FY23 Strategic University Research Partnership (SURP)

Fast planning under uncertainty with explicit operational and safety guarantees

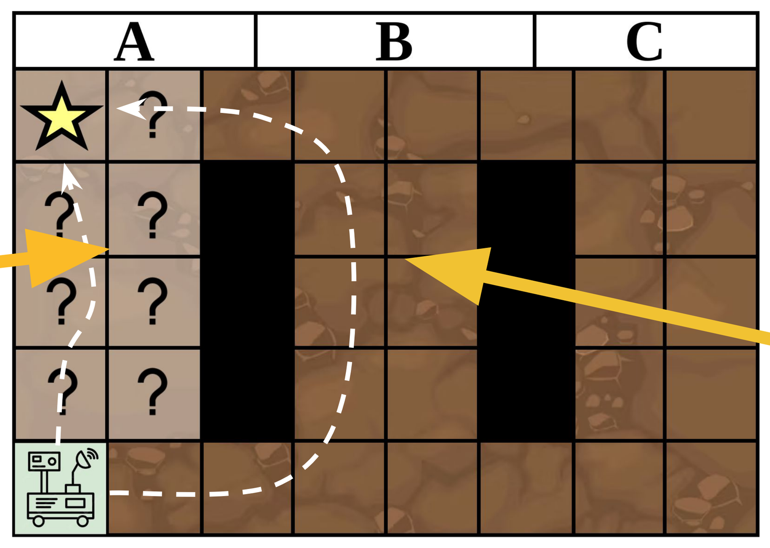
Principal Investigator: Federico Rossi (347)

Co-Investigators: Martin Feather (512), Zachary Sunberg (University of Colorado at Boulder), Morteza Lahijanian (University of Colorado at Boulder)

Goal and Objective

Develop efficient algorithms for *planning under uncertainty* with explicit operational and *safety guarantees*

Shorter but potentially sandy terrain



Longer but known traversable terrain

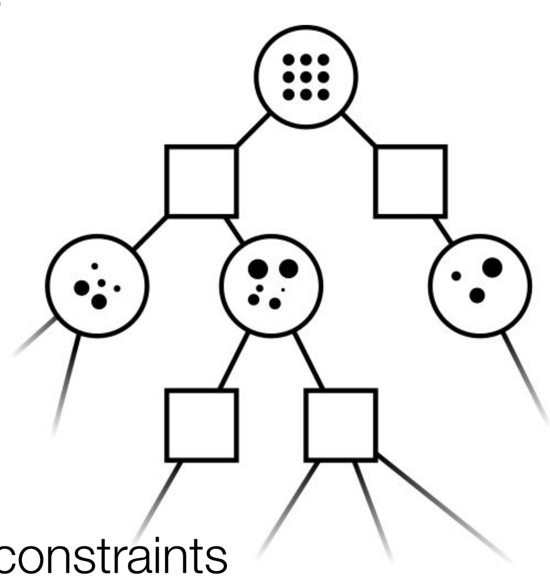
Examples of problem optimization objectives:

Min(drive time) *subject to p*(drive completed)*>0.99*

Max(data volume) *subject to p*(battery undervoltage)*<0.0001*

Max(good samples collected) *subject to p*(hardware failure)*<0.00001*

Background



Planning under uncertainty: POMDPs

- Efficient approximate algorithms
- No safety constraints

Safety-Aware POMDPs

- Assign extra penalty to "bad" end states
- Needs handcrafted penalties
- Asymptotic guarantees

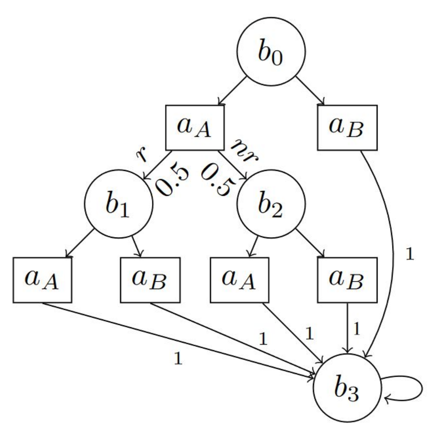
Constrained POMDPs

- Provide rigorous way of describing complex constraints
- Asymptotic guarantees
- Algorithms are not scalable with many assumptions
- Pathological behavior of algorithms

Our contribution: Recursively-Constrained POMDPs

- Provide rigorous way of describing complex constraints
- Finite sample guarantees
- Scalable algorithms
- Good behaving algorithms

Approach: Novel Problem Definition - Recursively Constrained POMDPs (RC-POMDPs)



$R(b,a)$	a_A	a_B
b_0	0	10
b_1	12	0
b_2	12	0

$C(b,a)$	a_A	a_B
b_0	0	5
b_1	8	5
b_2	2	5

Previous constrained POMDPs create **pathological behavior** due to **mismatch** between **operational requirements** and **constraint formulation**!

Recursively-Constrained POMDPs

- Obeys Bellman's principle of opimality - creates **better behavior**
- Able to do **re-planning**
- Deterministic (**more intuitively verified**) policies

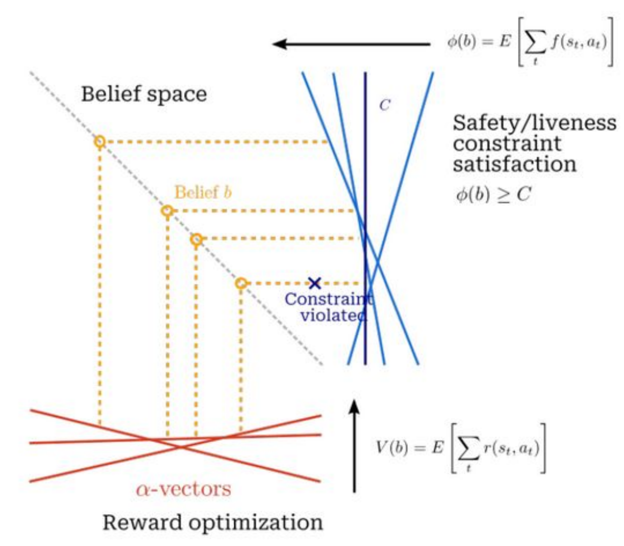
Problem 2 (RC-POMDP Planning Problem). *Given a C-POMDP and an admissibility horizon $k \in \mathbb{N}_0 \cup \{\infty\}$, compute optimal policy π^* that is k-admissible, i.e.,*

$$\pi^* = \arg\max_{\pi} V_R^{\pi}(b_0) \quad (11)$$

$$\text{s.t.} \quad W(h_t) + \gamma^t V_C^{\pi}(b_t) \le \hat{c} \quad \forall t \in \{0, \ldots, k\}. \quad (12)$$

Approach: Scalable Point-based Algorithm

- Developed a **point based value iteration algorithm** for Recursively-Constrained POMDPs
- Provide **explicit bounds** on the expected reward
- Provide **explicit guarantees** for constraint satisfactions
- Leverages recent advancements for POMDPs for **scalable planning**

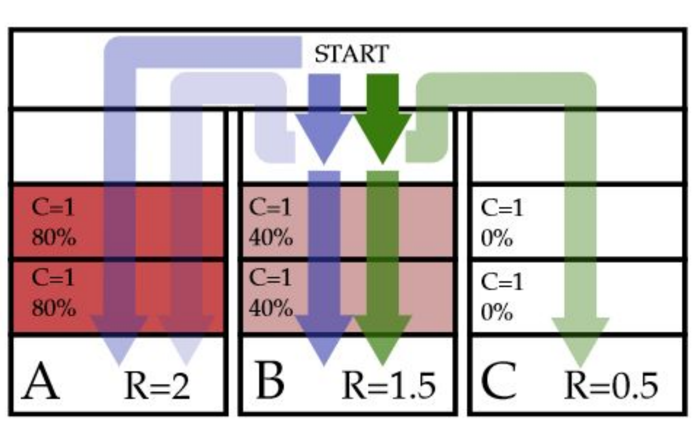


Future Work (Y2-Y3)

- Extend to **linear temporal logic** constraints
- Apply to **human-in-the-loop planning** under uncertainty
- Robot demonstration in Mars Yard

Selected Results

- **Scalable** up to state spaces of 10^4 states
- Results show that RC-POMDPs **align** more to operational specs



Example Planned Trajectories
Blue (prev C-POMDP formulation)
Green (ours)

Env.	Algorithm	Violation Rate	Reward	Cost
CE ($\hat{c}=5$)	CGCP	0.51	**12.00**	5.19
	CGCP-CL	**0.00**	6.12	3.25
	CPBVI	**0.00**	8.39	4.38
	CPBVI-D	**0.00**	6.10	3.54
	Ours	**0.00**	**10.00**	5.00
C-Tiger ($\hat{c}=3$)	CGCP	0.75	**−1.69**	3.00
	CGCP-CL	0.14	**−2.98**	2.93
	CPBVI	0.15	−11.11	2.58
	CPBVI-D	0.09	−9.49	2.76
	Ours (CL)	**0.00**	**−5.75**	2.98
CRS(4,4) ($\hat{c}=1$)	CGCP	0.51	**10.43**	0.51
	CGCP-CL	0.78	1.68	0.72
	CPBVI	**0.00**	−0.40	0.52
	CPBVI-D	**0.00**	0.64	0.47
	Ours	**0.00**	**6.96**	0.50
CRS(5,7) ($\hat{c}=1$)	CGCP	0.41	**11.98**	1.00
	CL-CGCP	0.18	9.64	0.99
	CPBVI	**0.00**	0.00	0.00
	CPBVI-D	**0.00**	0.00	0.00
	Ours	**0.00**	**11.62**	0.95
Tunnels ($\hat{c}=1$)	CGCP	0.50	1.61	1.01
	CL-CGCP	0.31	1.22	0.68
	CPBVI	0.90	**1.92**	1.62
	CPBVI-D	0.89	**1.92**	1.57
	Ours	**0.00**	**1.02**	0.44

Benefits to JPL and NASA

Uncertainty is ubiquitous

- Environmental conditions (e.g., terrain-wheel interaction, air density)
- Actuator performance (e.g., thruster valve timing, heater performance)

Planning under uncertainty can increase science returns

From conservative, *worst-case* margins to explicit, *tight* representation of uncertainty

Planning with formal guarantees helps performance as well as V&V

- Formal guarantees span the *wide range of conditions* in which the autonomy is intended to operate
- Testing explores in detail *specific scenarios*
- Formal guarantees *complement* testing to help instill confidence that the autonomy will perform correctly and sufficiently rapidly

National Aeronautics and Space Administration

Jet Propulsion Laboratory
California Institute of Technology
Pasadena, California

www.nasa.gov

Clearance Number: CL#00-0000
Poster Number: RPC#

Publications:

Q. H. Ho, et al. Recursively-Constrained Partially Observable Markov Decision Processes (submitted)

PI/Task Mgr. Contact Information:
Federico Rossi, federico.rossi@jpl.nasa.gov