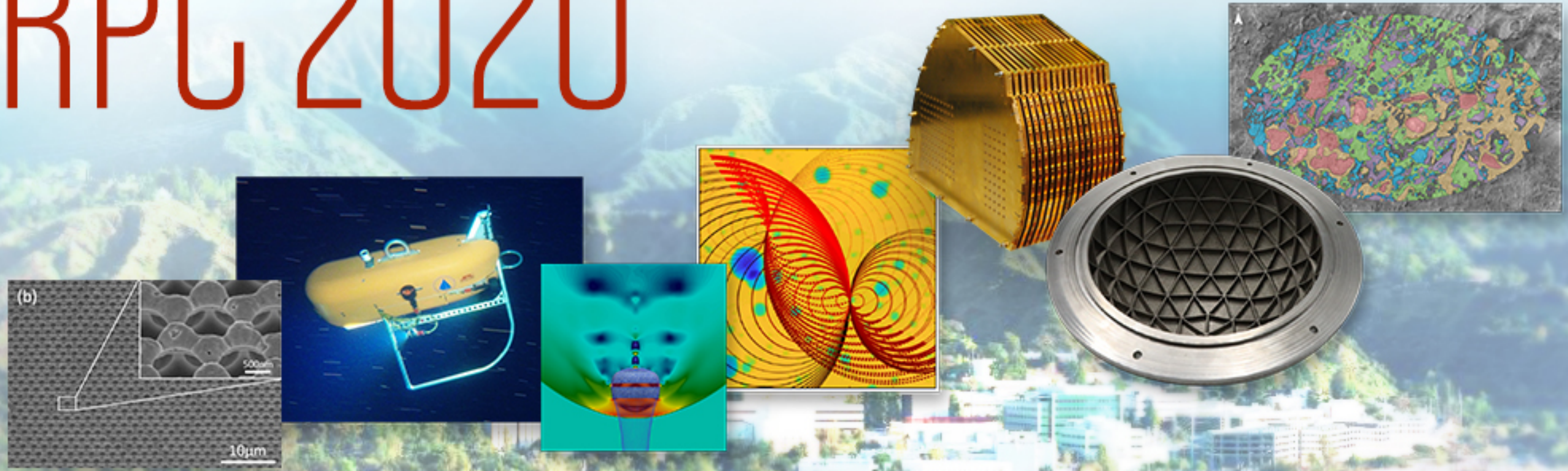
RPC 2020



Virtual Research Presentation Conference

Autonomous planning under uncertainty with infrequent, periodic state observations

Principal Investigator: Federico Rossi, Mobility and Robotics Systems Section (347)
Co-I: Dylan Shell, Department of Computer Science & Engineering, Texas A&M University
Program: Spontaneous Concept

Assigned Presentation # RPC-114

Jet Propulsion Laboratory
California Institute of Technology

Tutorial Introduction

Goal: design algorithm to solve autonomous planning problems under uncertainty with infrequent, periodic state observations.

Abstract

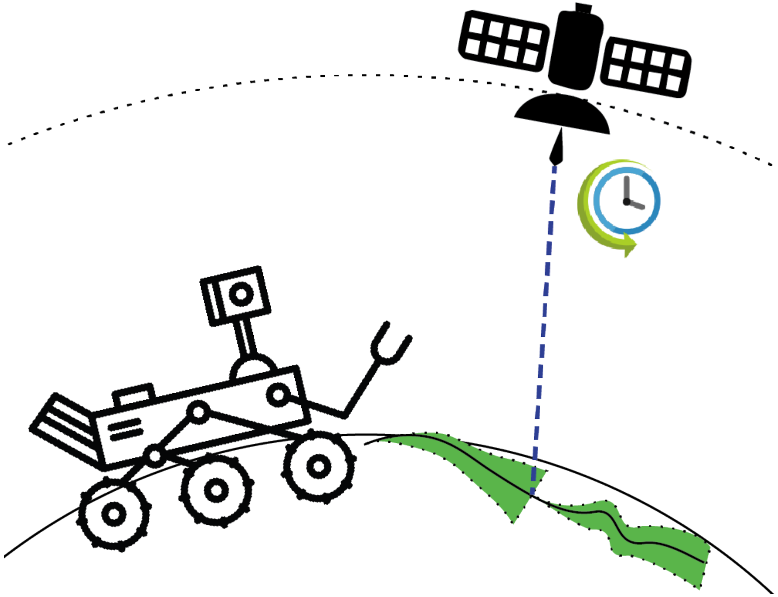
Consider a rover navigating autonomously across a planetary surface. Its goal is to efficiently arrive within a predesignated region of terrain.

