UNIVERSITY of VIRGINIA | ENGINEERING

LINK LAB

Predictive Monitoring and Safety Shielding for AI-Enabled Cyber-Physical Systems

Lu Feng

University of Virginia

lu.feng@virginia.edu

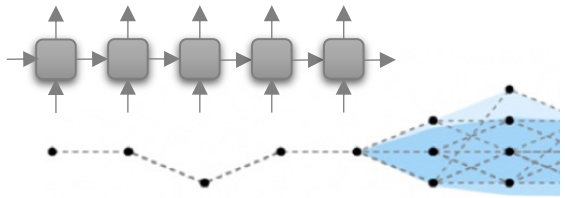
85th IFIP WG 10.4 Meeting, Feb 2024

Motivation

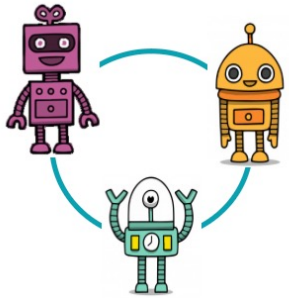
- Growing use of AI/ML technologies in safety-critical CPS
- Design-time verification is not sufficient for safety guarantees
- Need runtime techniques
  - Predictive monitoring
  - Safety shielding
  - …

Images generated by AI (OpenAI's DALL·E)

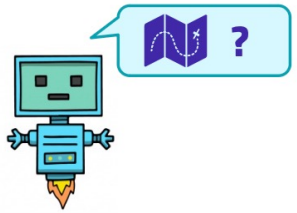
Outline



Predictive monitoring for Bayesian RNNs



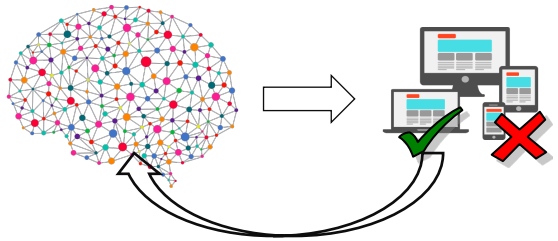
Safety shielding for multi-agent RL



Safe POMDP online planning via shielding and conformal prediction

Predictive monitoring for Bayesian RNNs

- Predictive monitoring enhances CPS decision-making support
  - Adapt traffic signals in response to predicted congestion from accidents
  - Lower insulin dosage automatically on predicting hypoglycemia risk
- Existing work mostly focus on monitoring individual predictions
- Our work monitors sequential predictions generated from Bayesian RNNs that can capture the inherent uncertainty in CPS

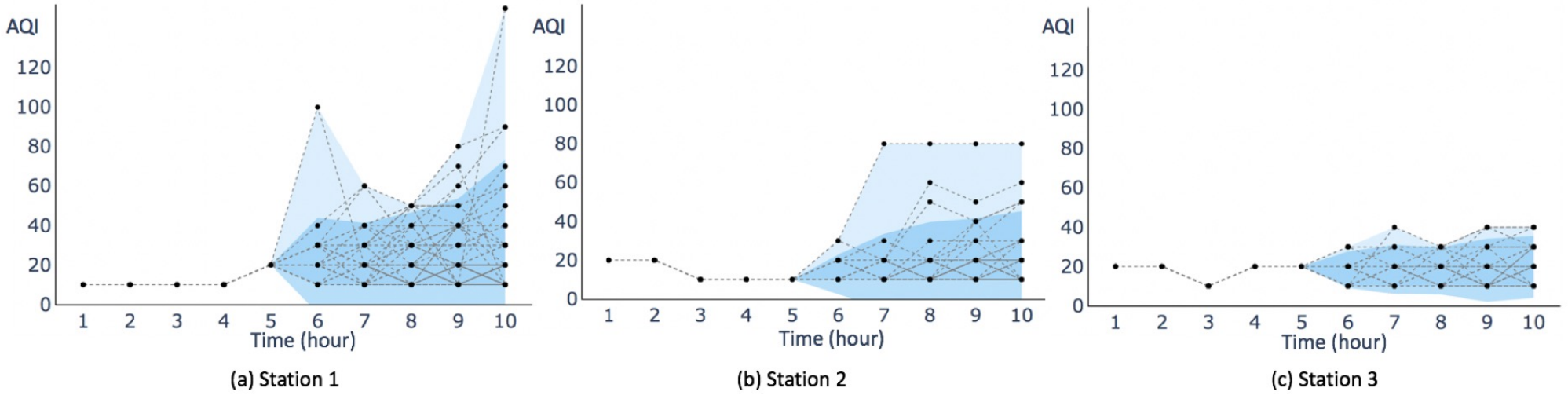


"Predictive monitoring with logic-calibrated uncertainty for cyber-physical systems". M Ma, J Stankovic, E Bartocci, L Feng. EMSOFT 2021.

