Real-time Planning as Data-driven Decision-making

Maximilian Fickert[1] and Tianyi Gu[2] and Leonhard Staut[1] and
Sai Lekyang[2] and Wheeler Ruml[2] and
Jörg Hoffmann[1] and Marek Petrik[2]

[1] UNIVERSITÄT DES SAARLANDES

[2] University of New Hampshire

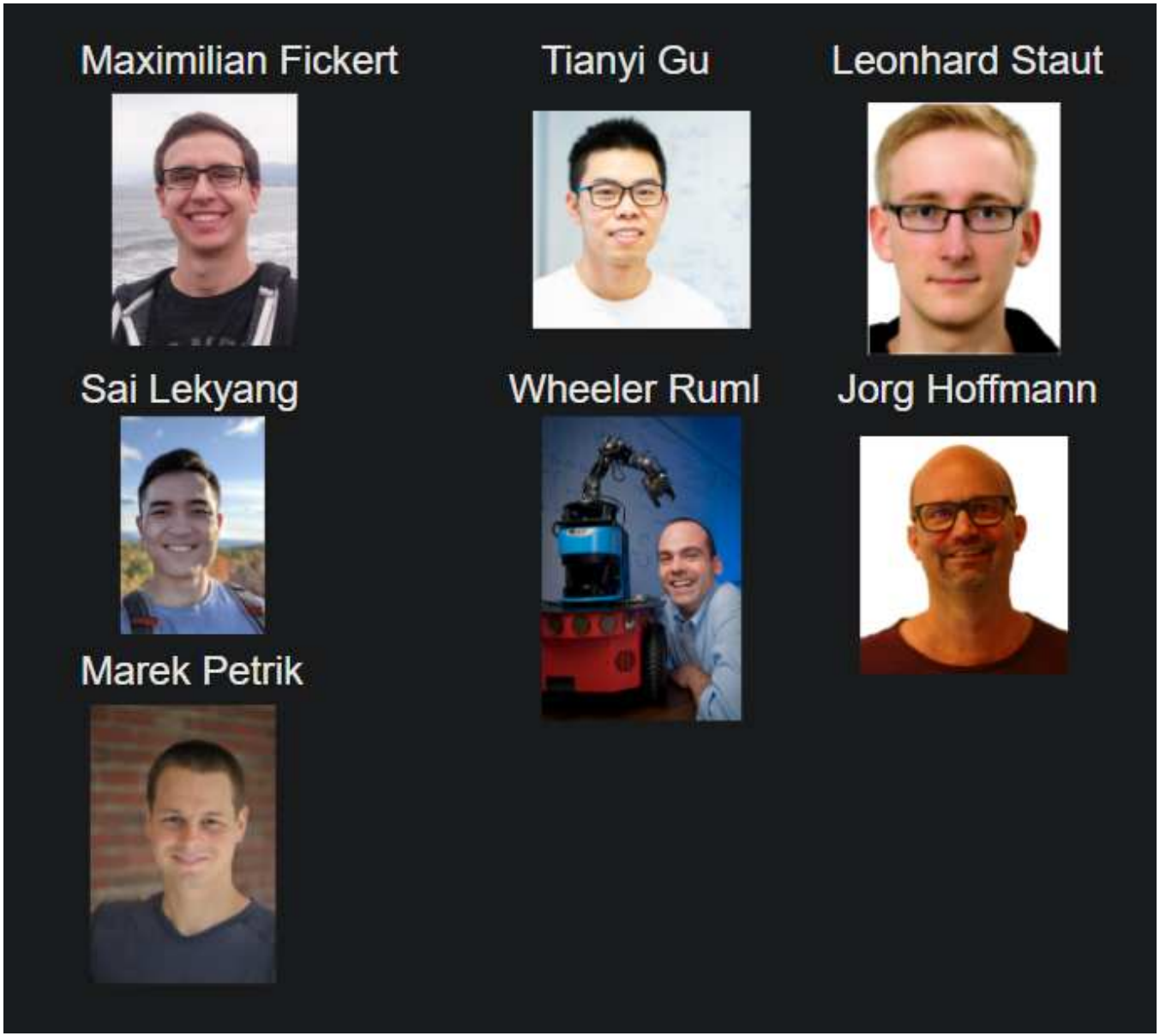
The authors

Introduction
- The Authors
- Real-time Search
- An Example
- This work

Nancy

Results

Conclusions



Problem Setting: Real-time Heuristic Search

Introduction
- The Authors
- Real-time Search
- An Example
- This work

Nancy

Results

Conclusions

Environments:

single agent
discrete state, discrete action
fully observability
online planning: interleaving planning and execution
deterministic state transition