Even if the rover departs from a known initial position, its knowledge of its own pose becomes unreliable before long. Without sufficiently informative sensors, uncertainty can accumulate until it impairs the rover's ability to pick traverses which are safe, efficient, or both.

But suppose that along with deployment of the rover, or perhaps as part of the surface deployment, a separate orbital device was also introduced. This satellite carries surface-directed sensors that include a detector capable of localizing the rover. As the satellite circles the body, it acquires data from its extrinsic perspective, obtaining imagery and information to provide the rover's position. When both devices can communicate, the rover has the possibility of a *check-in* with the orbiting device.

How might this affect the way the rover navigates? When should we continue driving under uncertainty, and when should we wait for more information from the orbiter?

In this project, we design efficient *autonomy algorithms* to answer this question.



Motivation

Infrequent, periodic state observations occur frequently in spaceflight application:

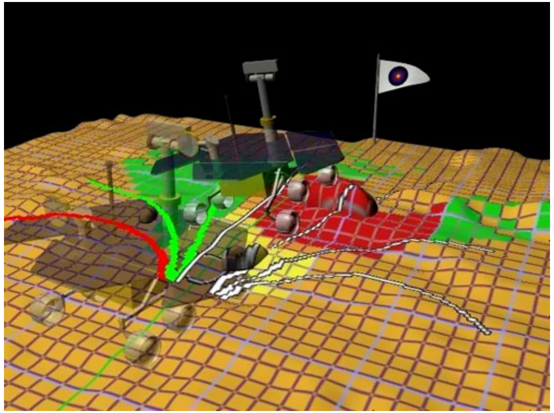
- Human-in-the-loop high-level planning with DSN delay
- Orbiter assist for localization, reconnaissance

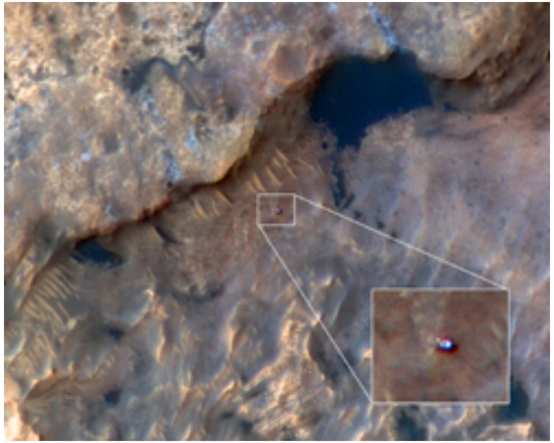
Future Solar System exploration **requires more autonomy**:

- Short mission duration
- Long communication delays
- Complex, ambitious science

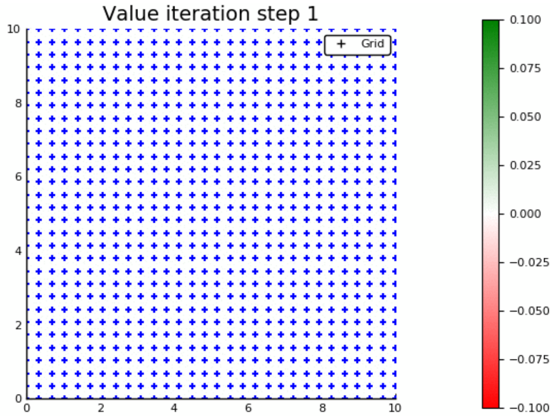
Examples: Europa Lander, PRIME, MSR

We need **computational tools** for autonomous planning problems under uncertainty with infrequent, periodic state observations





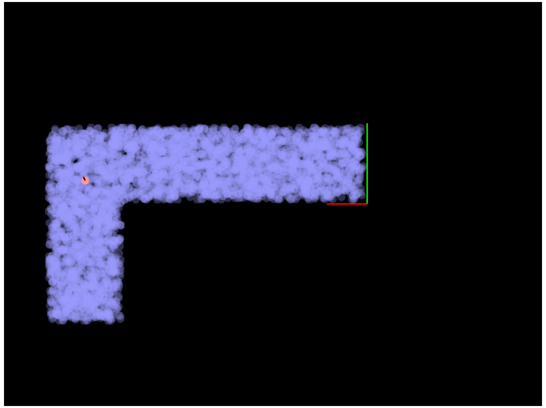
State of the Art



Markov Decision Processes (**MDPs**):