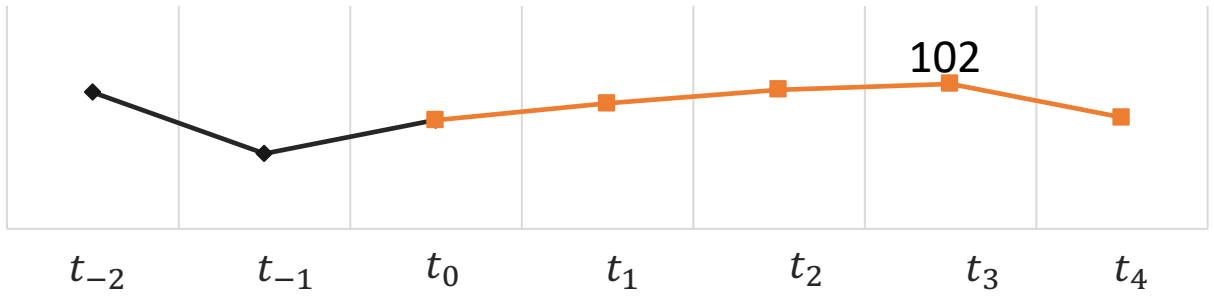
Insights from real-world CPS datasets

- Uncertainty in CPS data
  - Sensing noise
  - Environment
  - Human behavior

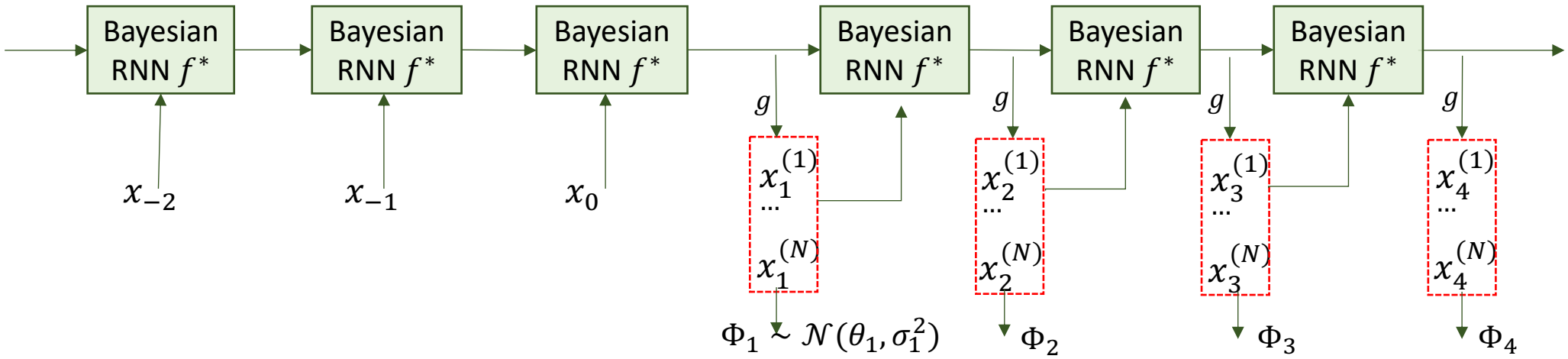
Dataset	Location	Period	# Records
Air quality	437 stations	5/2014-4/2015	2,891,393
Traffic volume	1,490 streets	9/2014-4/2018	514,776

