Making and Keeping Informed Commitments in Human-Machine Systems

(FA9550-15-1-0039)

PI: Edmund H. Durfee (University of Michigan)

Co-PI: Satinder Singh (University of Michigan)

AFOSR Program Review:
Computational Cognition and Machine Intelligence Program
(October 6-8, 2020, Arlington, VA)

Making and Keeping Informed Commitments in Human-Machine Systems

Edmund H. Durfee and Satinder Singh, University of Michigan

Objective:

- Improve ability of machine to know when to ask human for clarification, and what to ask
- Exploit autonomous capabilities while reliably meeting commitments to teammates

Approach:

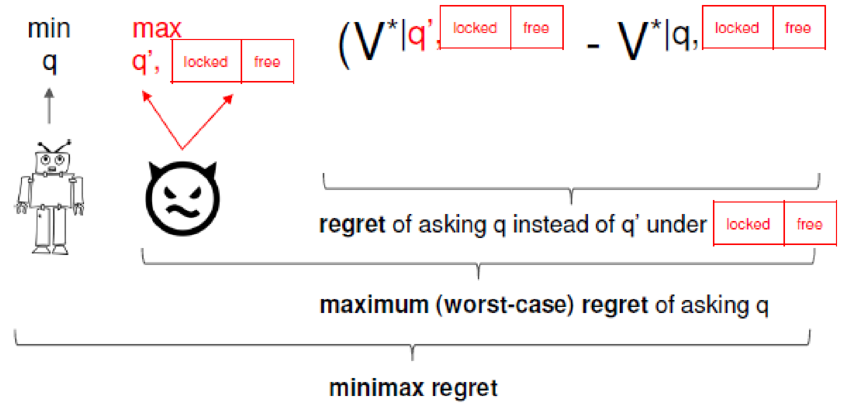
- Integrate sequential decision, expected value-of-information, and cover-set methods for querying
- Make probabilistic commitments and revise their pursuit in a constrained responsive manner

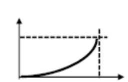
DoD Benefits:

- Avoid (bad) surprises to human operator
- Autonomous systems proactively seek clarification
- Flexible but reliable autonomous systems

Progress:

- Suite of methods for asking most informative queries to increase safety and reward
- Principled semantics for pursuing and interpreting commitments, and for querying to rapidly find optimal cooperative commitments



$\hat{P}_u(\cdot)$	Suboptimality / $(v^+ - v^-)$	
	Achievement	Maintenance
Min Enablement (Step Function)	≤ 1 Better than $\hat{P}_u^-$	> 1 Can be worse than $\hat{P}_u^-$
Max Enablement (Step Function)	> 1	> 1
Min Value Timing (Step Function)	≤ 1	> 1
Minimax Regret Timing (Step Function)	> 1	> 1
Constant Toggling 	> 1	> 1

List of Project Goals

1. Develop methods for finding (approximately-)optimal queries for resolving reward uncertainty
2. Develop methods for finding (approximately-)optimal queries for resolving safe side-effect uncertainty
3. Integrate querying for both reward and safety uncertainty
4. Develop principled semantics for probabilistic commitments in Bayesian settings, and beyond
5. Develop tractable techniques for finding, interpreting, and adhering to probabilistic commitments
6. Integrate commitments and querying into a unified framework for querying to efficiently converge on an optimal cooperative commitment

Progress Towards Goals (or New Goals)

1. Projection-based techniques for querying to resolve reward uncertainty (ICAPS-17)
2. Querying to resolve safe side-effect uncertainty using minimax-regret (AAMAS-18, IJCAI-18) and set-cover (AAAI-20, NeurIPS-19 Workshop) techniques
3. Integrate querying for both reward and safety uncertainty (S. Zhang PhD dissertation)
4. Develop principled semantics and tractable adherence techniques for probabilistic commitments in Bayesian settings (JAAMAS-20, ICAPS-17, IJCAI-16)
5. Develop interpretation strategies for commitments of achievement and maintenance (AAAI-20)
6. Unify commitments and querying into a framework for querying to efficiently converge on an optimal cooperative commitment (OptLearnMAS@AAMAS-20)

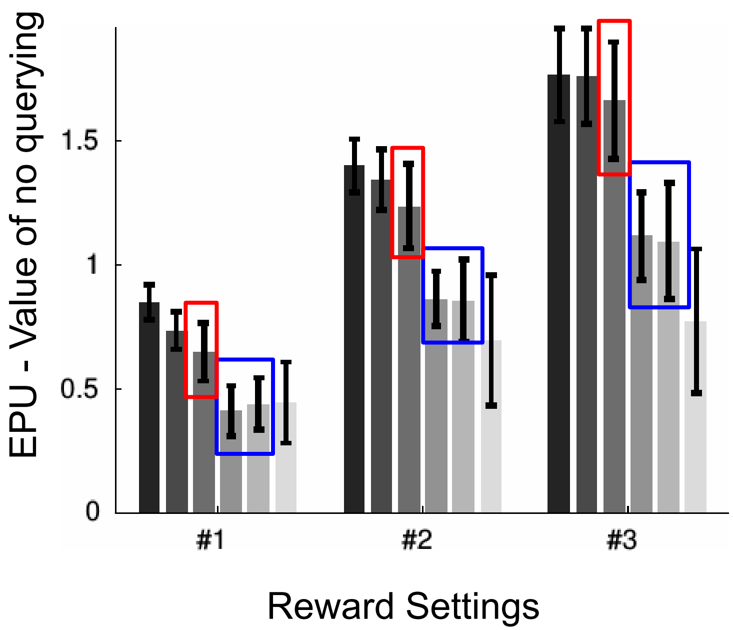
<u>Projection-Based Querying Under Reward Uncertainty</u>