+ Computationally efficient

− Assume full state knowledge before every action



Partially Observable Markov Decision Processes (**POMDPs**)

+ Explicitly represent state uncertainty

− Extremely computationally expensive

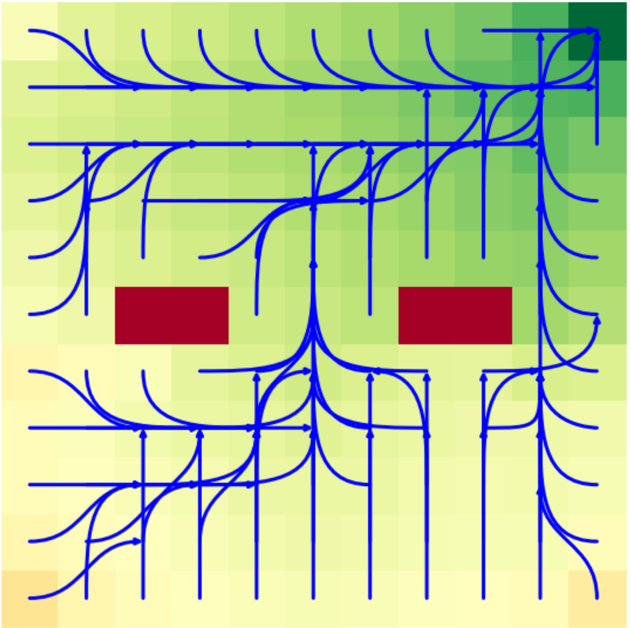
No POMDP formulation captures the "periodic observations" setting

Images courtesy of
https://github.com/JuliaPOMDP/POMDPGallery.jl

Approach

Periodically State Observed MDP (PSO-MDP)

$S = \{s_0, s_1, \ldots, s_{\lvert S \rvert}\}$	Finite set of states
$A = \{a_0, a_1, \ldots, a_{\lvert A \rvert}\}$	Finite set of actions
$T : S \times A \times S \rightarrow [0, 1]$	Transition dynamics
$R : S \times A \rightarrow \mathbb{R}$	Reward function
$\kappa \in \mathbb{N}_{>0}$	Check-in period
$\gamma \in [0, 1)$	Discounting factor



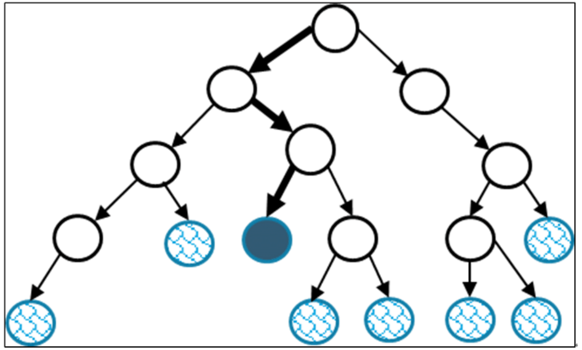
Dual solution approaches:

- MDP with large action space
- POMDP with periodic uncertainty reset

PSO-MDP as a large action-space MDP

Approach:

- ~~Naively solve with MDP tools~~
  - Complexity is **exponential** in time between check-ins
- Develop actionable **upper and lower bounds** on solution
- Develop **algorithms** to **tighten** and **exploit** the upper and lower bounds



Top image: Steve Buissinne via Pixabay
Bottom image: from Shokry, S. et al. (2018). Bandwidth Maximization Approach for Displaced Left-Turn Crossovers Coordination under Heterogeneous Traffic Conditions. Journal of Traffic and Transportation Engineering. 6. 183-196.