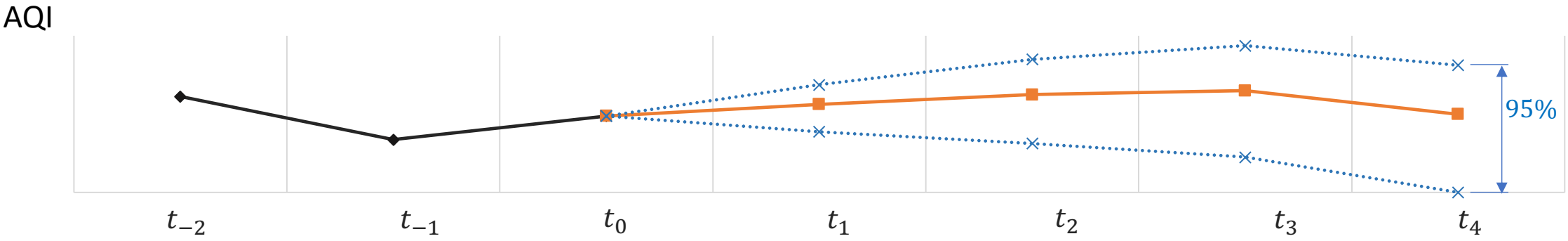


Example scenario

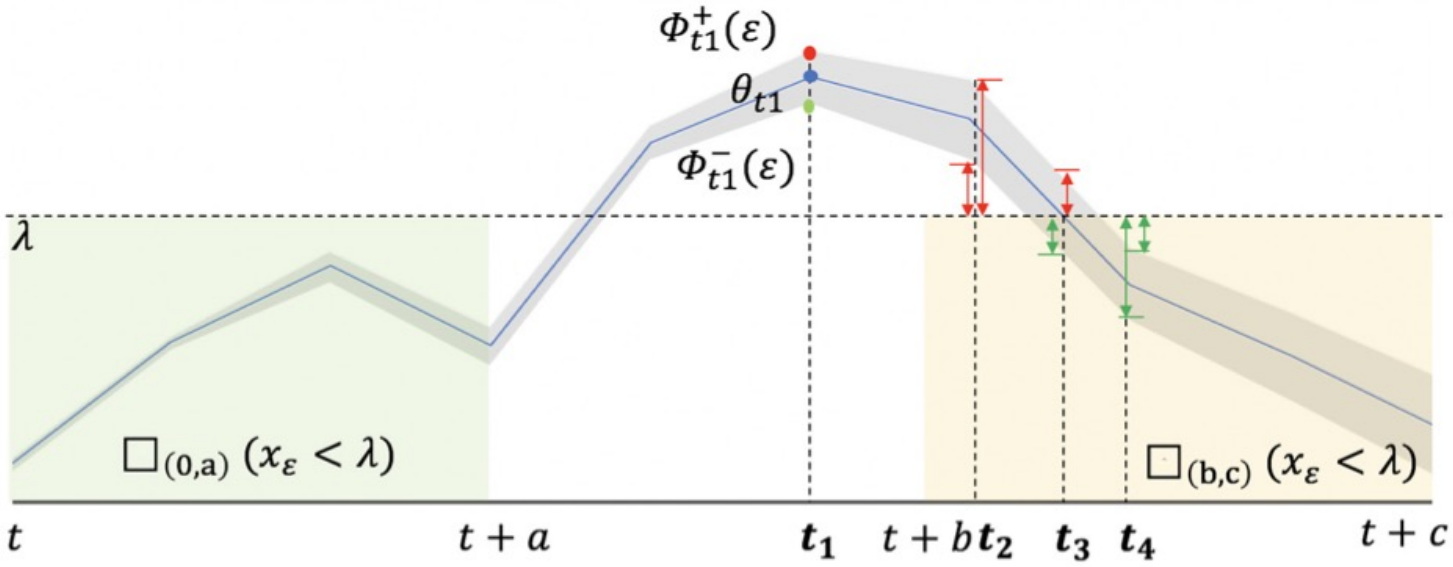


Uncertainty in deep learning

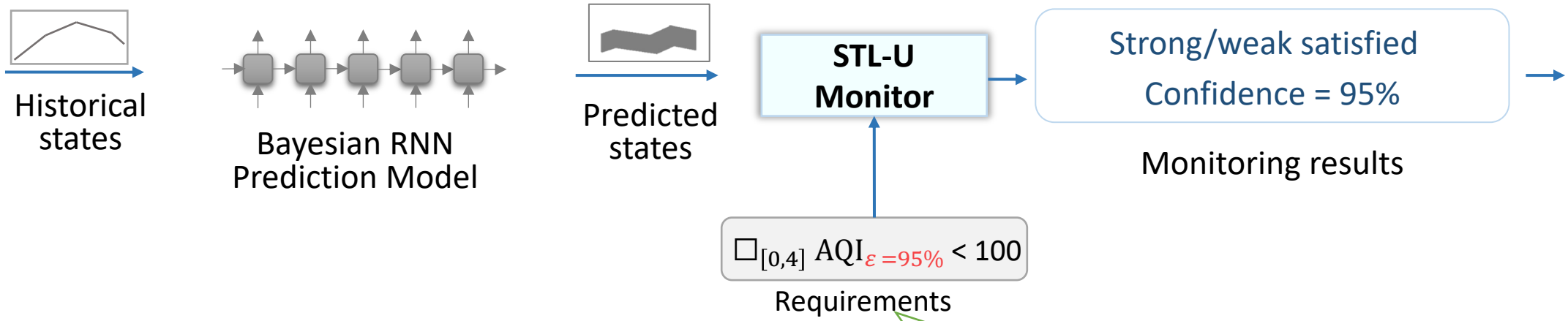




STL-U: Signal Temporal Logic with Uncertainty



Predictive monitoring with uncertainty

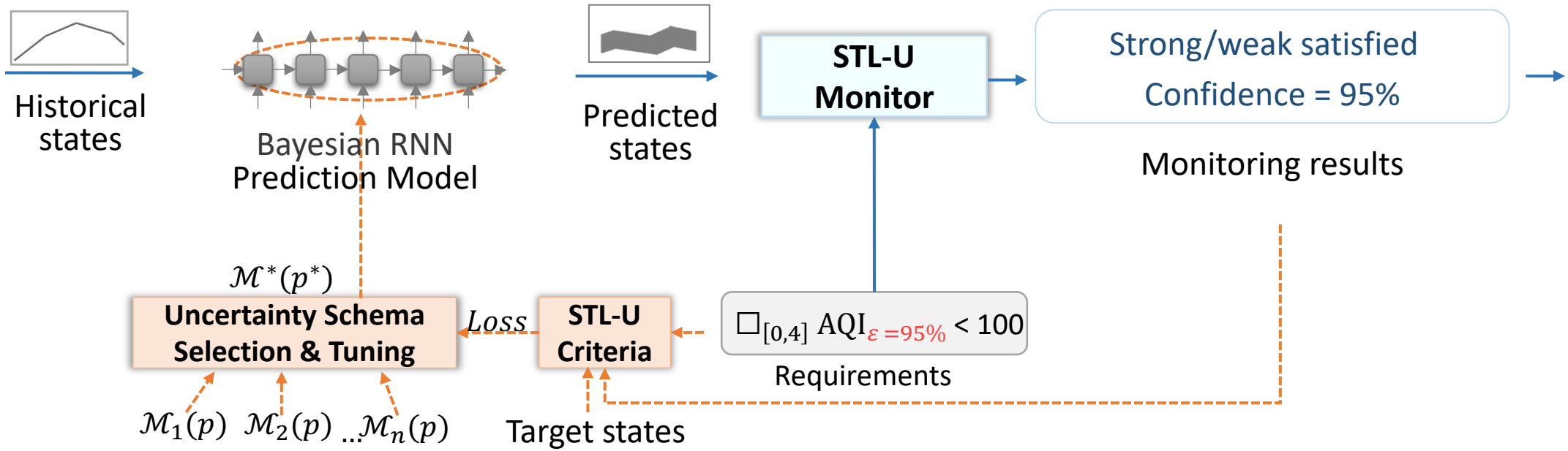


With 95% confidence level, the predicated air quality index in the next 4 hours should always be below 100

- A Novel Specification Logic: STL-U
  - Strong/Weak Semantics
  - Confidence Calculation

What is the confidence level that guarantees the predicated air quality index in the next 4 hours always be below 100

Logic-calibrated uncertainty estimation



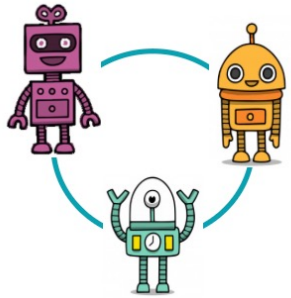
Evaluation

	No Monitor	STL Monitor	STL-U Monitor	
Number of Violation	undetected	267	189	
Air Quality Index	67.91	57.22	43.65	23.7%
Noise (db)	73.32	49.27	48.21	
Emergency Waiting Time (s)	20.32	14.87	10.65	28.3%
Vehicle Waiting Number	22	18	15	
Pedestrian Waiting Time (s)	190.2	148.9	121.1	
Vehicle Waiting Time (s)	112.12	89.77	80.31	

☞ City safety & performance ⬆

Outline

Predictive monitoring for Bayesian RNNs

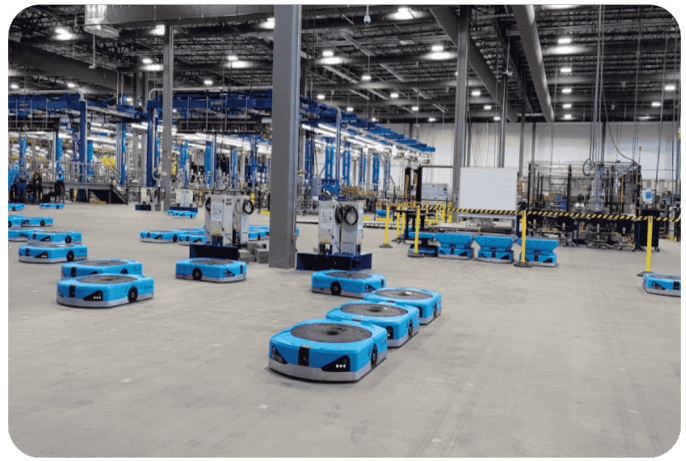


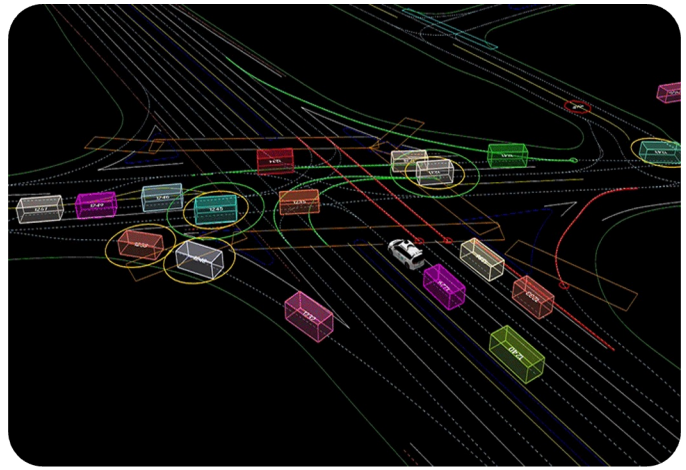
Safety shielding for multi-agent RL

Safe POMDP online planning via shielding and conformal prediction

Safety shielding for multi-agent RL (MARL)