Real-time heuristic search:

can't wait until have full plan
return the next action within a given time bound

An Example of Real-time Heuristic Search

Introduction
- The Authors
- Real-time Search
- An Example
- This work

Nancy

Results

Conclusions

An example: path finding

agent performs search for a bounded time

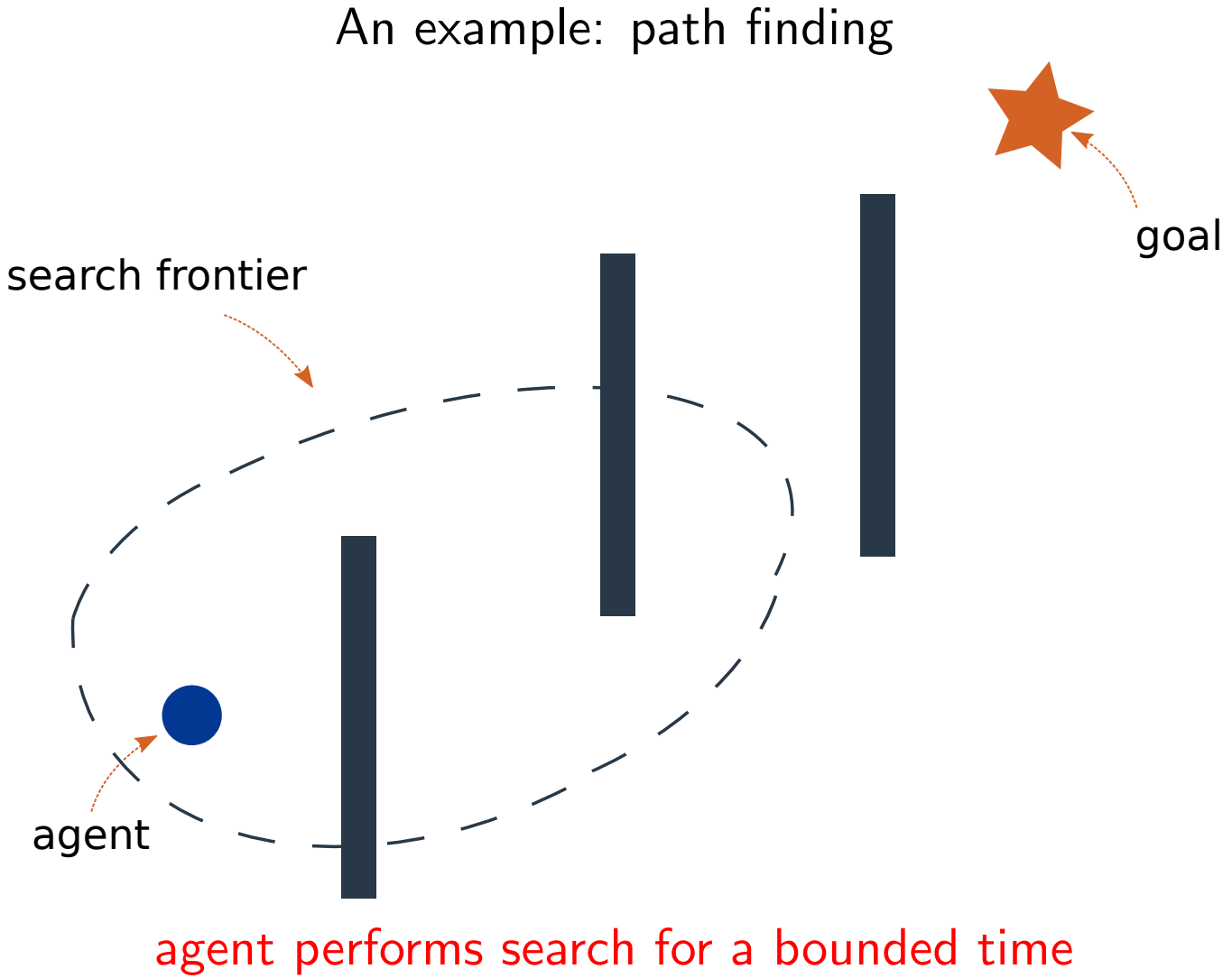
An Example of Real-time Heuristic Search

