Planning Domain Modelling Competition

Simon Dold

University of Basel, Switzerland

June 2, 2023

Motivation

IPC History

IPC1998 (called the 1998 AI Planning Systems Competition)

- compare different planning systems
  - measure progress in the filed
- put pressure on the community to develop planners further
- provide benchmark sets for evaluation

IPC History

IPC1998 (called the 1998 AI Planning Systems Competition)

- compare different planning systems
  - measure progress in the filed
- put pressure on the community to develop planners further
- provide benchmark sets for evaluation

The IPC needs benchmark domains

Where do domains come from?

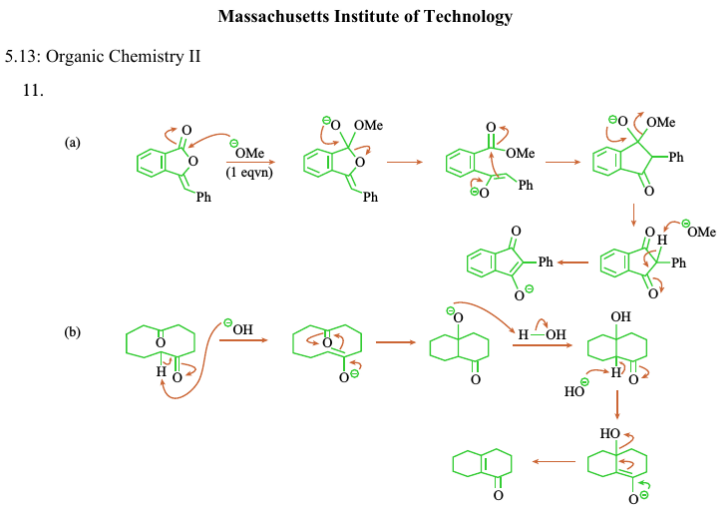
Every IPC has a call for domains

- participants are allowed to submit
- selected by the IPC organizers

IPC History

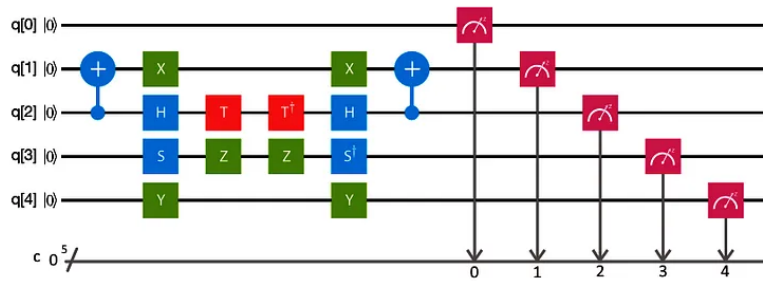
IPC2018 and IPC2023 provided further motivation with the Outstanding Domain Submission Award

Organic Synthesis



Submitted by: Hadi Qovaizi, Arman Masoumi, Anne Johnson, Russell Viirre, Andrew McWilliams, and Mikhail Soutchanski (Ryerson University)

Quantum Circuit Layout Synthesis



https://towardsdatascience.com/what-is-a-quantum-circuit-transpiler-ba9a7853e6f9

Submitted by: Irfansha Shaik and Jaco van de Pol (Aarhus University)

Planning Domain Modelling Competition (PDMC)

Award great Domain Submissions

- transparent criteria
- transparent process

Planning Domain Modelling Competition (PDMC)

Hopes

Planning Domain Modelling Competition (PDMC)

Hopes

- People from field X formalize interesting problems in pddl

Planning Domain Modelling Competition (PDMC)

Hopes

- People from field X formalize interesting problems in pddl
- PDMC awards interesting domains

Planning Domain Modelling Competition (PDMC)

Hopes

- People from field X formalize interesting problems in pddl
- PDMC awards interesting domains
- IPC uses PDMC winner

Planning Domain Modelling Competition (PDMC)

Hopes

- People from field X formalize interesting problems in pddl
- PDMC awards interesting domains
- IPC uses PDMC winner
- IPC domains are used for planner research and development

Planning Domain Modelling Competition (PDMC)