- MARL has been used in many CPS applications
- Traditional MARL methods focus on optimizing returns and do not prevent unsafe actions
- Our methods provide safety guarantees during learning and execution

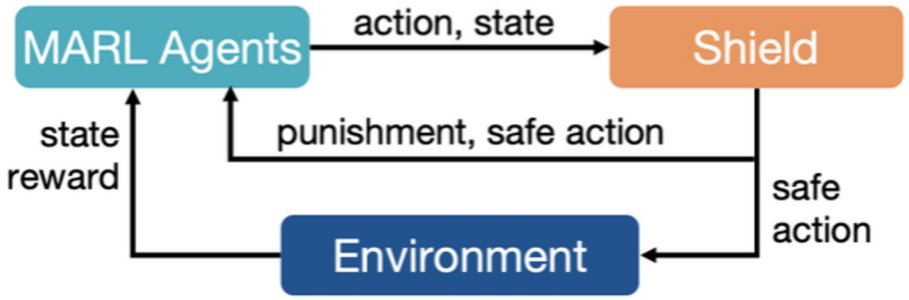




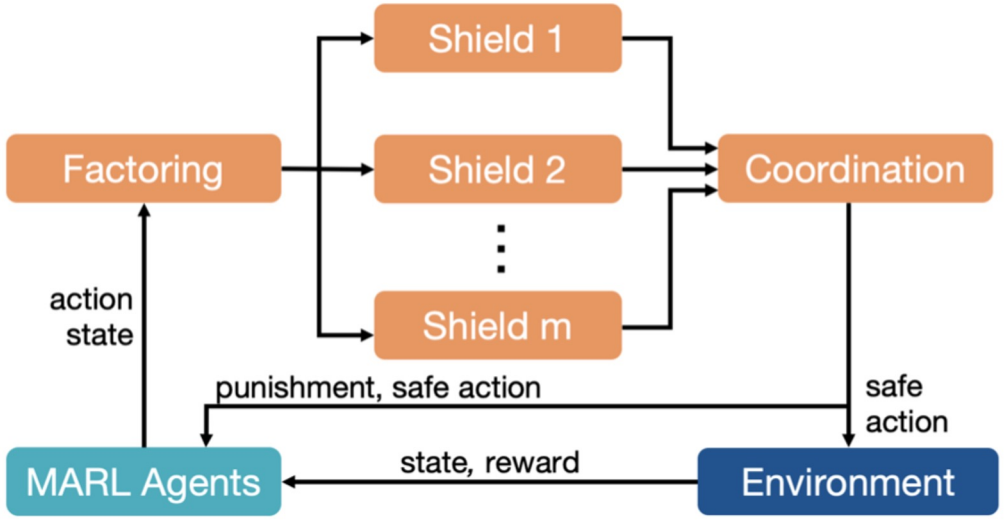
"Safe multi-agent reinforcement learning via shielding".
I ElSayed-Aly, S Bharadwaj, C Amato, R Ehlers, U Topcu, L Feng. AAMAS 2021.

Safety shielding for multi-agent RL (MARL)

Centralized Shielding



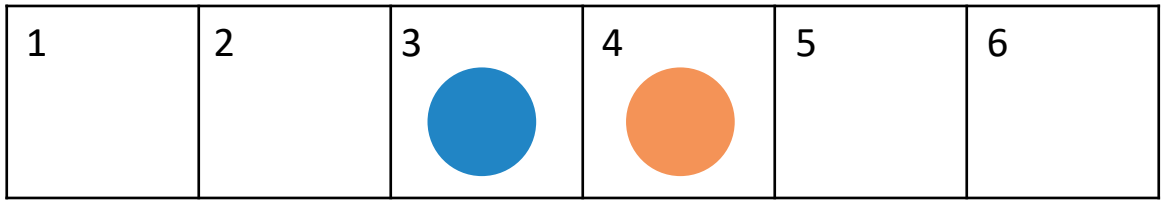
Factored Shielding

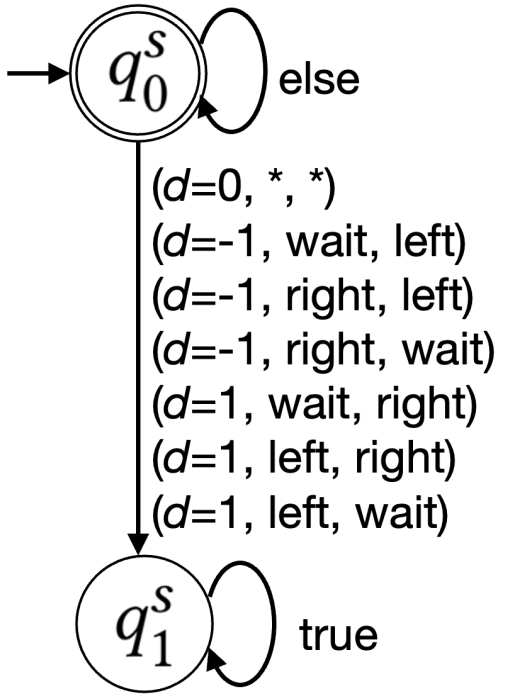


"Safe multi-agent reinforcement learning via shielding".
I ElSayed-Aly, S Bharadwaj, C Amato, R Ehlers, U Topcu, L Feng. AAMAS 2021