Introduction
- The Authors
- Real-time Search
- An Example
- This work

Nancy

Results

Conclusions



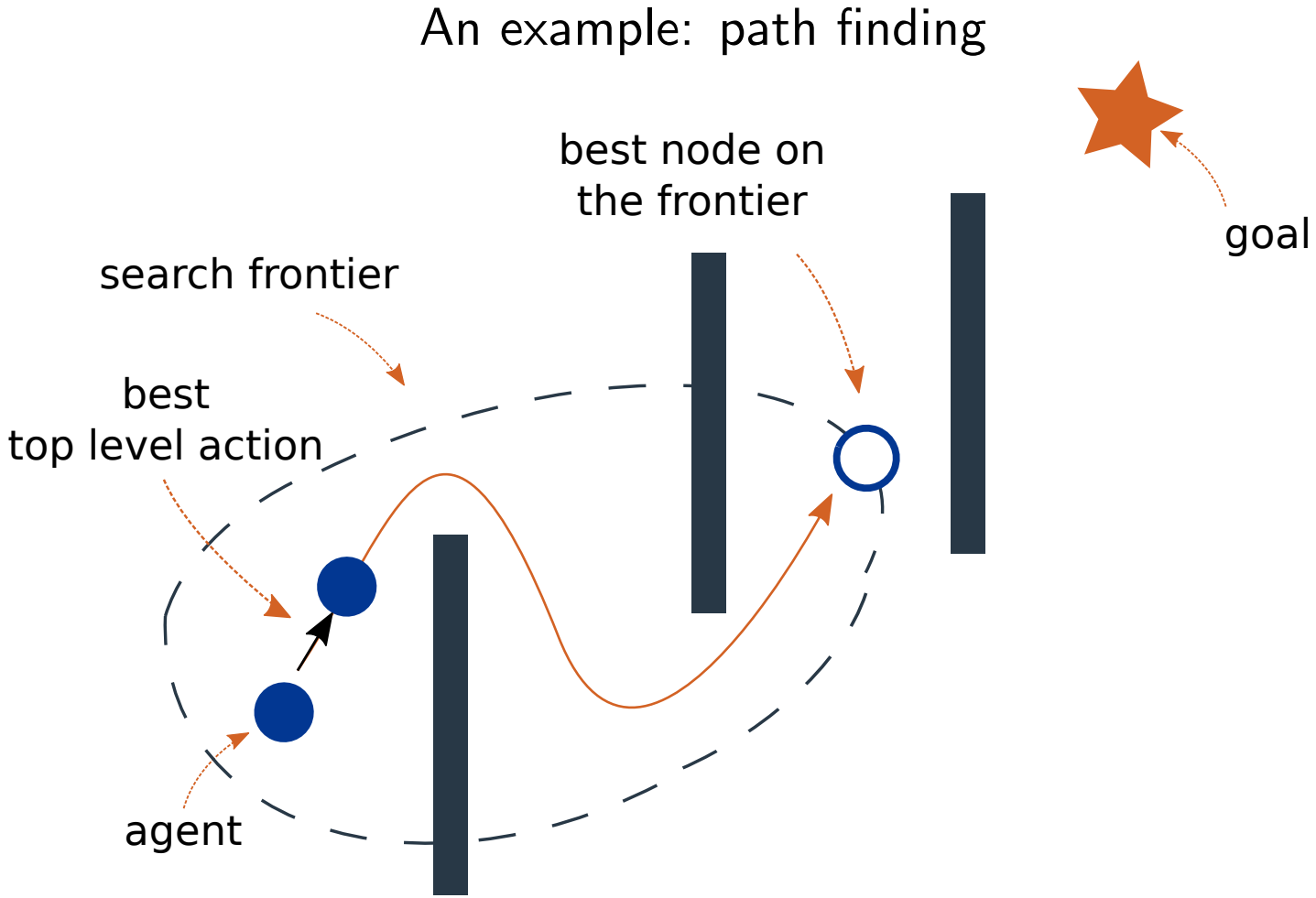
An Example of Real-time Heuristic Search

Introduction
- The Authors
- Real-time Search
- An Example
- This work

Nancy

Results

Conclusions



agent commits to best action and executes

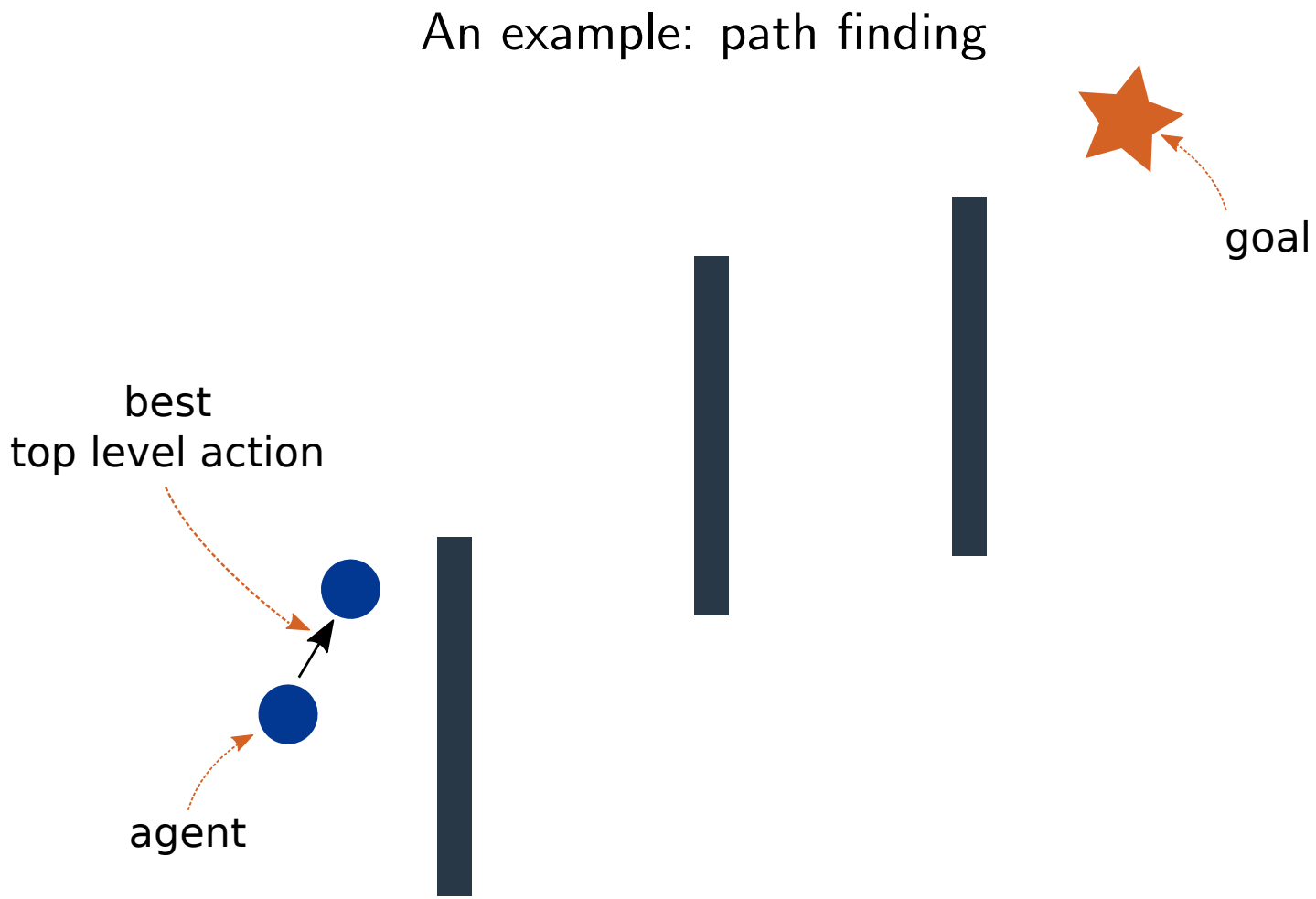
An Example of Real-time Heuristic Search