Actionable upper and lower bounds

Lower bound

Take one action per check-in

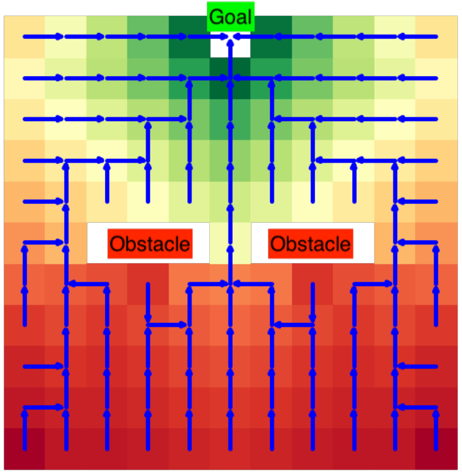
Upper bound

Information check-in at every time step

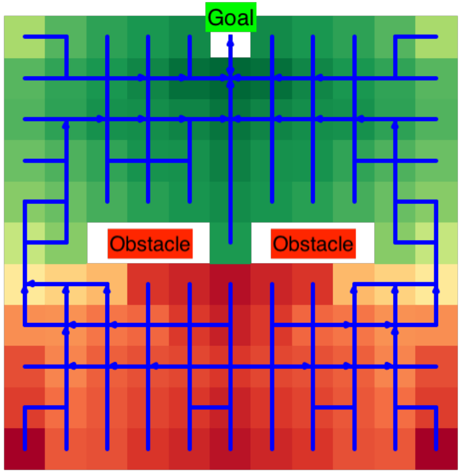
Gap between UB and LB is ~1/κ *

Theorem: Lower bound is a *1/κ approximation* for PSO-MDP

Improvement: “blind rollout” policy



Lower bound



Optimal solution

* See paper for details

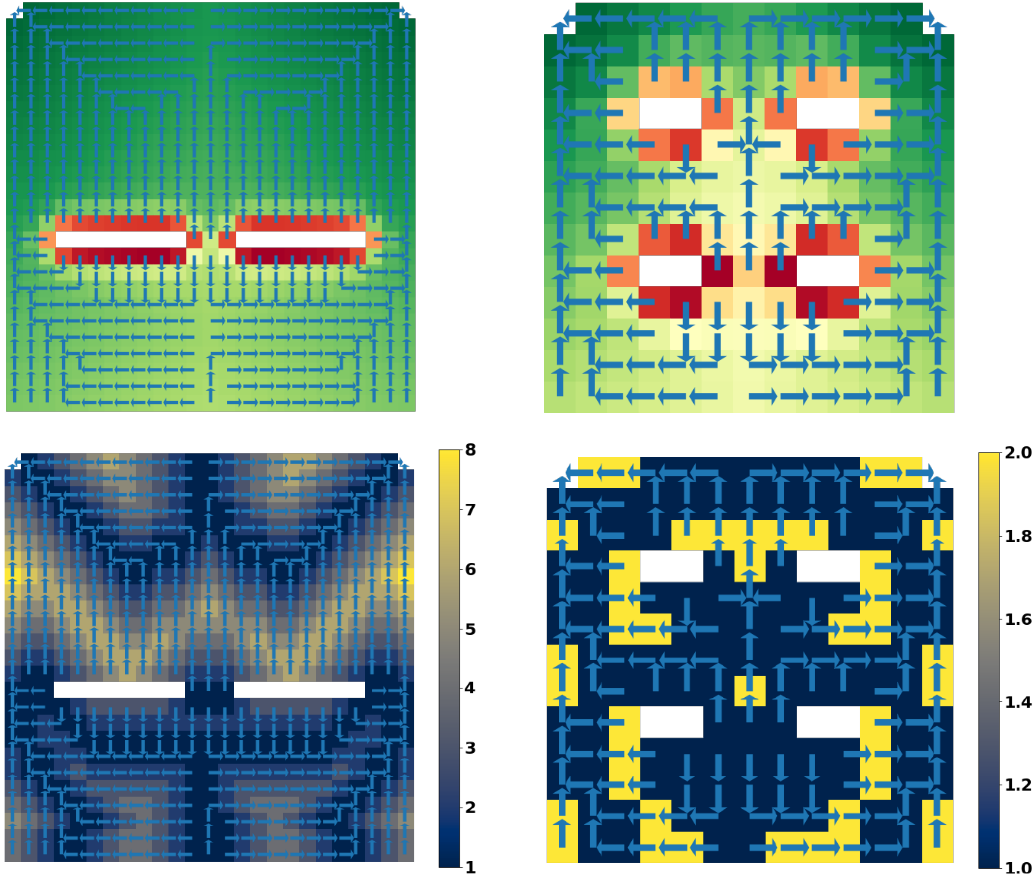
Blind rollout policy

Idea: "chain" multiple actions when waiting for a check-in does not add additional information

Extension (planned): chain multiple actions when picking the wrong action has well-bounded downside

+ *Highly effective* in uncluttered environments

− *Not so good* in cluttered environments



Uncluttered

Cluttered

A Branch-and-Bound Approach

Action family: all composite actions with a given prefix

Idea:

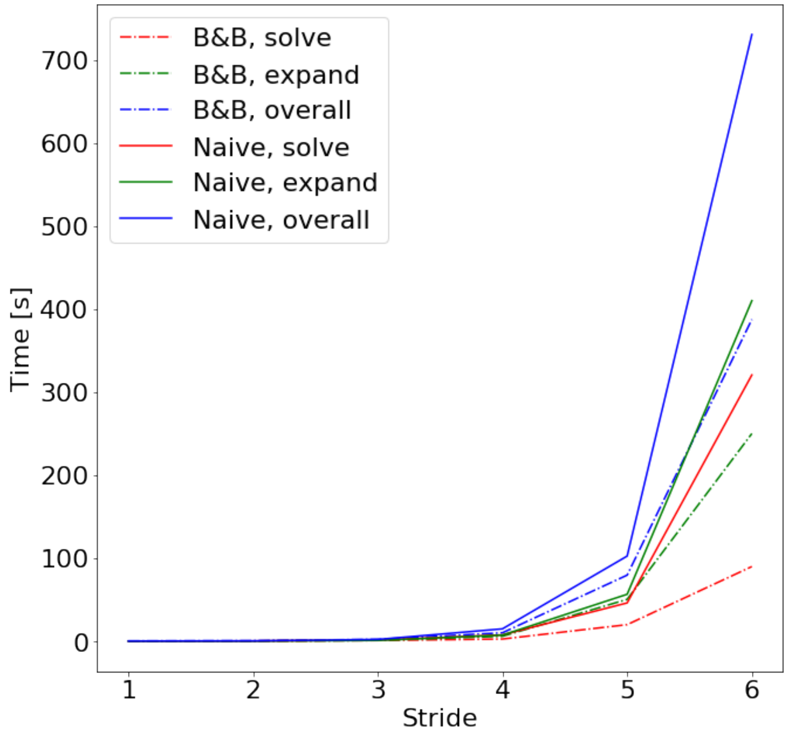
1. Find upper and lower **bounds** for action families of length *s*
2. **Prune** suboptimal actions families
3. **Increase** *s*

Upper Bound

*Add extra check-ins (idea: more information is better)**

Lower Bound

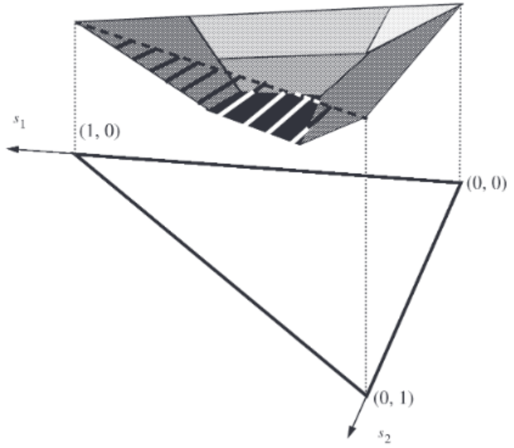
Add an *arbitrary suffix* to the action family prefix

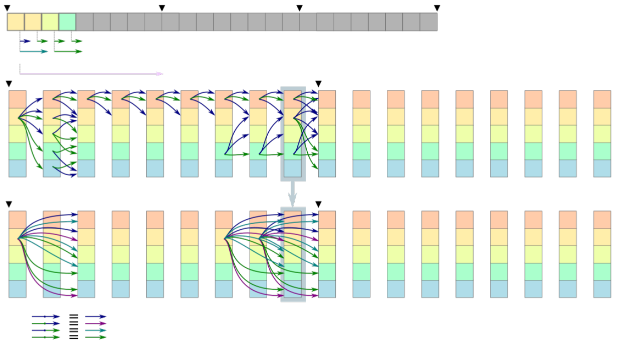


* But you need check-ins at the right times! See paper for details

Looking Forward

- **Heuristics** and classifiers to select a subset of actions
    - Suboptimal, but *bounded* complexity
- Adapt online **POMDP** solvers to the PSO-MDP setting
    - Information check-ins *reset* the agent's belief
    - Run multiple online solvers from different states, *rewire* to states
- **Refine** upper bounds for PSO-MDP
    - Location of the extra information check-ins is a *decision variable*





Publications and References

F. Rossi* and D. Shell*, “*Autonomous planning under uncertainty with infrequent, periodic state observations*”, in preparation for conference submission.