Real-time Planning as Data-driven Decision-making

Maximilian Fickert, Tianyi Gu, Leonhard Staut,
Sai Lekyang, Wheeler Ruml, Jörg Hoffmann, Marek Petrik

UNIVERSITÄT DES SAARLANDES

University of New Hampshire

Motivation

- Real-time Search: agent has bounded time to search next action for execution (deterministic, single agent)
- Must efficiently allocate limited number of search node expansions. Classical solutions are often intuitive adaptions of offline search, such as RTA* and LSS-LRTA*

What if we design for real-time planning from scratch?

Contribution

AAAI-19: The Nancy Framework

- Nancy Backup, Risk-based lookahead

AAAI-20: Data-Driven Nancy

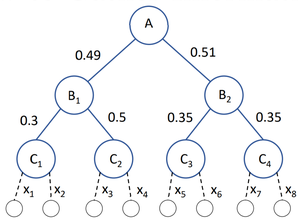
- Replace assumptions with data, completeness proof

PRL-20: More Experiments

- Compare against other methods that try to explore value uncertainty, visualization of actual data

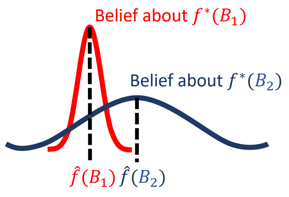
Troublesome Example

How to Gather Information?



Given these search nodes, should agent at A move to B_1 or B_2

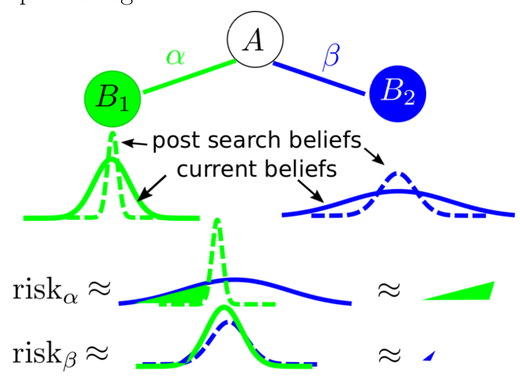
Which Node to Expand?



Should the agent expand nodes on the frontier under B_1 or B_2

Searching use beliefs

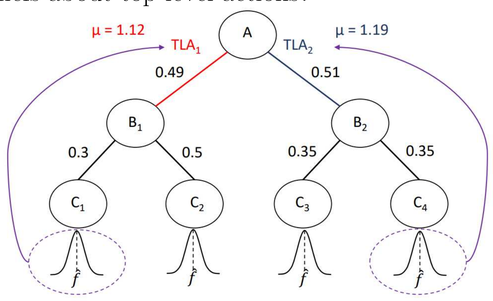
Risk-based Expansion: given beliefs about top-level action values, expand nodes on the frontier under top-level action that minimizes risk, the expected regret



expand under β !

Where do beliefs come from?

Purpose of search is to gather information to inform decision-making process. Which information on the search frontier should be used to form beliefs about top-level actions?



$$\hat{f} = g + \hat{h} = g + h + \epsilon d$$

- Assumption-based Nancy: Trucated Gaussian based on h and d
- Data-Driven Nancy: replace the assumptions with data

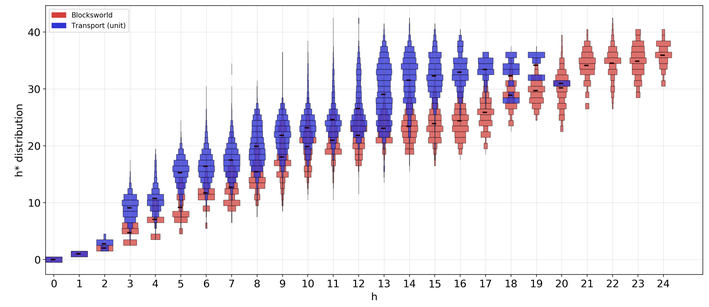
Gathering Data

- Run weighted-A* on random problems
- Collect all states
- For each observed h value pick common states

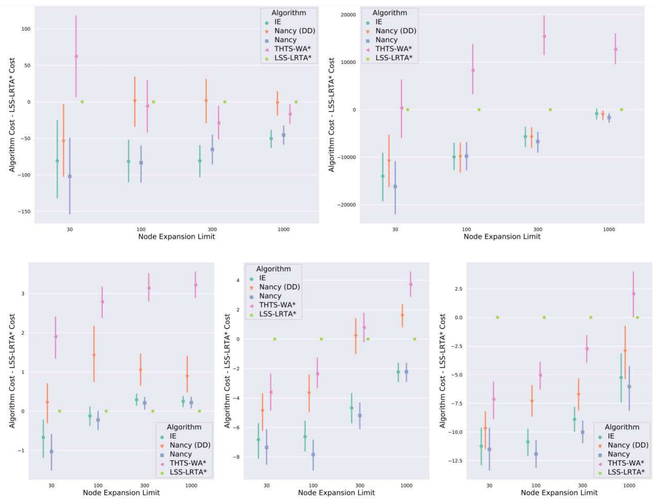
  Compute h^*

Experiments

Example h^*: Transport vs Blocks World



Compare against MCTS and Interval Estimation:



MCTS:

- Trial-based Heuristic Tree Search (THTS): An MCTS adaption to deterministic planning problems.
- For real-time search, we replace the A* expansion strategy with THTS algorithm.

Interval Estimation:

- Interval Estimation applies the philosophy of optimism in the face of uncertainty.
- For real-time search, IE choose the TLA with the lowest lower bound on the 95% confidence interval of the backed up belief.
- IE naturally practices the spirit of uncertainty-based exploration in a very computationally efficient way.