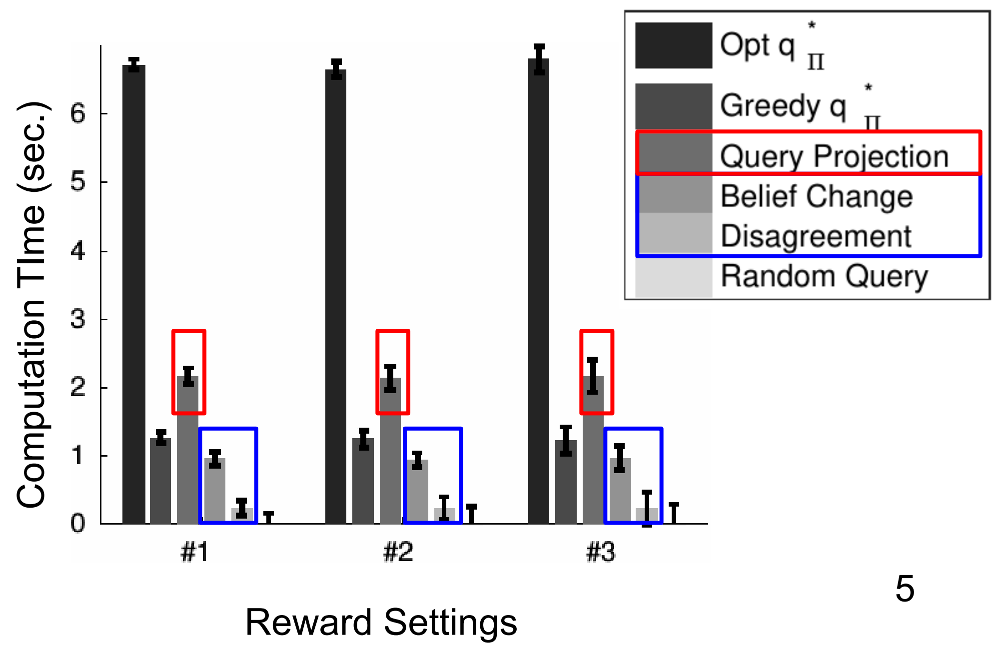
Reward Uncertainty: Which behaviors would most/least satisfy user?

Query optimization: What multiple (*k*-ary) choice query will elicit the most valuable information to resolve the uncertainty.

Key **contributions/results**:

- **<u>Prove</u> that no query can outperform a query offering *k* <u>policies</u> to choose from**
- Submodularity supports greedy approximation, but still too many policies to check
- **Define an appropriately-constrained MILP to <u>construct</u> next policy to greedily add**
- **Develop method to <u>project</u> policy queries to <u>easily answerable</u> trajectory queries**





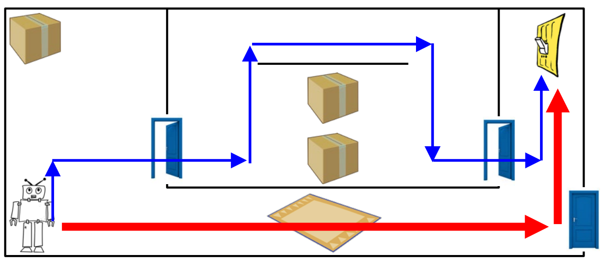
Minimax-Regret Querying for Safely-Optimal Plans

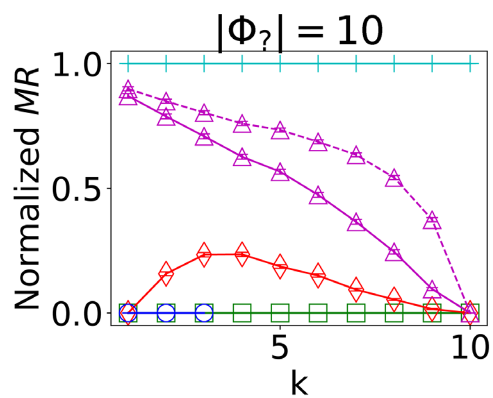
Safety Uncertainty: Which side-effects are acceptable, and which aren't?

Query optimization: Beginning <u>with a safe plan</u>, which *k* possible side-effects to ask if they are free to do, so as to maximally improve the safe plan

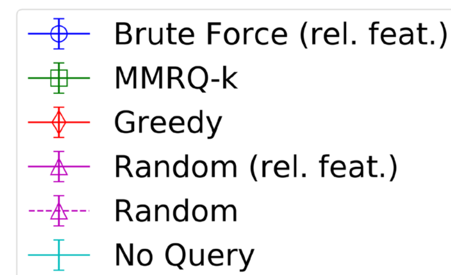
Key **contributions/results**:

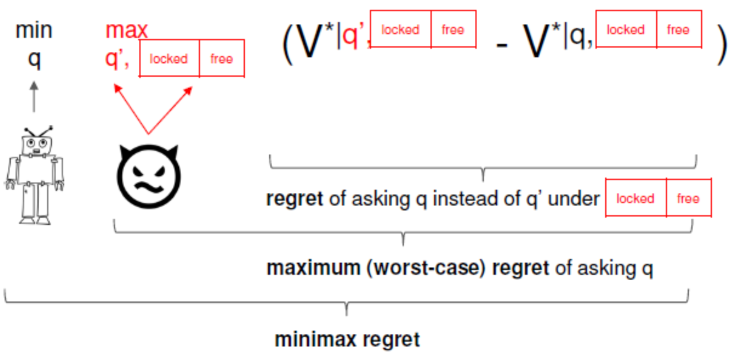
- **Efficiently finding <u>dominating policies</u> to provably identify <u>relevant side-effects</u>**
- **MMRQ-k: Adversarial search method to quickly find minimax-regret *k*-ary query**

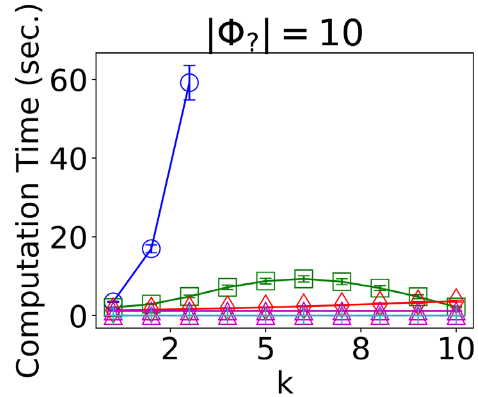




MMRQ-k always finds the minimax-regret query.







MMRQ-k finds the minimax-regret query much faster

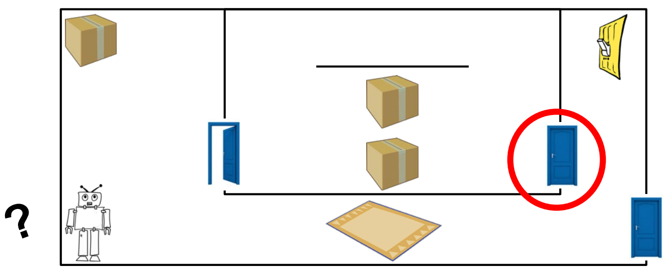
Set-Cover Methods for Safe Side-Effect Planning

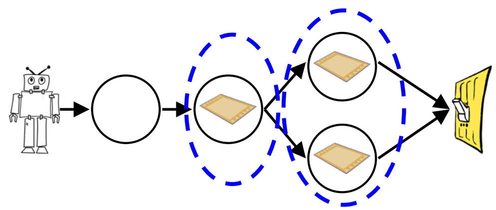
Safety Uncertainty: Which side-effects are acceptable, and which aren't?

Safe Planning: With no initial safe plan, ask about as few side-effects as possible (in expectation) to find one or prove none exists

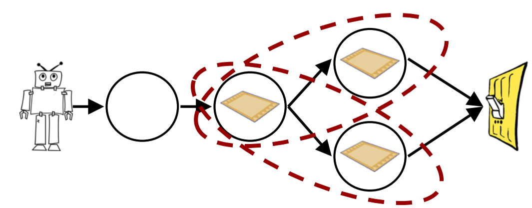
Key **contributions/results**:

- **Formulate the search problem as a pair of parallel set-cover problems**
- **Develop heuristics to find best side-effect to ask to progress on both problems**

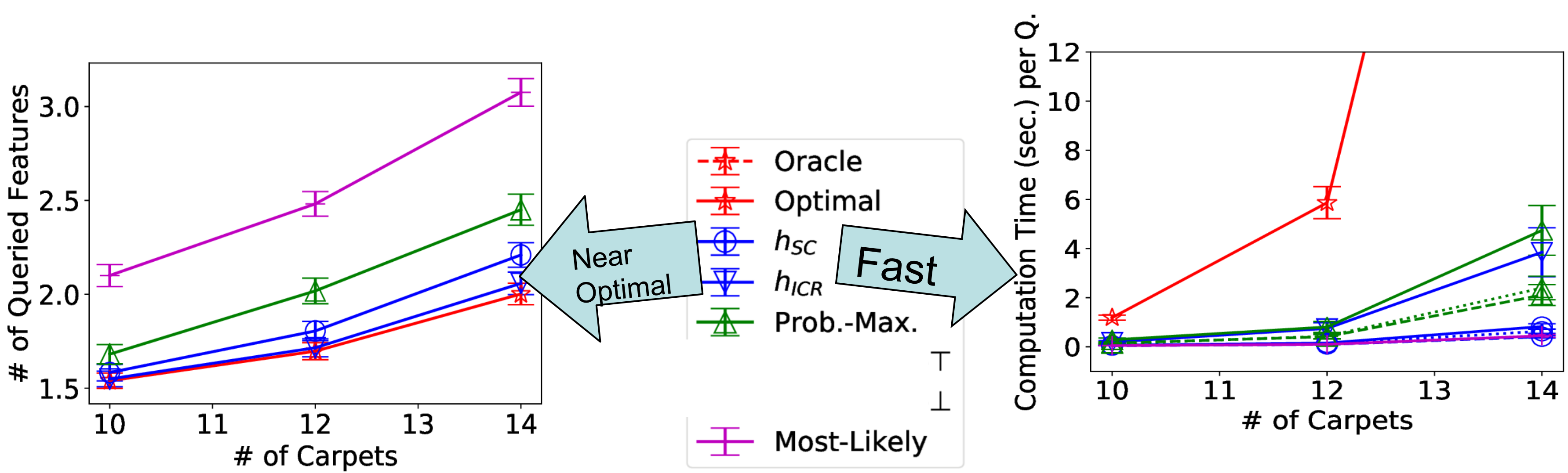




If both sets contain **safe** carpets, then **a safe policy exists**.