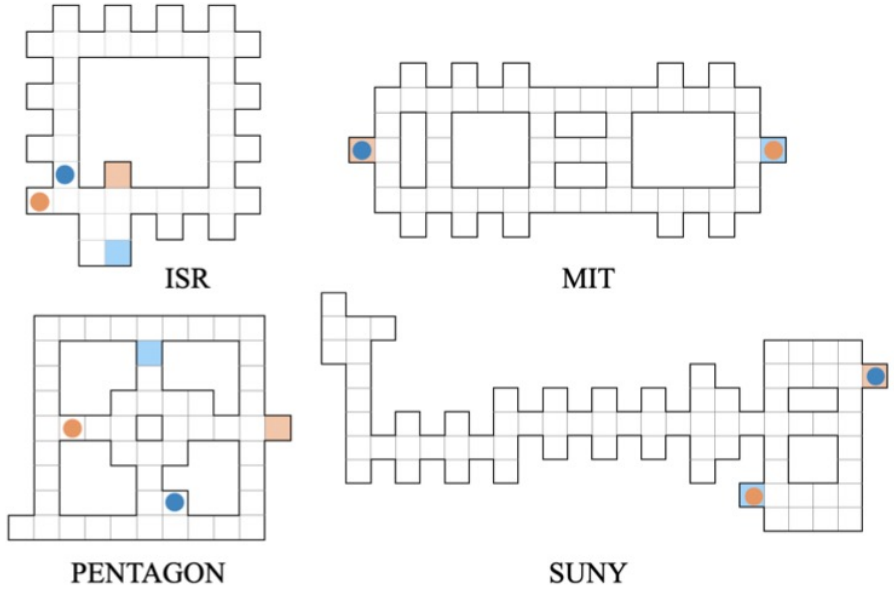
Safety shielding for multi-agent RL (MARL)

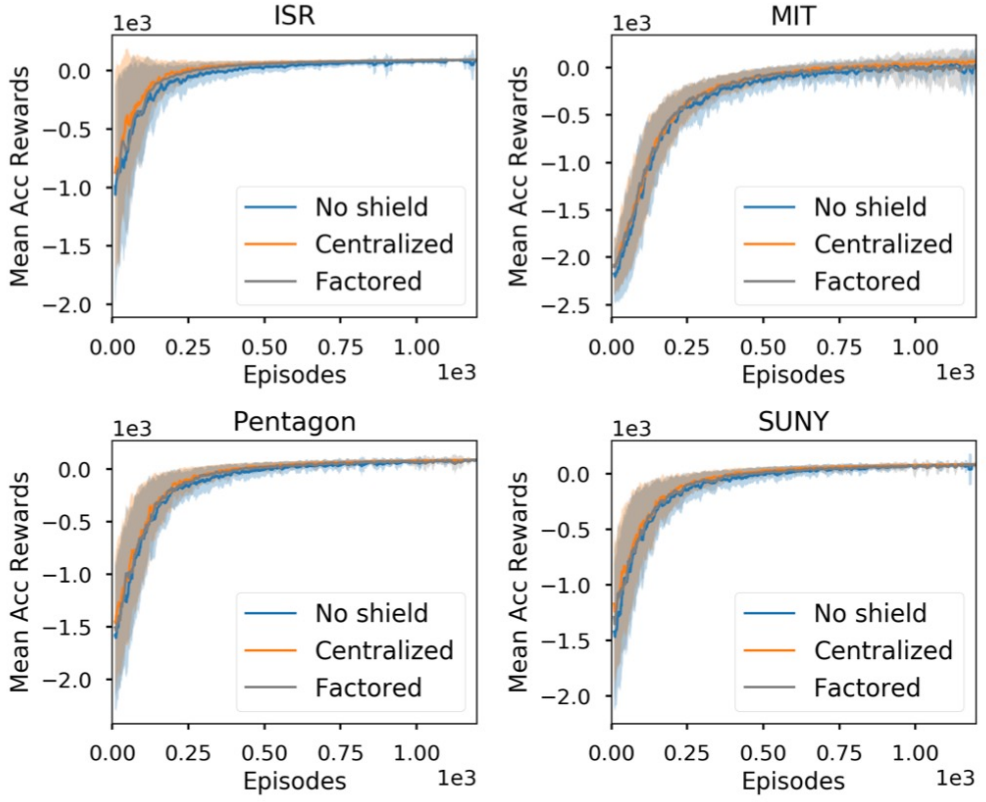
- Safety specification in Linear Temporal Logic
- Synthesizing shields by solving two-player safety games





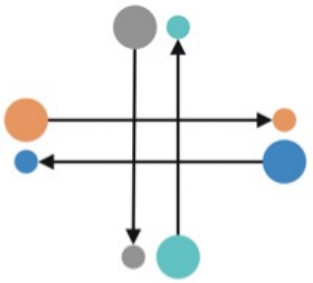
Evaluation on discrete environments



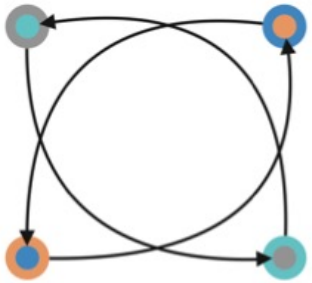


		IQL			CQ			CQ with centralized shield			CQ with factored shield		
Maps	Optimal Steps	Steps	Reward	Collisions	Steps	Reward	Collisions	Steps	Reward	Collisions	Steps	Reward	Collisions
ISR	5	30.35	-10.20	20.30	8.66	89.53	0.40	7.03	93.85	**0.00**	7.31	93.74	**0.00**
Pentagon	10	46.58	-19.17	11.60	10.96	88.96	0.20	12.08	88.44	**0.00**	13.20	84.88	**0.00**
MIT	18	20.84	77.33	0.00	42.93	30.38	0.90	28.38	73.94	**0.00**	29.96	37.96	**0.00**
SUNY	10	34.80	-160.175	72.60	13.97	84.78	0.30	11.97	88.44	**0.00**	14.02	83.77	**0.00**