Introduction
- The Authors
- Real-time Search
- An Example
- This work

Nancy

Results

Conclusions



agent commits to best action and executes

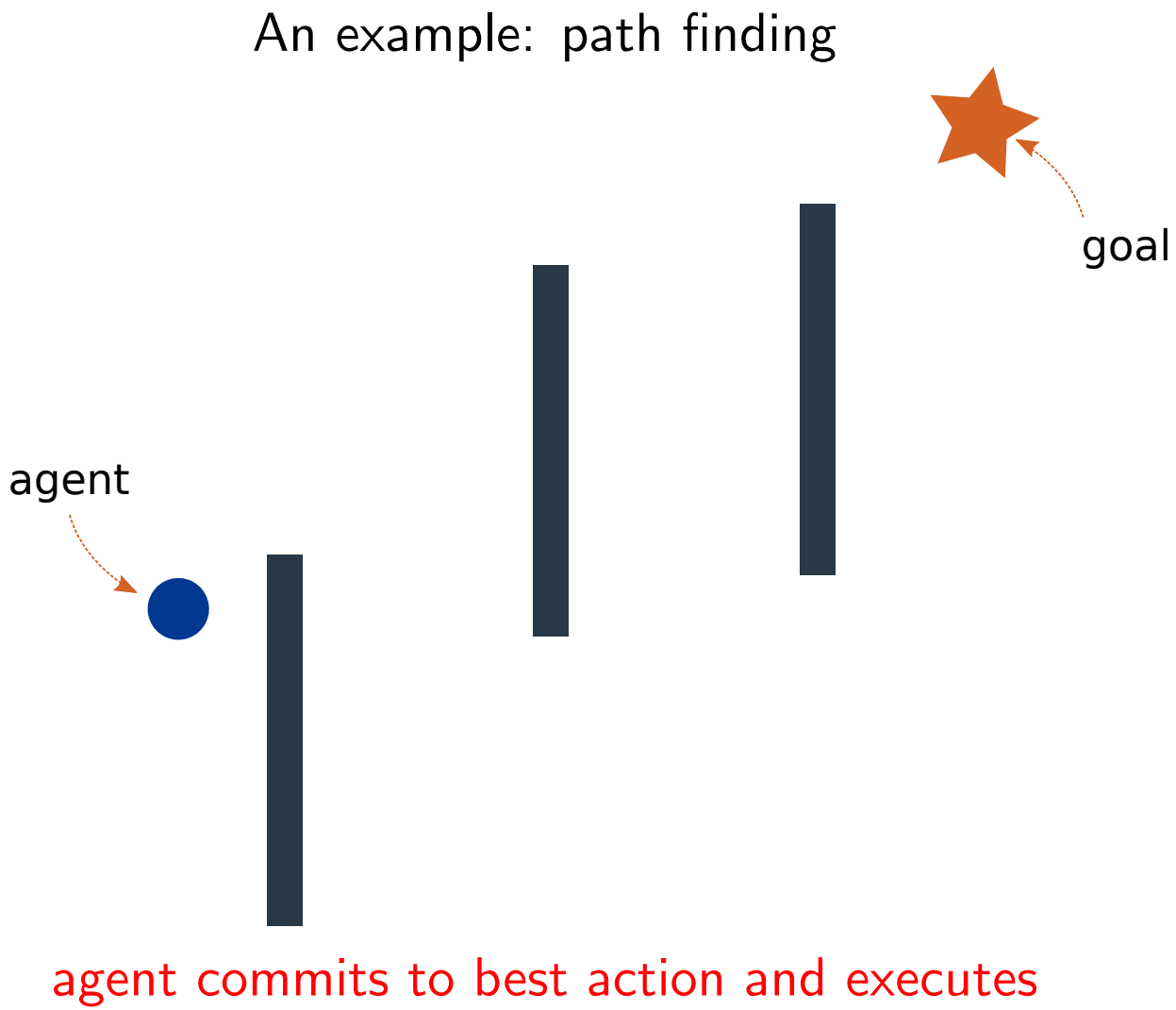
An Example of Real-time Heuristic Search