If both sets contain **unsafe** carpets, then **no safe policy exists**.



Querying Under Both Safety and Reward Uncertainty

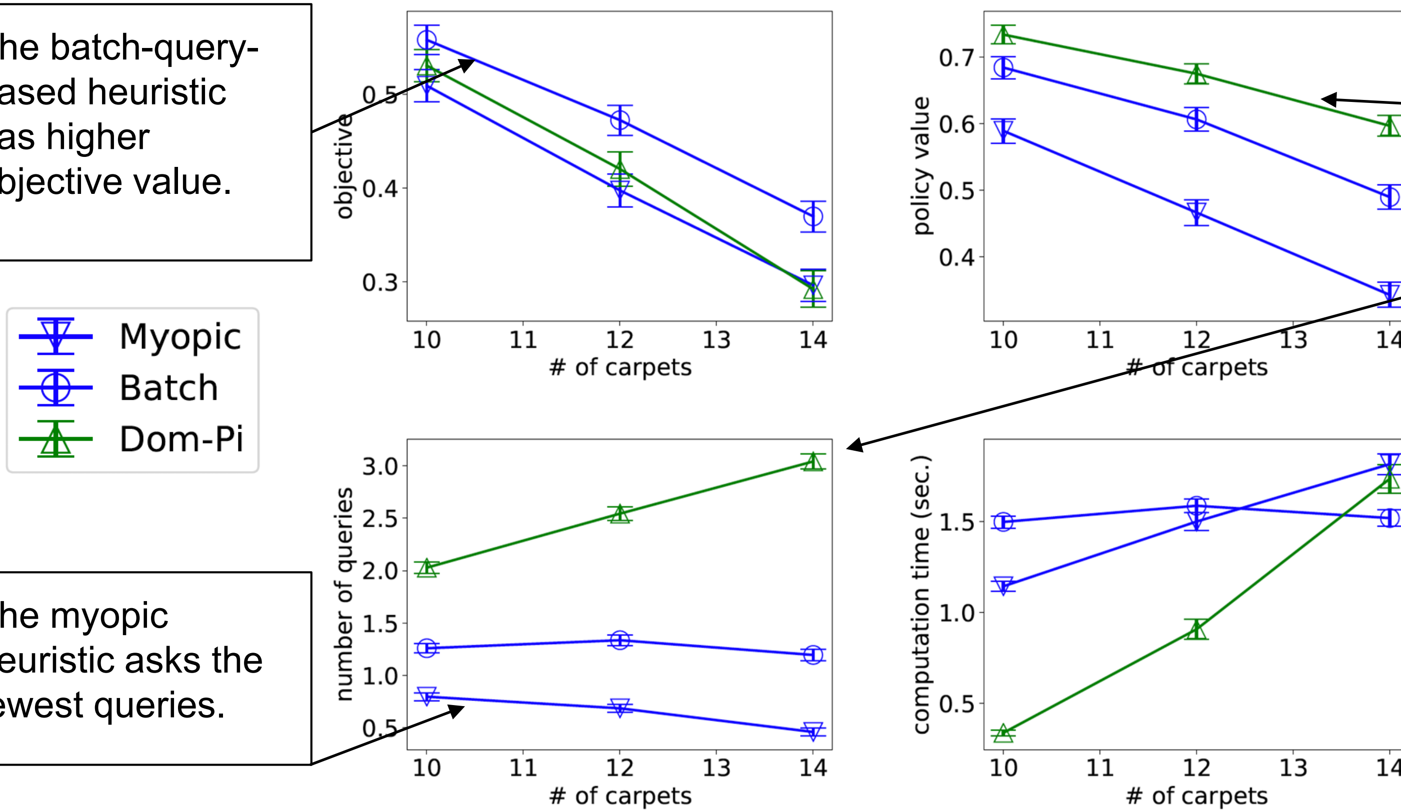
Reward & Safety Uncertainty: Unsure what to achieve and what to avoid!

Sequential Querying: Keep asking a reward or safety query until cost of querying exceeds expected improvement. Challenge: What to ask next?

Key **contributions/results**:

- **Formulate the querying problem as a <u>joint optimization</u> problem**
- **Develop heuristics that strike <u>different tradeoffs</u> between cost and effectiveness**

The batch-query-based heuristic has higher objective value.



The dominating-policy-based heuristic finds a safe policy with a higher value, but also poses more queries.

The myopic heuristic asks the fewest queries.

Tractable Commitment-Constrained Autonomy

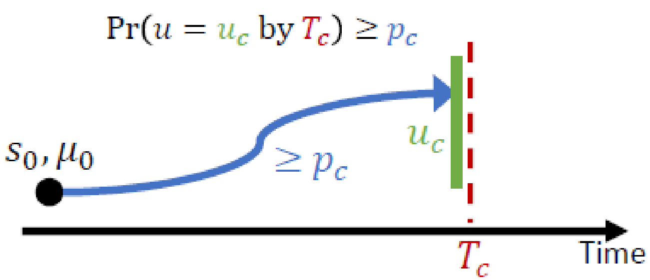
Probabilistic Commitment: Achieve condition u_c by time T_c with prob p_c

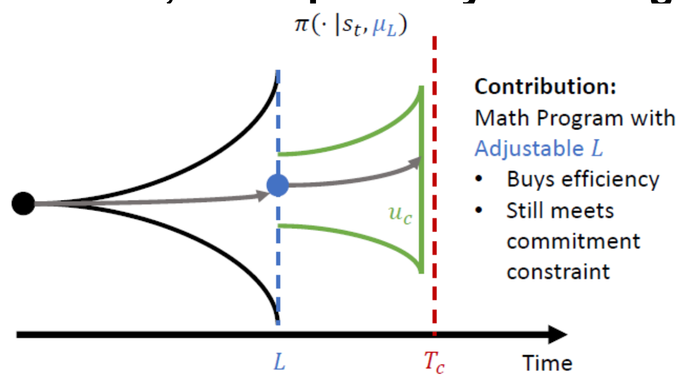
Model Uncertainty: During execution, learn more about goals and actions

CC Autonomy: Autonomously adjust to new model but respect commitment

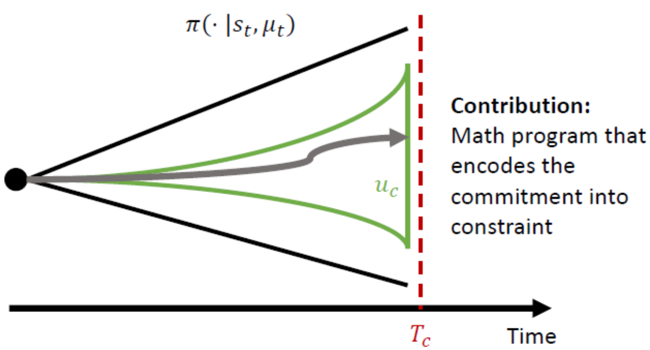
Key **contributions/results**:

- **Prescriptive semantics for commitment constrains autonomy to be dependable**
- **Planning balances local reward and cost, while provably assuring commitment**

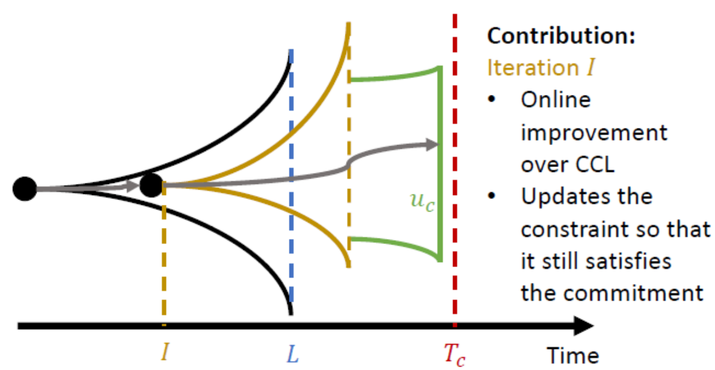




Commitment Constrained Lookahead (CCL)



Commitment Constrained Full Lookahead (CCFL)



Commitment Constrained Iterative Lookahead (CCIL)

Probabilistic Commitment: Recipient's Interpretation

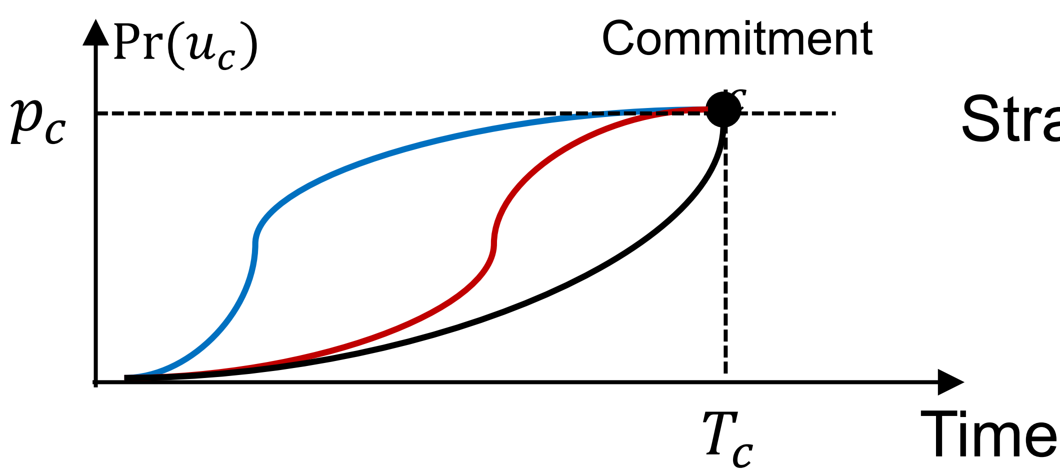
Probabilistic Commitment: Achieve condition u_c by time T_c with prob p_c

Provider's Behavior: Any policy that satisfies commitment. Also provider can flexibly change its policy in response to model updates.

Recipient's Behavior: Policy to preferably take advantage of the commitment by exploiting the establishment of u_c by time T_c with prob p_c

Recipient's Challenge: How to model condition u_c between times T_0 and T_c

Suboptimality: Difference between performance of recipient's policy given its hypothesis about how to model the condition u_c and its real behavior

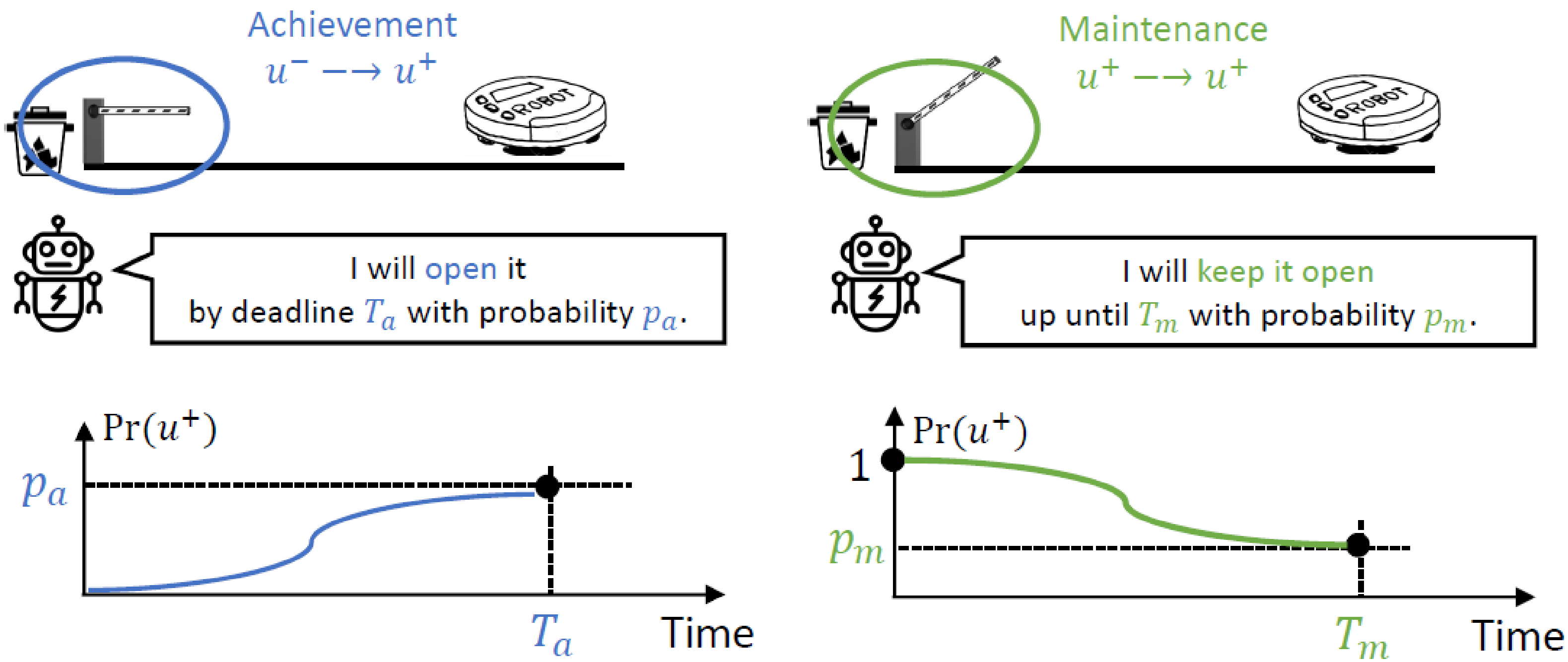


Strategy $\hat{P}_u(\cdot) \xrightarrow{c} \hat{P}_u(c) \rightarrow \hat{\pi}$

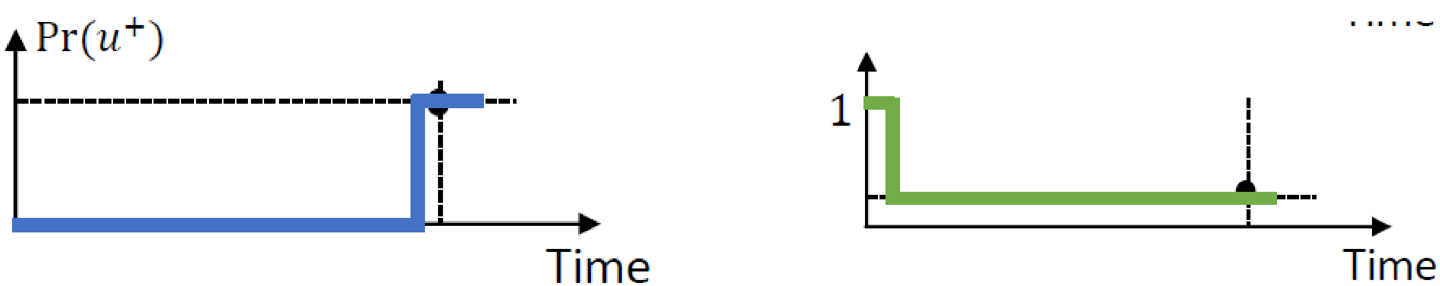
$$\text{Suboptimality} = v^* - v^{\hat{\pi}}$$

Probabilistic Commitment: Recipient's Interpretation

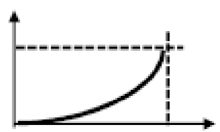
Achievement vs Maintenance

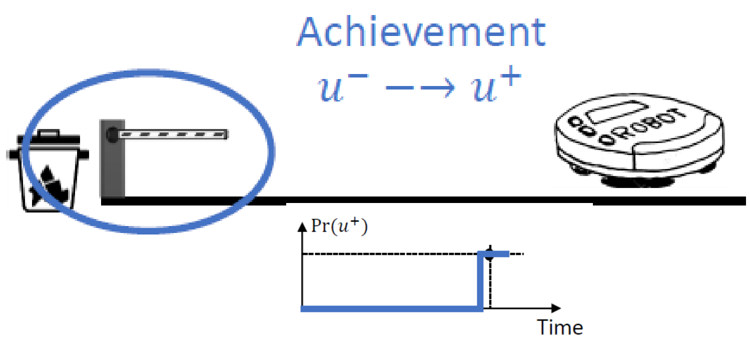


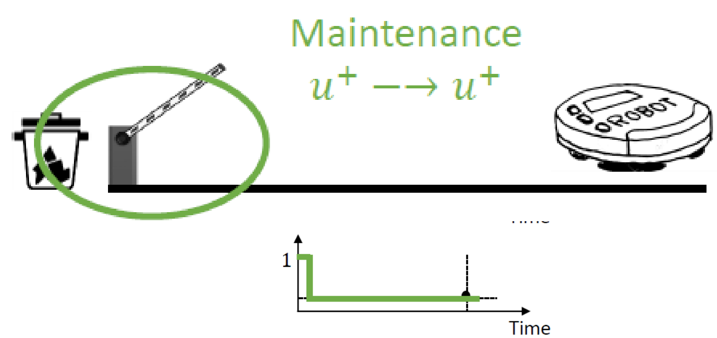
Pessimistic (?) Hypothesis: Minimum Enablement Duration



Suboptimality for Achievement vs Maintenance

$\hat{P}_u(\cdot)$	Suboptimality / $(v^+ - v^-)$	
	Achievement	Maintenance
Min Enablement (Step Function)	≤ 1 Better than $\hat{P}_u^-$	> 1 Can be worse than $\hat{P}_u^-$
Max Enablement (Step Function)	> 1	> 1
Min Value Timing (Step Function)	≤ 1	> 1
Minimax Regret Timing (Step Function)	> 1	> 1
Constant Toggling 	> 1	> 1

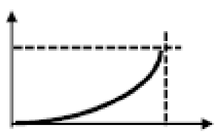




If gate unexpectedly opens earlier than modeled, performance can only improve.

If gate unexpectedly stays open longer than modeled, performance can get worse!

Suboptimality for Achievement vs Maintenance

$\hat{P}_u(\cdot)$	Suboptimality / $(v^+ - v^-)$	
	Achievement	Maintenance
Min Enablement (Step Function)	≤ 1 Better than $\hat{P}_u^-$	> 1 Can be worse than $\hat{P}_u^-$
Max Enablement (Step Function)	> 1	> 1
Min Value Timing (Step Function)	≤ 1	> 1
Minimax Regret Timing (Step Function)	> 1	> 1
Constant Toggling 	> 1	> 1

Conclusion:

- Maintenance commitments are inherently harder…
- But they are harder for the recipient, not for the provider!

Implications:

- On limitations of existing commitment-based coordination frameworks
- On maintenance commitments needing more details
  - Sacrifice some flexibility of the provider to improve performance of the recipient

Querying for Optimal Cooperative Commitments

Cooperative Commitments: Find T_c and p_c maximizing joint reward

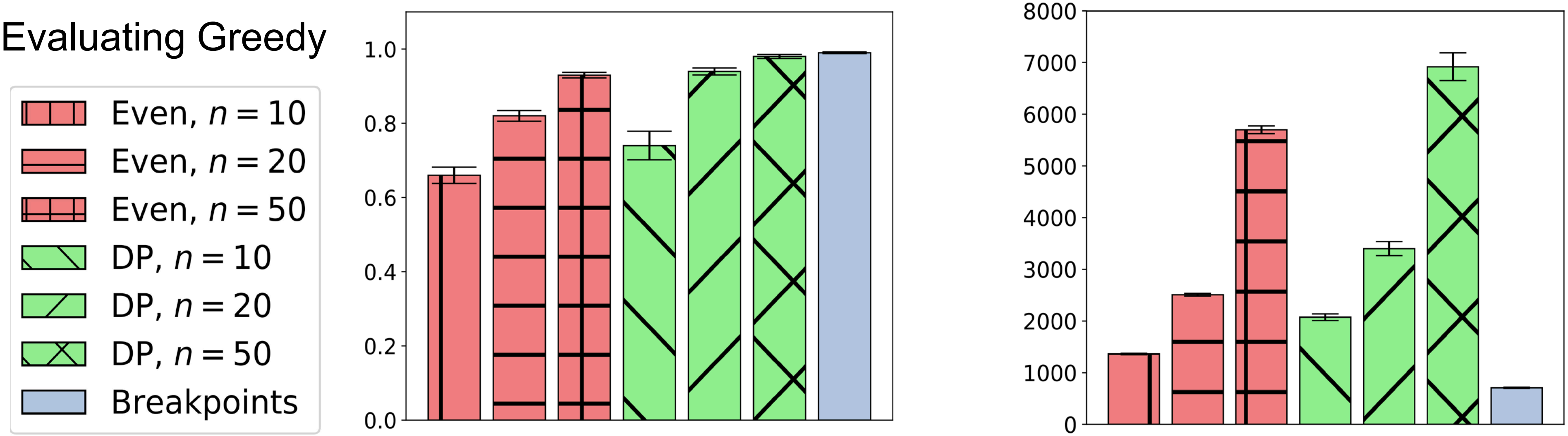
Decentralized: Each agent only knows its own reward function

Querying Challenge: Too many combinations of T_c and p_c to ask about!