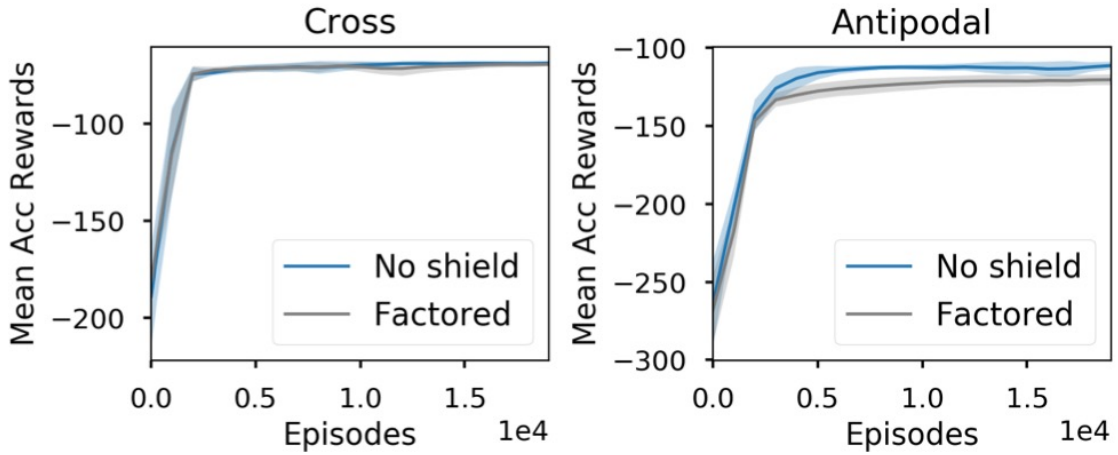
Evaluation on continuous environments



Cross



Antipodal

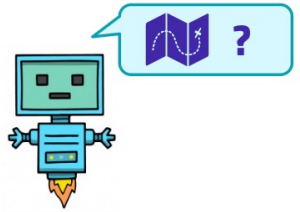


	MADDPG	MADDPG with Shield
Cross	207.20	**0.00**
Antipodal	14,419.20	**0.00**

Outline

Predictive monitoring for Bayesian RNNs

Safety shielding for multi-agent RL

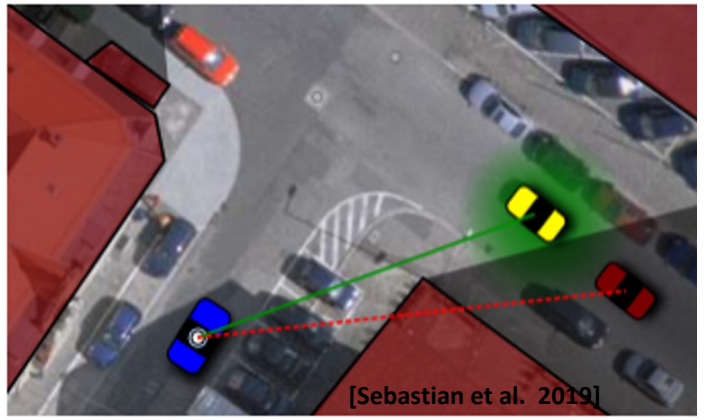


Safe POMDP online planning via shielding and conformal prediction

Safe POMDP online planning via shielding

- POMDP provides a general modeling framework for decision-making under uncertainty

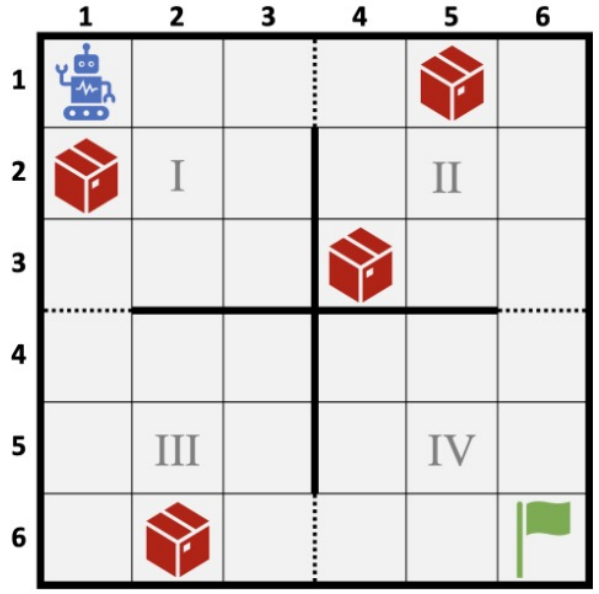
- POMDP online planning
  - Policy computation and execution are interleaved
  - Can scale up to solve very large POMDPs than offline planning



"Safe POMDP online planning via shielding". S Sheng, D Parker, L Feng. ICRA 2024.

Safe POMDP online planning via shielding

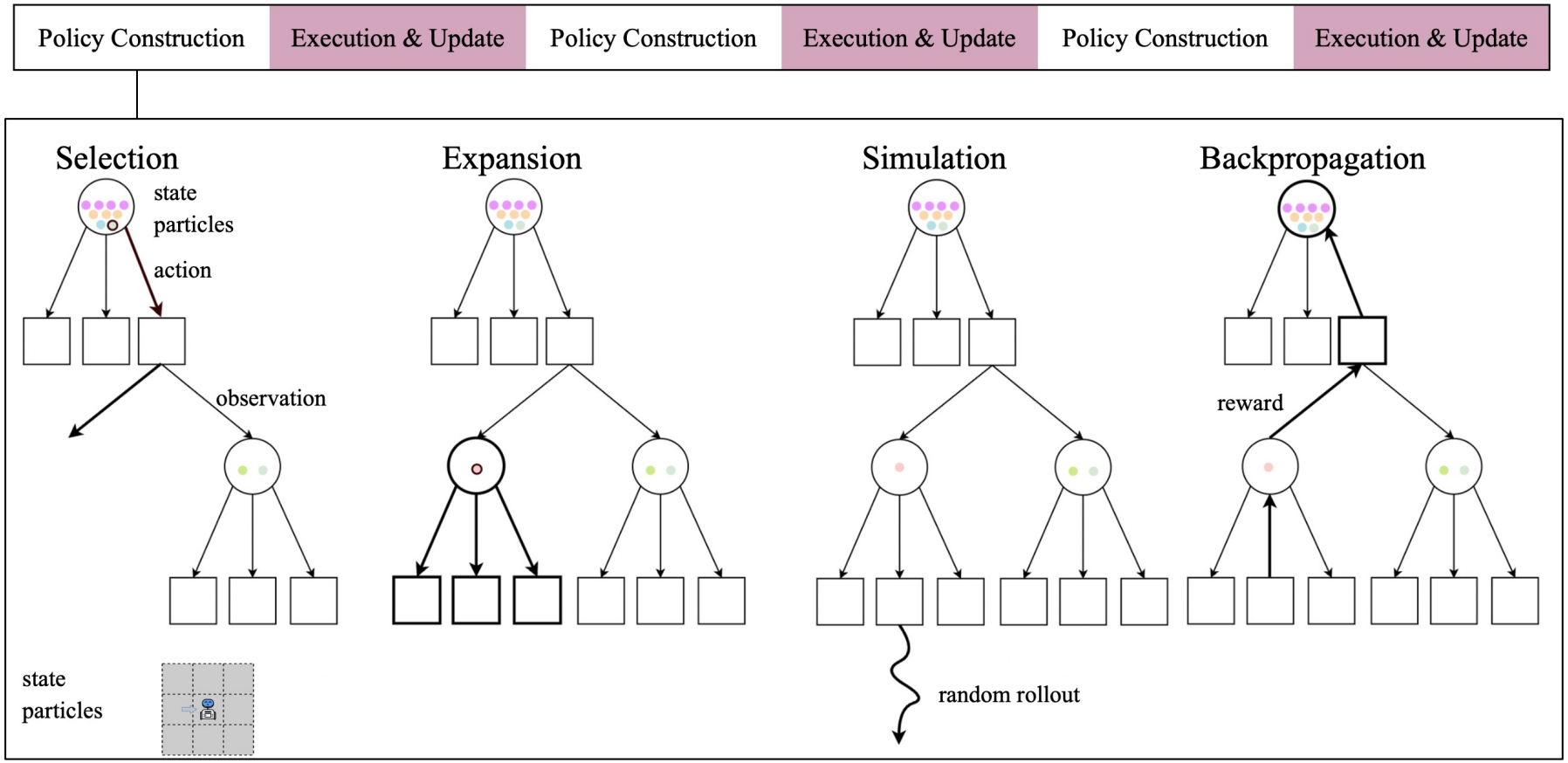
- Existing methods consider various safety requirements
  - Cost-constrained
  - Chance-constrained
- Our work focuses on stricter safety requirements
  - Almost-sure reach-avoid specifications (i.e., the probability of reaching goal states while avoiding unsafe states is 1)
  - Shield synthesis via computing maximal winning regions with a SAT-based method (Junges et al. 2021)
  - Centralized shield vs. factored shields