Introduction
- The Authors
- Real-time Search
- An Example
- This work

Nancy

Results

Conclusions



An Example of Real-time Heuristic Search

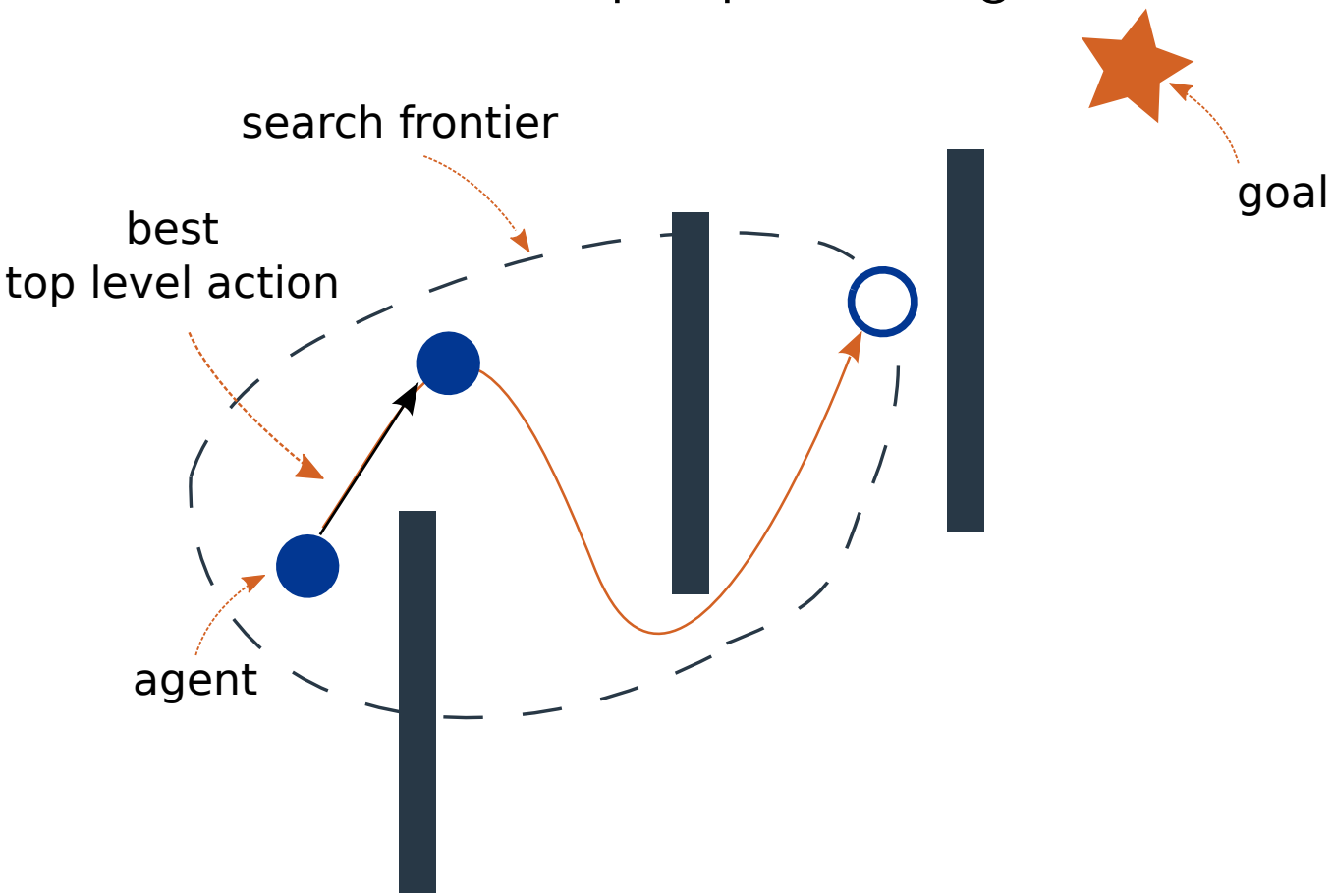
Introduction
- The Authors
- Real-time Search
- An Example
- This work

Nancy

Results

Conclusions

An example: path finding



online planning: interleaving search and action execution

"receding horizon control"

Our Focus: Rational Real-time Planning

Introduction
- The Authors
- Real-time Search
- An Example
- This work

Nancy

Results

Conclusions

Challenge:

limited time $\rightarrow$ limited lookahead $\rightarrow$ value uncertainty

This work:

A beginning of rational approach for real-time planning

Borrowing idea from Bayesian RL

Simplified setting

Put focus on bounded rationality

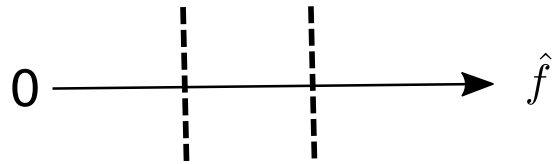
Introduction
Nancy
- Lookahead
- Our Approach