Hopes

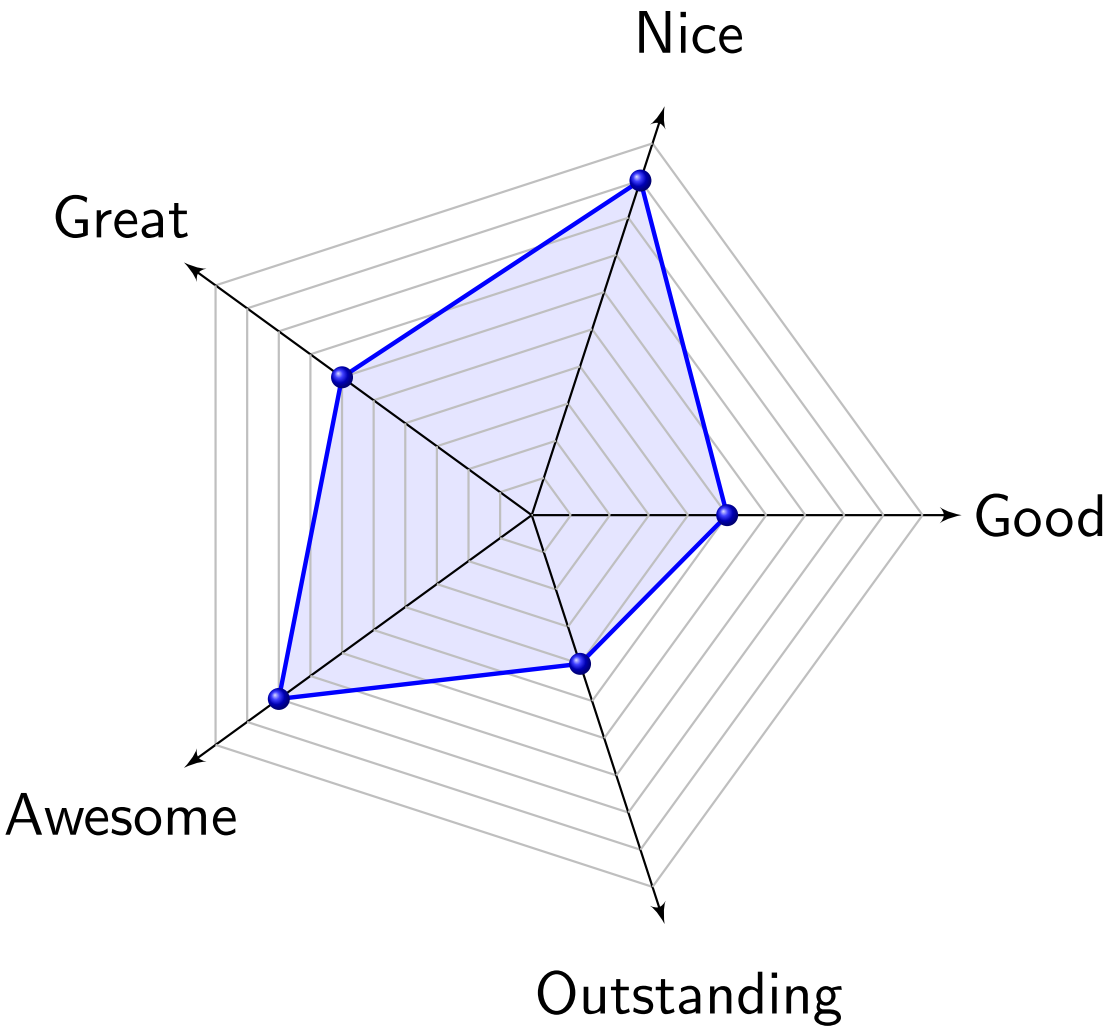
- People from field X formalize interesting problems in pddl
- PDMC awards interesting domains
- IPC uses PDMC winner
- IPC domains are used for planner research and development
- Planning becomes more relevant for field X

Planning Domain Modelling Competition (PDMC)

Hopes

- People from field X formalize interesting problems in pddl
- PDMC awards interesting domains
- IPC uses PDMC winner
- IPC domains are used for planner research and development
- Planning becomes more relevant for field X
- Repeat ...

What makes a domain good?



Property Discussion

- Why do we want this property?
- Qualitative: How do we detect this property?
- Quantitative: How do we measure this property?

Overview

- **Challenging:** not too easy nor to hard
- **Interesting:** real-world relevance
- **Diverse:** different to existing ones
- **Natural Encoding:** less auxiliary mechanisms
- **Adjustable:** difficulty granularity
- **Intrinsic Difficulty:** no scaled up problem
- **Tricky:** provoke shortcomings
- **Optimality Bounds:** optimal plan cost is known

Challenging

Time	box-land
ALPACA 2011	0.18s
EE	**0.14s**
Lowerpifted	0.16s
MyfaSA*-2	0.15s
Stick Salad	0.19s

Time	box-land
ALPACA 2011	ooM
EE	ooM
Lowerpifted	ooT
MyfaSA*-2	**1799s**
Stick Salad	ooM

Challenging

Time	box-land
ALPACA 2011	0.18s
EE	**0.14s**
Lowerpifted	0.16s
MyfaSA*-2	0.15s
Stick Salad	0.19s

Time	box-land
ALPACA 2011	ooM
EE	ooM
Lowerpifted	ooT
MyfaSA*-2	**1799s**
Stick Salad	ooM

Time	box-land
ALPACA 2011	900s
EE	1700s
Lowerpifted	1600s
MyfaSA*-2	1400s
Stick Salad	ooM

Challenging

- **Why?** Total order
- **Qualitative:** domain contains instances...
  - ...not trivial for baseline planner
  - ...not too hard for one planner

  $\rightsquigarrow$ similar to SAT competition
- **Quantitative:** Variance σ

Interesting

Interesting

`movie`

Challenging - **Interesting** - Diverse - Natural Encoding - Adjustable - Intrinsic Difficulty - Tricky - Bounds

Interesting

movie

organic-synthesis

Challenging - **Interesting** - Diverse - Natural Encoding - Adjustable - Intrinsic Difficulty - Tricky - Bounds

Interesting

- **Why?** Relevance outside of planning
- **Qualitative:** Use cases are presented
  - for other research fields
  - for industry
  - ...
- **Quantitative:** ???

Diverse

Diverse

`logistics`

Challenging - Interesting - **Diverse** - Natural Encoding - Adjustable - Intrinsic Difficulty - Tricky - Bounds

Diverse

`logistics`

`miconic`

Challenging - Interesting - **Diverse** - Natural Encoding - Adjustable - Intrinsic Difficulty - Tricky - Bounds

Diverse

logistics

miconic

ferry

Challenging - Interesting - **Diverse** - Natural Encoding - Adjustable - Intrinsic Difficulty - Tricky - Bounds

Diverse

- **Why?** To cover different aspects
- **Qualitative:** Discuss differences to similar existing domains
- **Quantitative:** ???

Natural Encoding

Natural Encoding

```
:action push-to-goal
:action push-to-nongoal
sokoban
```

Natural Encoding

```
:action push-to-goal
:action push-to-nongoal
sokoban
```

Natural Encoding

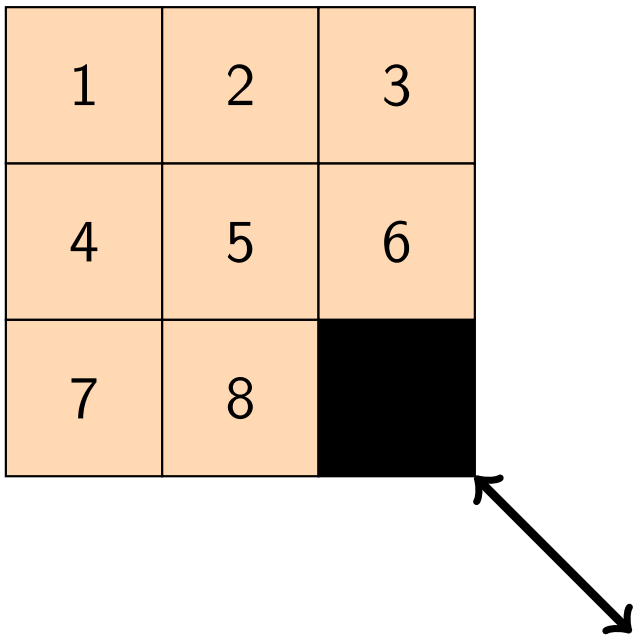
```
:action push-to-goal
:action push-to-nongoal
```
sokoban

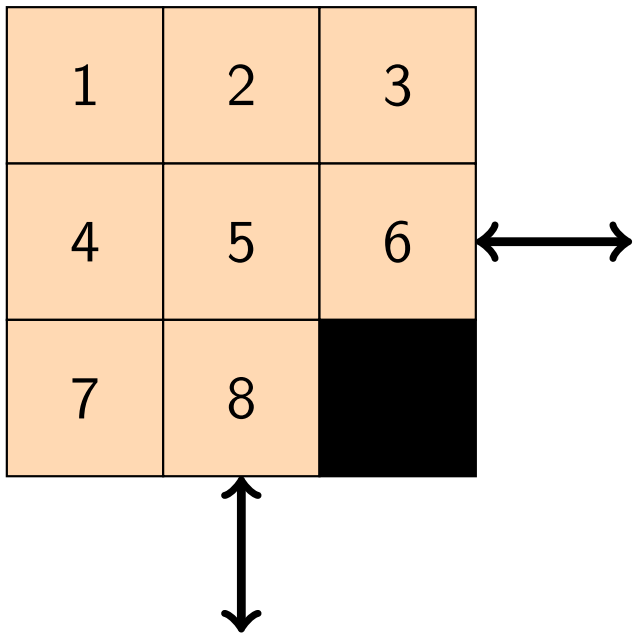
```
:action push
```

Natural Encoding

- **Why?** PDDL closer to the actual problem
- **Qualitative:** Do auxiliary actions/predicates/... exist?
- **Quantitative:** ???

Adjustable





`sliding-tiles`

Challenging - Interesting - Diverse - Natural Encoding - **Adjustable** - Intrinsic Difficulty - Tricky - Bounds

Adjustable

- **Why?** Variance of difficulty within one domain
- **Qualitative:** Is incrementally harder for planners with different generator parameters
  - no sudden phase-transition from trivially easy to too hard
- **Quantitative:** Different planning times for different instances
  - evenly spaced

Intrinsic Difficulty

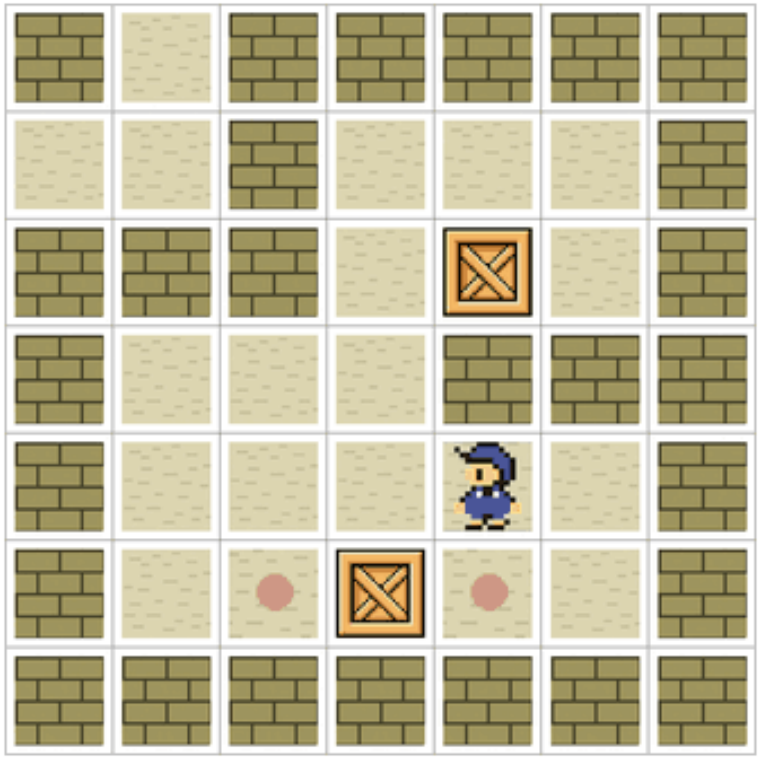
Intrinsic Difficulty

`childsnack`

Challenging - Interesting - Diverse - Natural Encoding - Adjustable - **Intrinsic Difficulty** - Tricky - Bounds

Intrinsic Difficulty

childsnack



sokoban

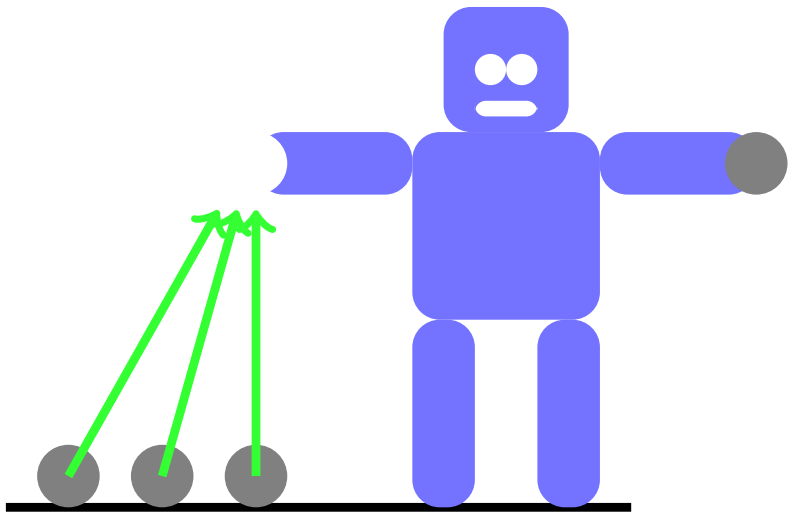
Challenging - Interesting - Diverse - Natural Encoding - Adjustable - **Intrinsic Difficulty** - Tricky - Bounds

Intrinsic Difficulty

- **Why?** Shows capability of planners
- **Qualitative:** Complexity class
- **Quantitative:** $P \subseteq NP \subseteq PH \subseteq PSPACE$

Tricky

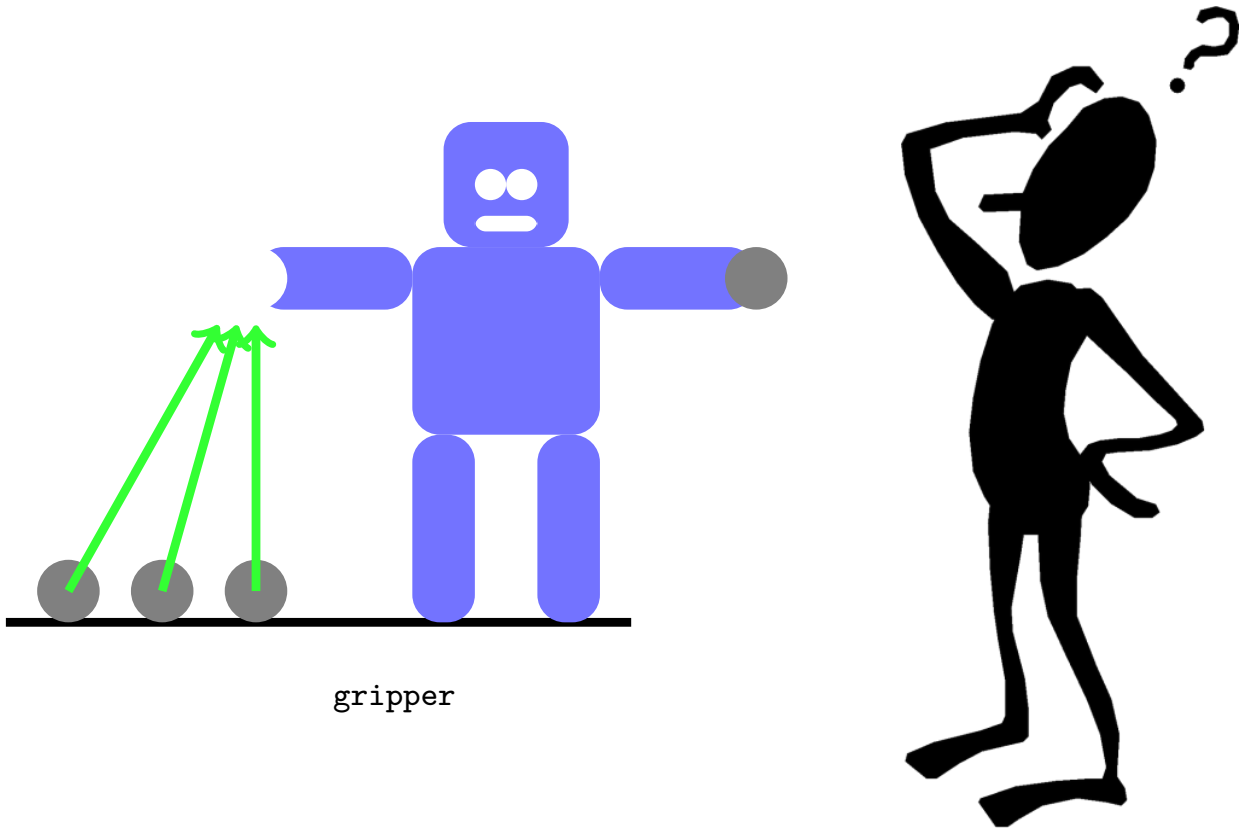
Tricky



`gripper`

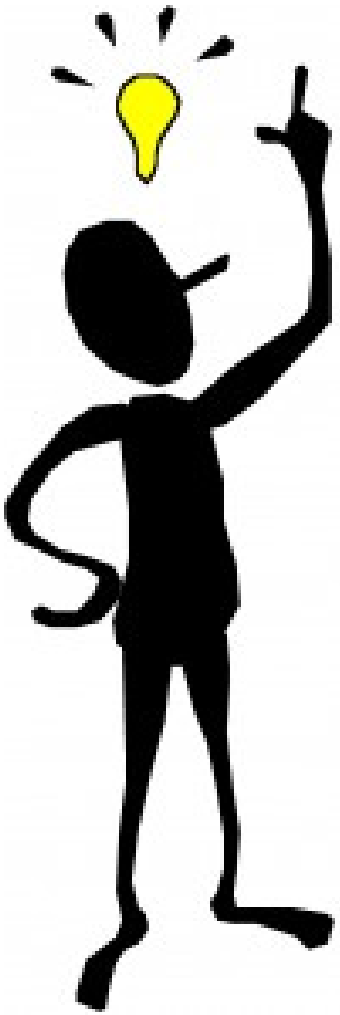
Challenging - Interesting - Diverse - Natural Encoding - Adjustable - Intrinsic Difficulty - **Tricky** - Bounds

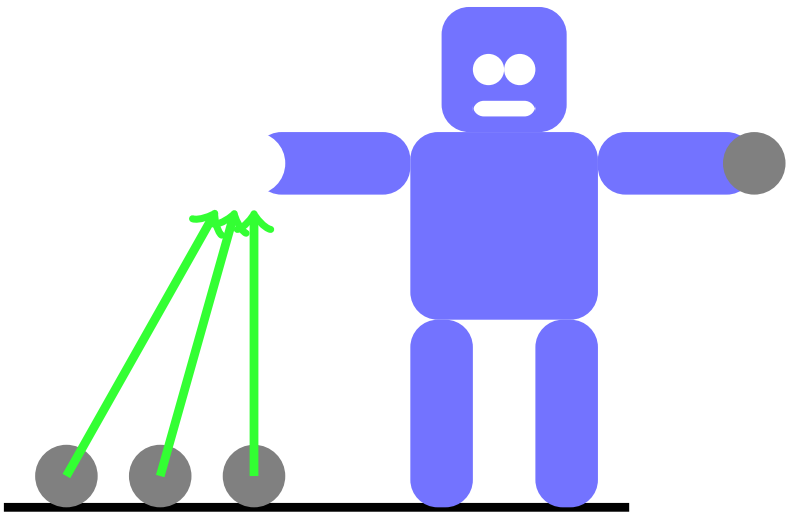
Tricky



Challenging - Interesting - Diverse - Natural Encoding - Adjustable - Intrinsic Difficulty - **Tricky** - Bounds

Tricky





gripper

Challenging - Interesting - Diverse - Natural Encoding - Adjustable - Intrinsic Difficulty - **Tricky** - Bounds

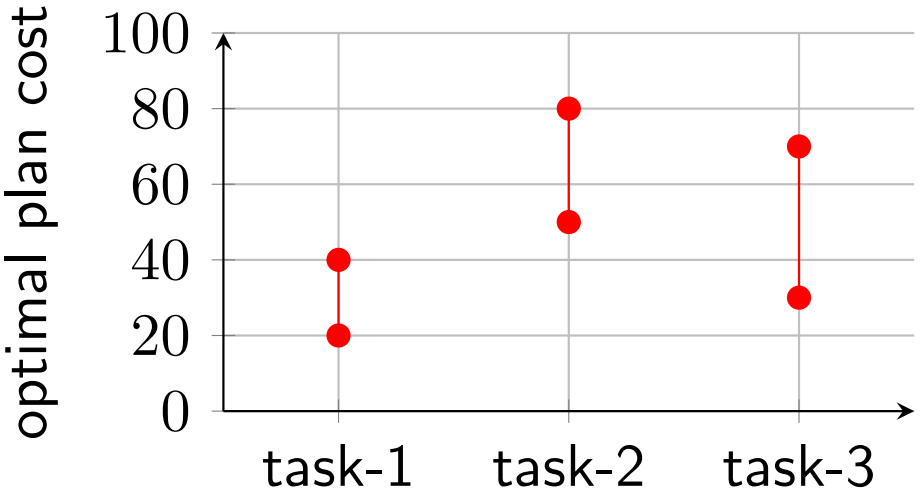
Tricky

- **Why?** Reveal shortcomings to overcome them in the future
- **Qualitative:** Separates planner in those that understand the trick and those that do not
- **Quantitative:** ???

Optimality Bounds

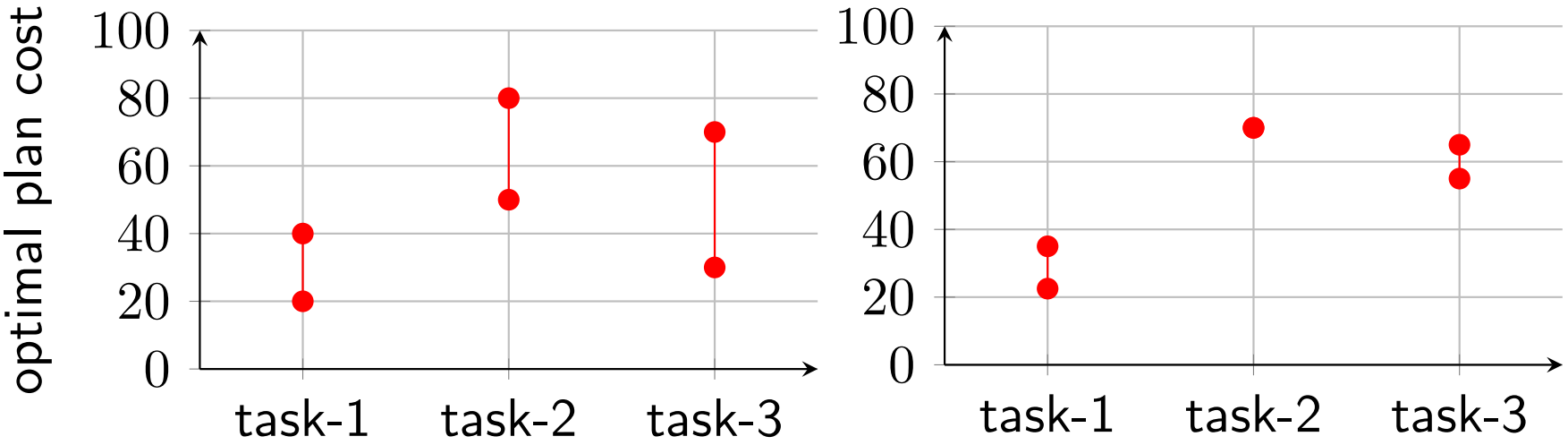
Challenging - Interesting - Diverse - Natural Encoding - Adjustable - Intrinsic Difficulty - Tricky - **Bounds**

Optimality Bounds



Challenging - Interesting - Diverse - Natural Encoding - Adjustable - Intrinsic Difficulty - Tricky - **Bounds**

Optimality Bounds



Challenging - Interesting - Diverse - Natural Encoding - Adjustable - Intrinsic Difficulty - Tricky - **Bounds**

Optimality Bounds

- **Why?** To evaluate optimal planners
- **Qualitative:** A bound exists
- **Quantitative:** The tighter the better

More?

- More properties?

Evaluation Discussion

- Who evaluates the discussed properties?
  - Community voting
  - Peer reviews
  - Committee

More open questions

- How to weigh different properties?
- How many tracks/winners?
- What features are allowed?

More open questions

- How to weigh different properties?
- How many tracks/winners?
- What features are allowed?

Are you interested in helping or participating?

$\rightsquigarrow$ get in contact: `simon.dold@unibas.ch`

Thank You!