Key **contributions/results**:

- **Prove structural regularities between parameters T_c and p_c and agents' rewards**
- **Use regularities to prune large portions of the space of T_c and p_c combinations**
- **Exploit submodularity to form approximately optimal *k*-ary commitment query**

Evaluating Greedy



EUS, k = 5

Runtime (sec.)

Best performance ……………….at lowest cost

List of Publications, Awards, Honors, etc. Attributed to the Grant

1. Qi Zhang, Edmund H. Durfee, Satinder P. Singh. "Semantics and algorithms for trustworthy commitment achievement under model uncertainty." *Autonomous Agents and Multi-Agent Systems* 34(19):1-35, 2020.
2. Prashant Doshi, Piotr Gmytrasiewicz, and Edmund Durfee. "Recursively Modeling Other Agents for Decision Making: A Research Perspective," *Artificial Intelligence*, 279, 2020.
3. Qi Zhang, Edmund H. Durfee, and Satinder Singh. "Modeling Probabilistic Commitments for Maintenance Is Inherently Harder Than for Achievement," *In Proceedings of the Association for the Advancement of AI Conference (AAAI20)*, February 2020.
4. Shun Zhang, Edmund H. Durfee, and Satinder Singh. "Querying to Find a Safe Policy Under Uncertain Safety Constraints in Markov Decision Processes," In *Proceedings of the Association for the Advancement of AI Conference (AAAI20)*, February 2020.
5. Qi Zhang, Richard Lewis, Satinder Singh, and Edmund H. Durfee. "Learning to Communicate and Solve Visual Blocks-World Tasks." In *Proceedings of the Association for the Advancement of AI Conference (AAAI19)*, pages 5781-5788, January 2019.
6. Shun Zhang, Edmund H. Durfee, and Satinder Singh. "Minimax-Regret Querying on Side Effects for Safe Optimality in Factored Markov Decision Processes." In *Proceedings of the International Joint Conference on Artificial Intelligence (IJCAI18)*, pages 4867-4873, July 2018.
7. Shun Zhang, Edmund H. Durfee, and Satinder Singh. "Approximately-Optimal Queries for Planning in Reward-Uncertain Markov Decision Processes." In *Proceedings of the International Conference on Automated Planning and Scheduling (ICAPS17)*, pages 339-347, 2017.
8. Qi Zhang, Satinder Singh, Edmund H. Durfee. "Minimizing Maximum Regret in Commitment Constrained Sequential Decision Making." In *Proceedings of the International Conference on Automated Planning and Scheduling (ICAPS17)*, pages 348-356, 2017.
9. Qi Zhang, Edmund H. Durfee, Satinder Singh, Anna Chen, and Stefan J. Witwicki. "Commitment Semantics for Sequential Decision Making under Reward Uncertainty." In *Proceedings of the International Joint Conference on Artificial Intelligence (IJCAI16)*, pages 3315-3323, August 2016.

List of Publications, Awards, Honors, etc. Attributed to the Grant

10. Qi Zhang, Edmund Durfee, and Satinder Singh. "Efficient Querying for Cooperative Commitments." In *OptLearnMAS Workshop at AAMAS 2020*, May 2020.
11. Shun Zhang, Efficiently *Finding Approximately-Optimal Queries for Improving Policies and Increasing Safety*, PhD Thesis, University of Michigan, April 2020.
12. Qi Zhang, *Making and Keeping Probabilistic Commitments for Trustworthy Multiagent Coordination*, PhD Thesis, University of Michigan, April 2020.
13. Shun Zhang, Edmund H. Durfee, and Satinder Singh. "Querying to Find a Safe Policy Under Uncertain Safety Constraints in Markov Decision Processes." *In Notes of the NeurIPS-2019 Workshop on Safety and Robustness in Decision Making*, December 2019.
14. Qi Zhang, Edmund H. Durfee, and Satinder P. Singh. "Computational Strategies for the Trustworthy Pursuit and the Safe Modeling of Probabilistic Maintenance Commitments." In *Notes of the IJCAI 2019 Workshop on Artificial Intelligence Safety (AISafety)*, August 2019.
15. Shun Zhang, Edmund H. Durfee, and Satinder Singh. "On Querying for Safe Optimality in Factored Markov Decision Processes (Extended Abstract)," *Proceedings of the 2018 Autonomous Agents and Multiagent Systems (AAMAS)*, pages 2168-2170, July 2018.
16. Qi Zhang, Edmund H. Durfee, and Satinder Singh. "Challenges in the Trustworthy Pursuit of Maintenance Commitments Under Uncertainty." *2018 TRUST@AAMAS Workshop*, pages 75-86, July 2018.
17. Edmund H. Durfee and Satinder Singh. "On the Trustworthy Fulfillment of Commitments." In Nadine Osman and Carles Sierra (eds.) *Autonomous Agents and Multiagent Systems: AAMAS 2016 Workshops Best Papers*, Springer Lecture Notes in Artificial Intelligence, pages 1-13, 2016.
18. Edmund H. Durfee and Satinder Singh. "On the Trustworthy Fulfillment of Commitments." In 2016 TRUST@*AAMAS Workshop*, May 2016. (Best Paper Award)
19. Edmund H. Durfee and Satinder Singh. "Commitment Semantics for Sequential Decision Making Under Reward Uncertainty." In *AAAI Fall Symp on Sequential Decision Making for Intelligent Agents*, Nov 2015.