Results
Conclusions

The Nancy Framework

Lookahead Strategy: A Troublesome Example

Introduction
Nancy
Lookahead
Our Approach
Results
Conclusions



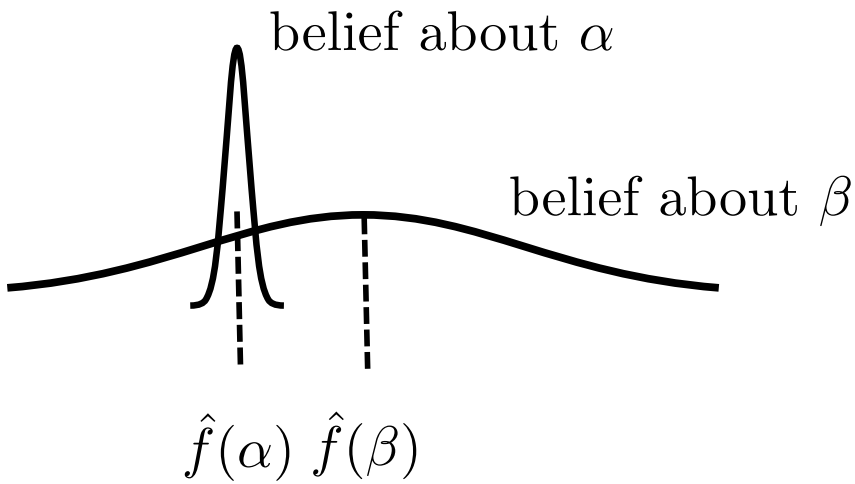
$\hat{f}(\alpha)$ $\hat{f}(\beta)$

$\hat{f}$ is expected value

Should an agent expand nodes under α or β?

Lookahead Strategy: A Troublesome Example

Introduction
Nancy
Lookahead
Our Approach
Results
Conclusions



$\hat{f}$ is expected value

Should an agent expand nodes under α or β?

$\hat{f}$ is not the answer: what to do?
want to maximize value of information
need to consider uncertainty of estimates

Lookahead Strategy: A Troublesome Example

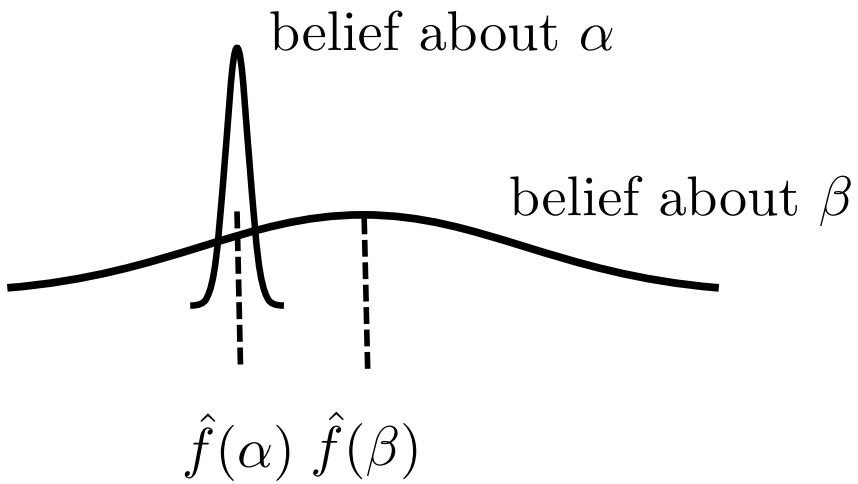
Introduction

Nancy

- Lookahead
- Our Approach

Results

Conclusions



$\hat{f}$ is expected value

Should an agent expand nodes under α or β?

Our approach: expand nodes which minimize expected regret

$$\mathbb{E}\left[\underbrace{f^*(\alpha) - f^*(\beta)}_{\text{our regret}} \;\middle|\; \underbrace{f^*(\beta) < f^*(\alpha)}_{\text{in cases when } \alpha \text{ not best}}\right]$$

Our Approach: The Nancy Framework

Introduction

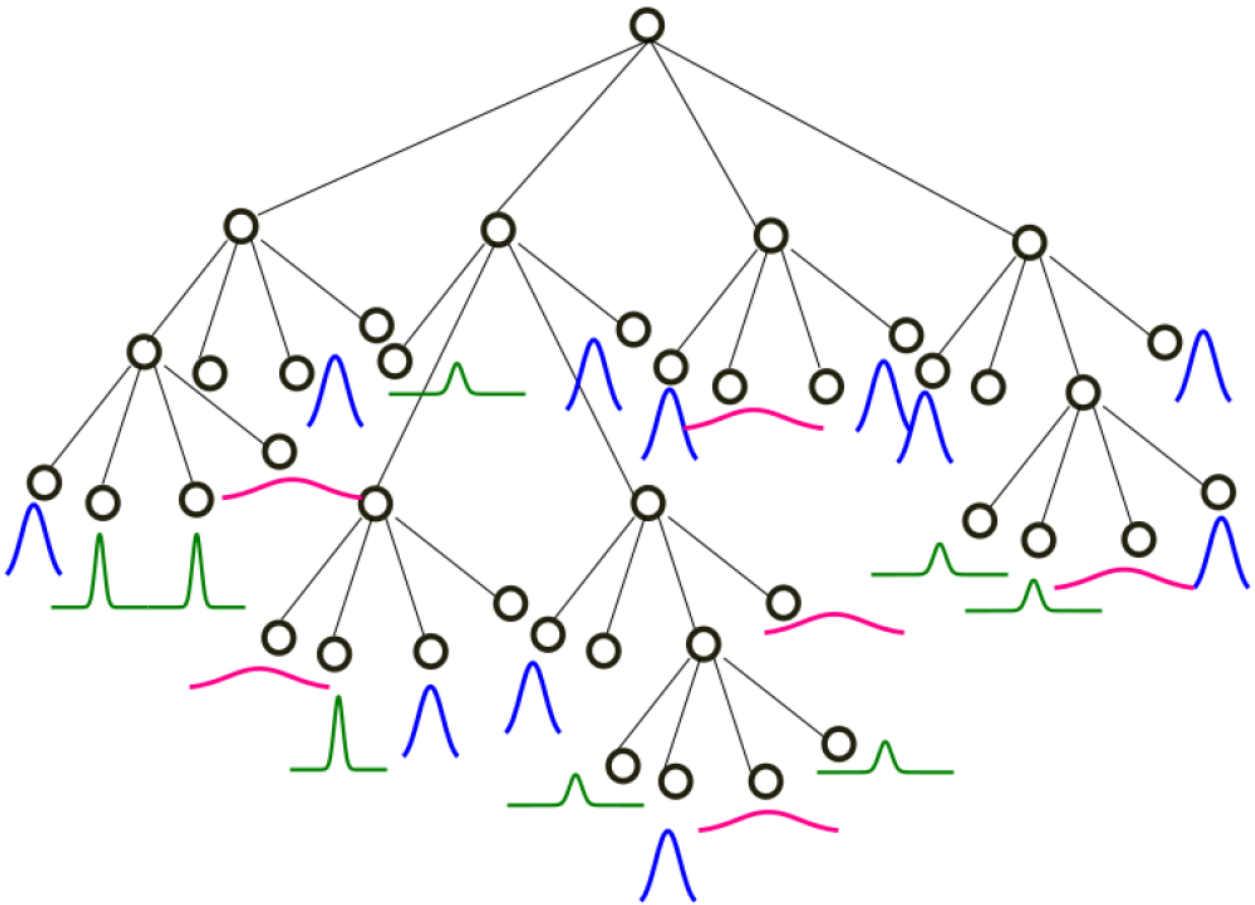
Nancy

- Lookahead
- Our Approach

Results

Conclusions

How to best use limited number of expansions?



Our Approach: The Nancy Framework

Introduction
Nancy
■ Lookahead
■ Our Approach
Results
Conclusions

How to best use limited number of expansions?

estimate value uncertainty

one option: uncertainty correlate distance-to-go

see details in Mitchell et al., AAAI 2019

use the estimate to guide search

see details in Mitchell et al., AAAI 2019

can use data to estimate distributions

see details in Fickert et al., AAAI 2020

This PRL-20 paper:

comparison against other methods that try to exploit value uncertainty

visualization of distributions estimated from data

Introduction
Nancy
Results
- Data Viz
- Results
Conclusions

Results

Example h^* distribution: Transport vs Blocks World

Introduction

Nancy

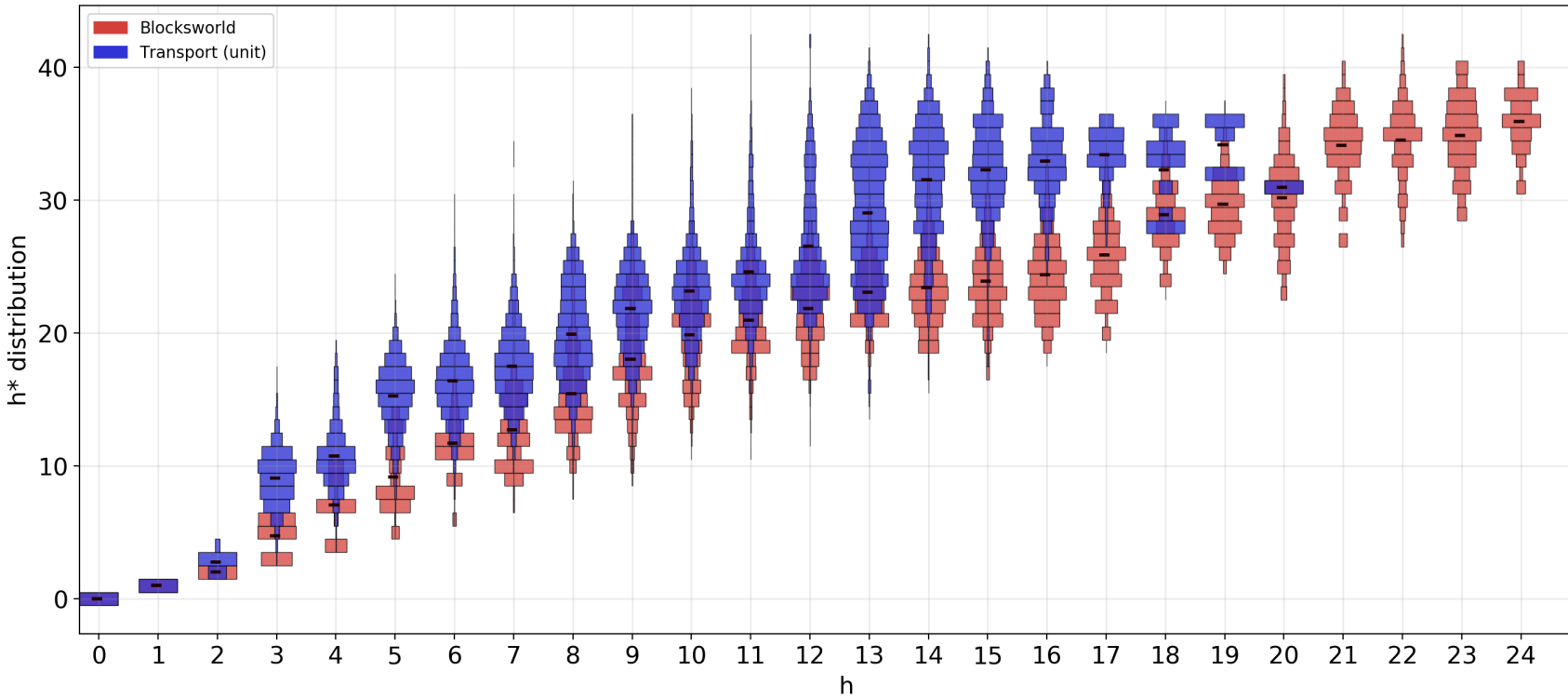
Results

Data Viz

Results

Conclusions

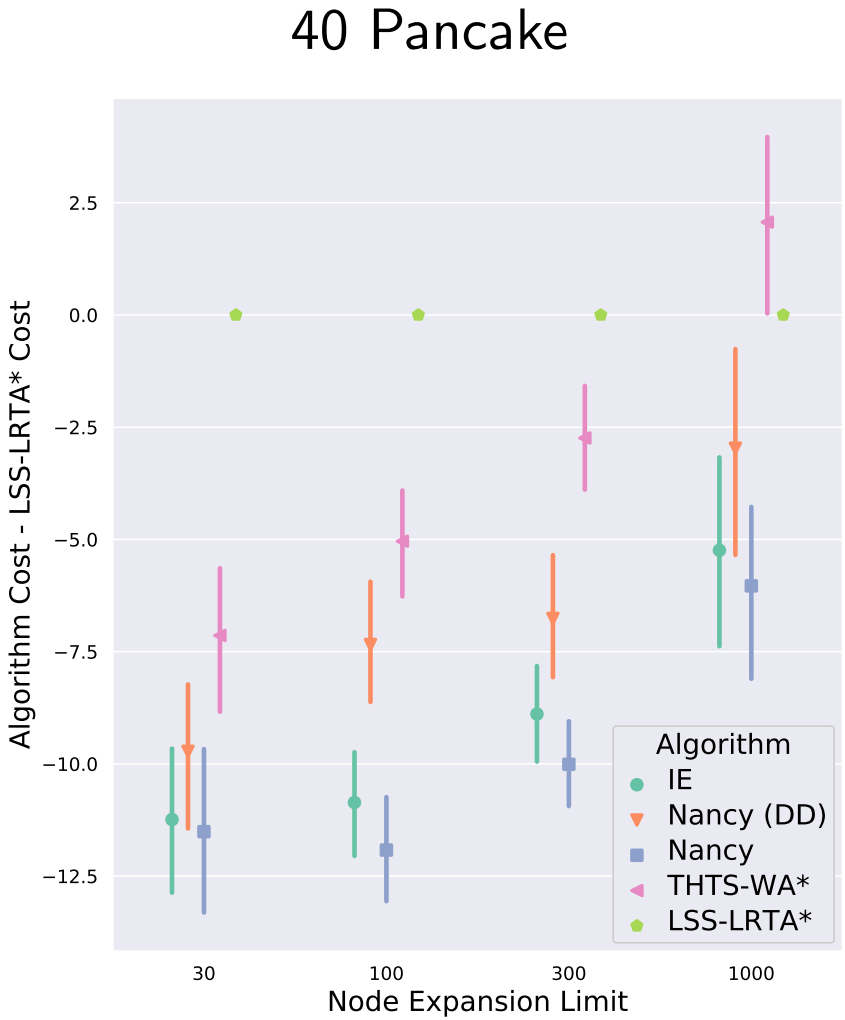
What does the actual cost-to-go value uncertainty distribution look like?



Beliefs are different from domain to domain

Comparison to IE and MCTS on Classic Search Domains

Introduction
Nancy
Results
Data Viz
Results
Conclusions



Nancy outperforms conventional approaches

Introduction
Nancy
Results
Conclusions
■ Summary

Conclusions

Summary

Introduction
Nancy
Results
Conclusions
Summary

- Nancy starts to explore an optimal way of doing online heuristic search
- Nancy is competitive or better than conventional approaches

More broadly:

- Setting isolates the issue: unlike in MDPs or RL, all uncertainty is due to bounded rationality
- Metareasoning about uncertainty pays off, even for deterministic domains!

Questions?

Introduction
Nancy
Results
Conclusions
Questions
Questions?