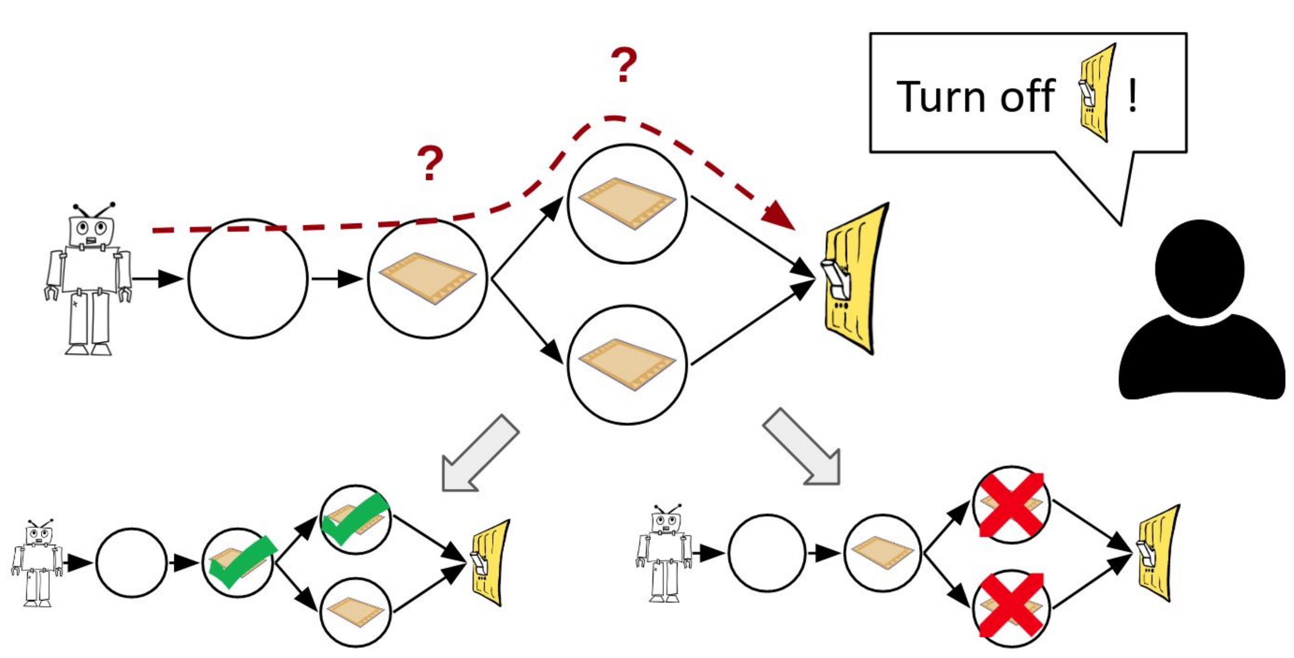
When a human user specifies a goal for a robot to achieve, the robot may find its policy cause **side effects** that the user may think **unsafe**. How should the robot **efficiently query** the human to find a **guaranteed-safe** policy (if one exists)?

Querying to Find a Safe Policy Under Uncertain Safety Constraints in Markov Decision Processes

Shun Zhang, Edmund H. Durfee, and Satinder Singh. University of Michigan

MOTIVATION

- Robot's policy to optimize its user's reward may have unexpected, possibly unsafe, **side effects**.
- Robot can **query** the user to find out which (if any) side effects are safe.
- Robot queries until it finds a safe policy, or proves that none exists.

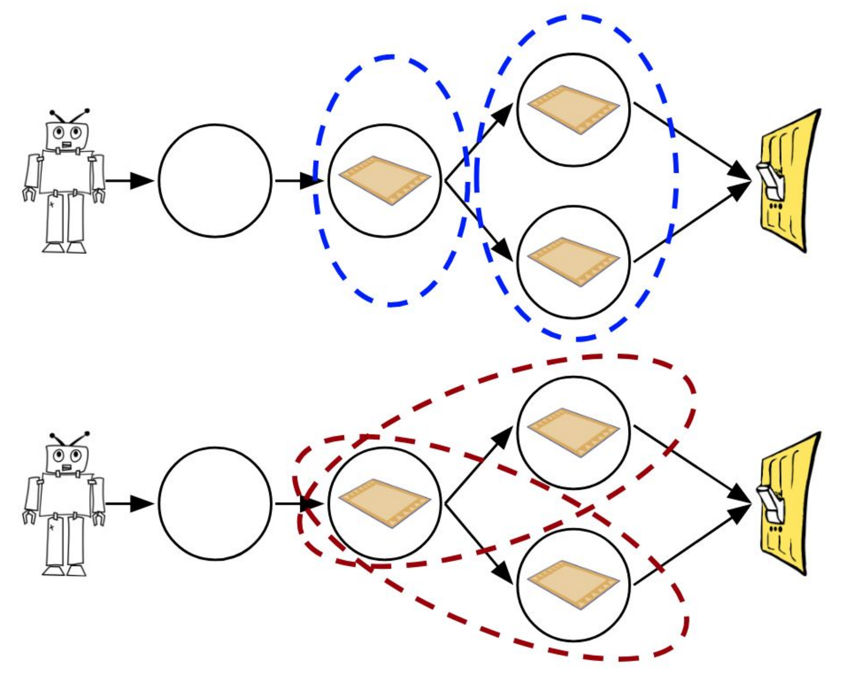


OBJECTIVE

Minimize the number of queries needed, in expectation, to either find a safe policy or prove none exists.