"Safe POMDP online planning via shielding". S Sheng, D Parker, L Feng. ICRA 2024.

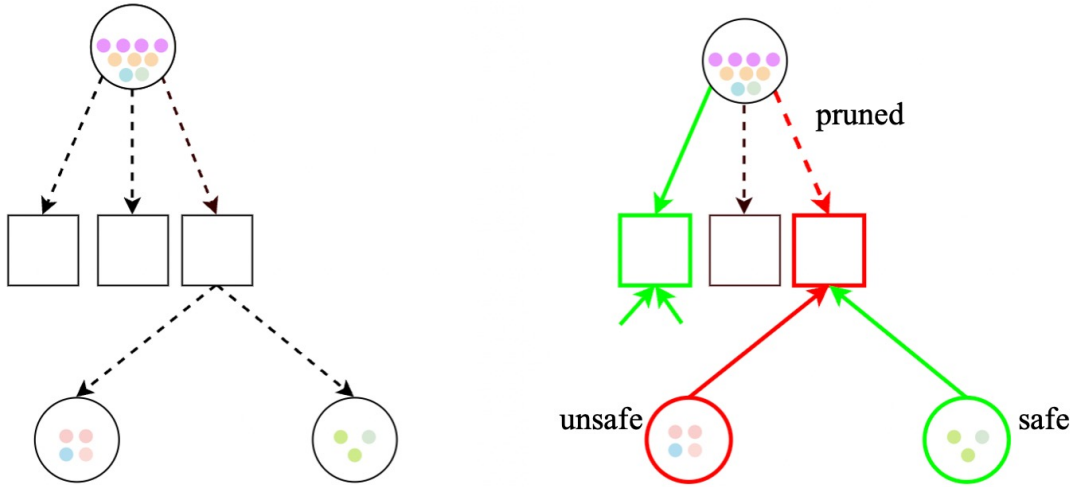
Partially observable monte-carlo planning

- A widely used POMDP online planning algorithm (Silver et al. 2008)



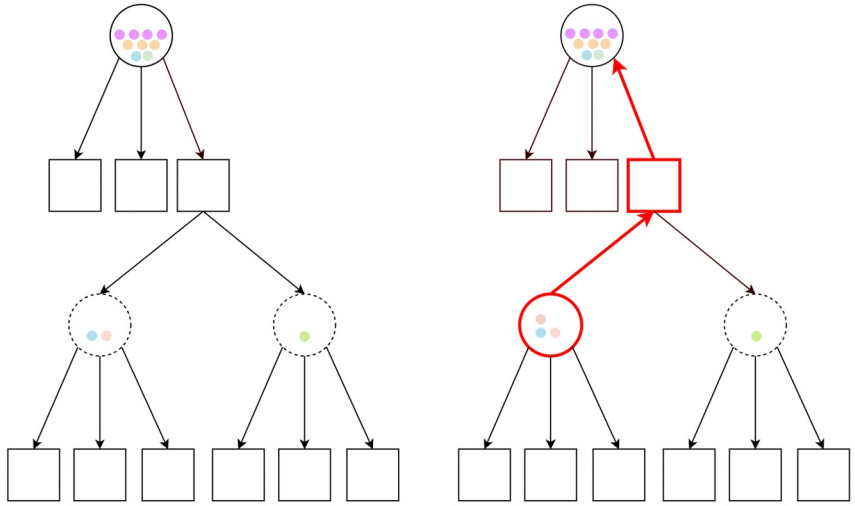
Prior pruning

- At each time step t, before the POMCP algorithm iterations, find all actions disallowed by the shield and prune the corresponding tree branches from the root node



On-the-fly backtracking

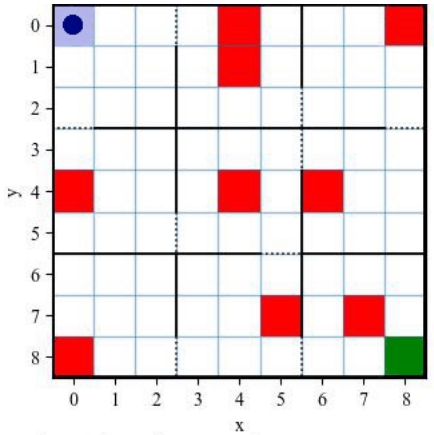
- During the POMCP simulation phase, check if every updated particle set along the path is contained in the shield's winning region. If not, prune the tree branch.