METHOD

Observation: Finding a safe policy and proving that no safe policy exists each corresponds to a **set cover problem**.



Solution: **Efficient iterative query selection algorithms** that solve both set cover problems simultaneously.

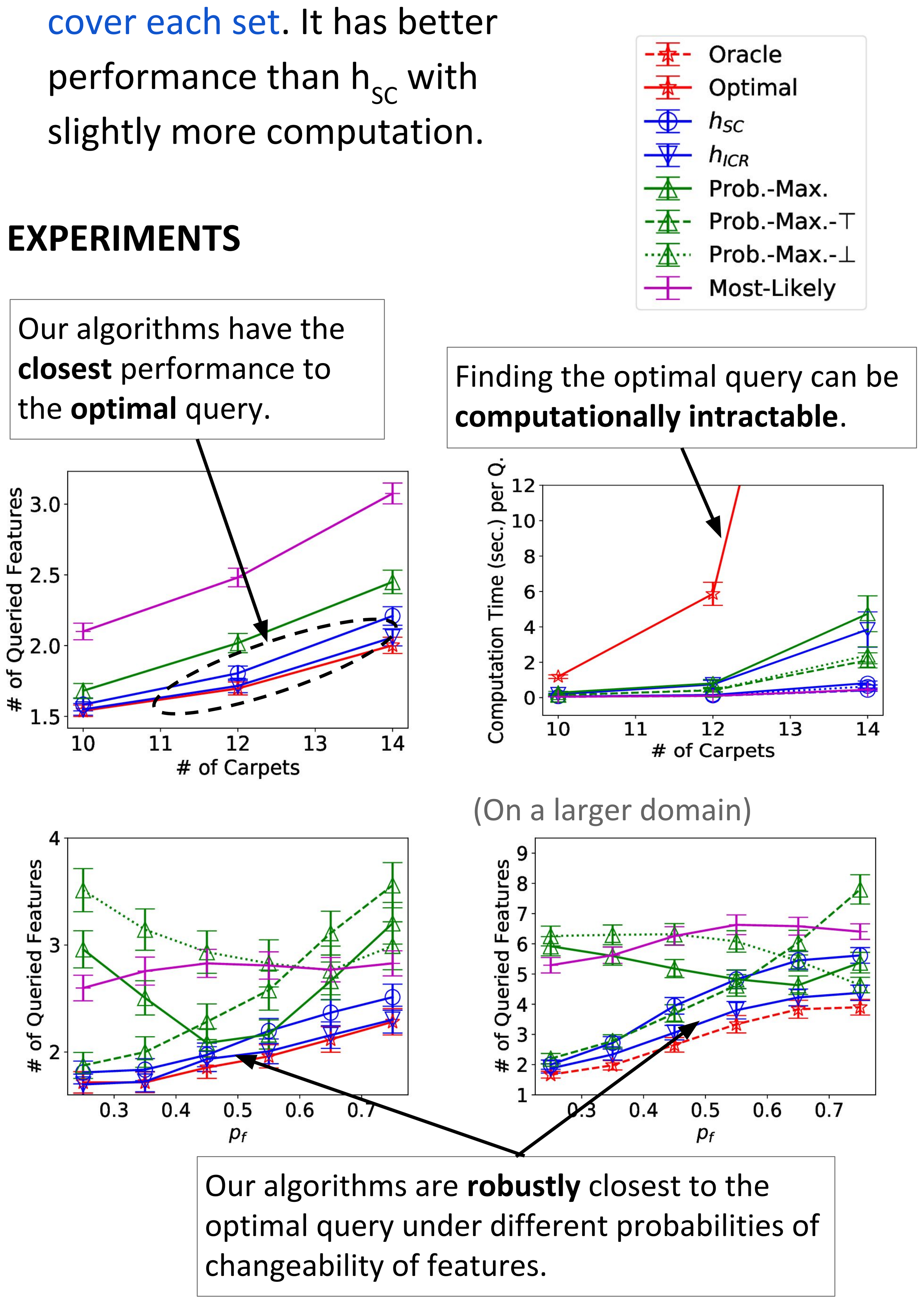
RESULTS

Our query algorithms find **better queries** than greedy-heuristic algorithms and are **computationally cheaper** than brute-force methods.

OUR QUERY-SELECTION ALGORITHMS

- Our algorithms are based on irreducible infeasible sets (IIS) (Chinneck, 2007) and adaptive submodularity (Golovin and Krause 2011).
- **h_{SC} (set cover).** Robot selects the query that makes the most progress in covering both sets in expectation.
- **h_{ICR} (inverse cover ratio).** Robot selects a query by estimating the number of queries needed to cover each set. It has better performance than h_{SC} with slightly more computation.

EXPERIMENTS



Scan me to download the paper