Legend:

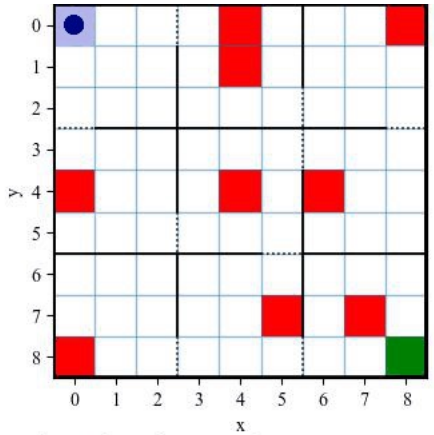
- Ground-truth state
- Belief State
- Target state
- Obstacle

Centralized Shielding With Prior Pruning



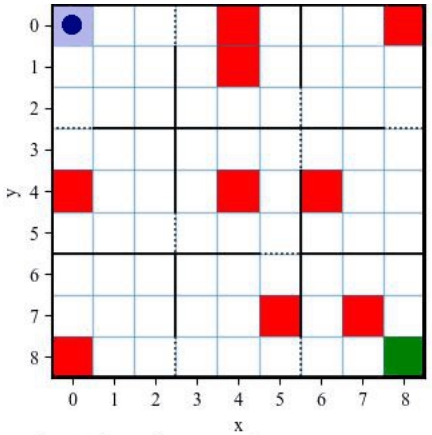
Selected action: south
Disallowed actions: []
Cumulative reward: -1.0 Step: 001

Centralized Shielding With On-the-fly Pruning



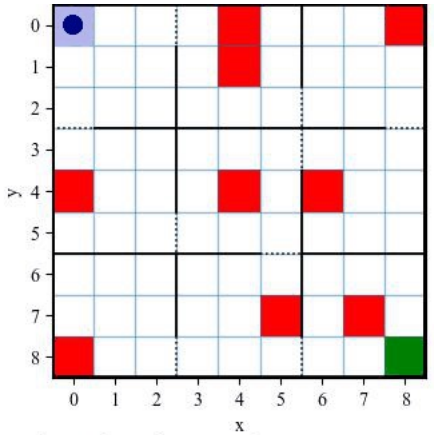
Selected action: south
Disallowed actions: []
Cumulative reward: -1.0 Step: 001

No Shielding



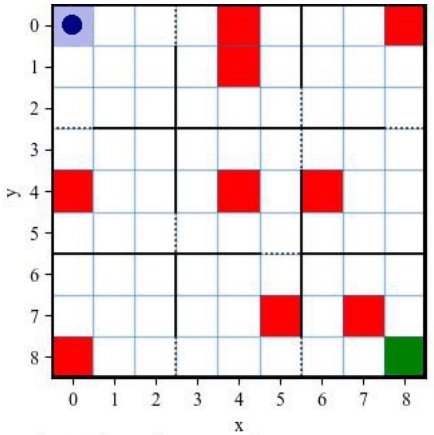
Selected action: south
Disallowed actions: []
Cumulative reward: -1.0 Step: 001

Factored Shielding With Prior Pruning



Selected action: south
Disallowed actions: []
Cumulative reward: -1.0 Step: 001

Factored Shielding With On-the-fly Pruning

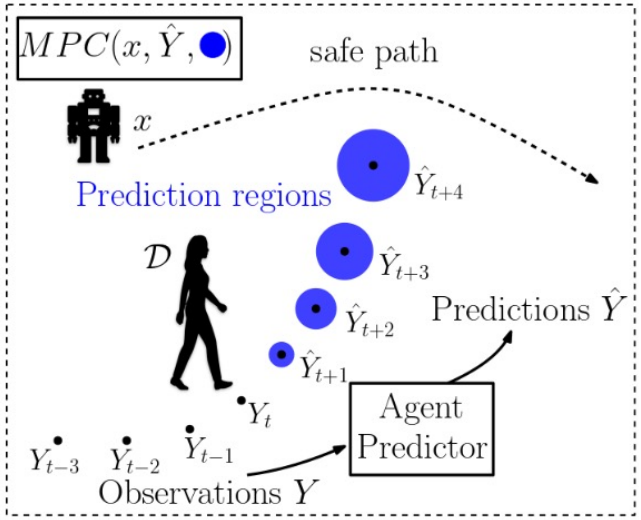


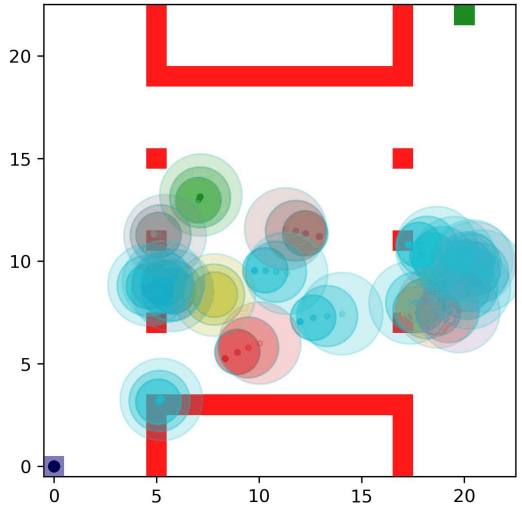
Selected action: south
Disallowed actions: []
Cumulative reward: -1.0 Step: 001

Accounting for pedestrians (ongoing work)

- Predict pedestrians' future trajectories using trained LSTM models and quantify prediction uncertainty with adaptive conformal prediction
- Online computation of winning regions
- Shielding POMCP based on winning regions

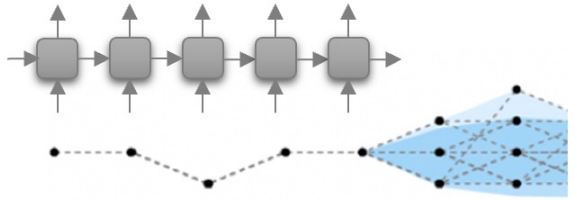
(Dixit et al. 2023)



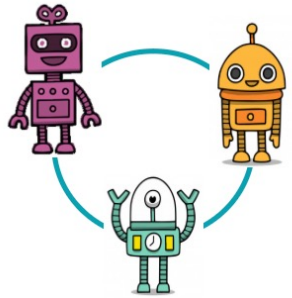


Conclusion

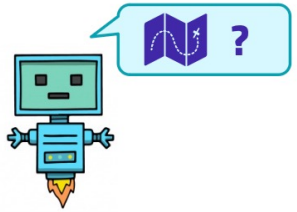
- Safe AI-enabled CPS necessitate runtime techniques like predictive monitoring and safety shielding



- Various AI methods require different safety guarantees
  - Bayesian RNNs
  - Multi-agent RL
  - POMDPs



- Many interesting open research questions …



Thank you!
Questions and Comments?

Lu Feng
lu.feng@virginia.edu

Sponsors: