

Counterfactual Task Analysis

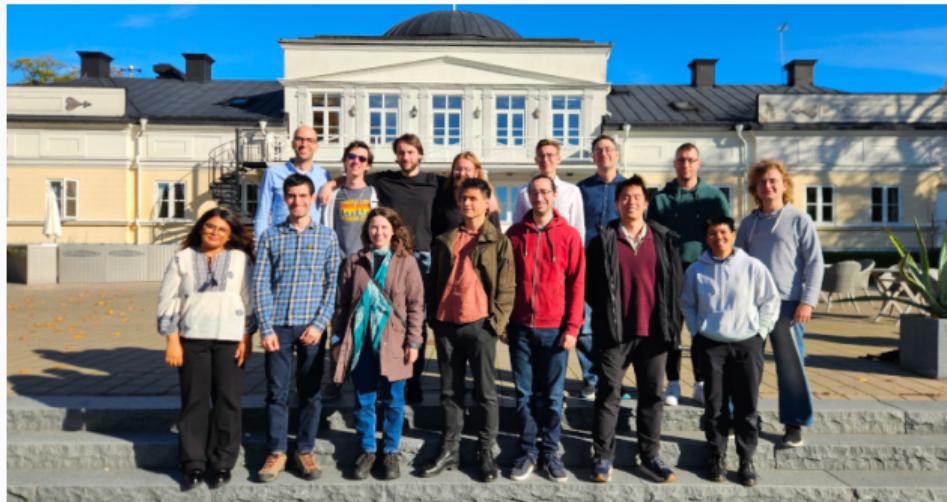
Explainable Planning for the Beluga Challenge and Beyond

Elliot Gestrin¹, Gustav Söderholm¹, Paul Höft¹, Mauricio Salerno², Jendrik Seipp¹, Daniel Gnad¹

¹Linköping University, ²Universidad Carlos III de Madrid

10 November, 2025

Machine Reasoning Lab at Linköping University



Beluga Challenge as a PhD-level course

The Beluga Challenge

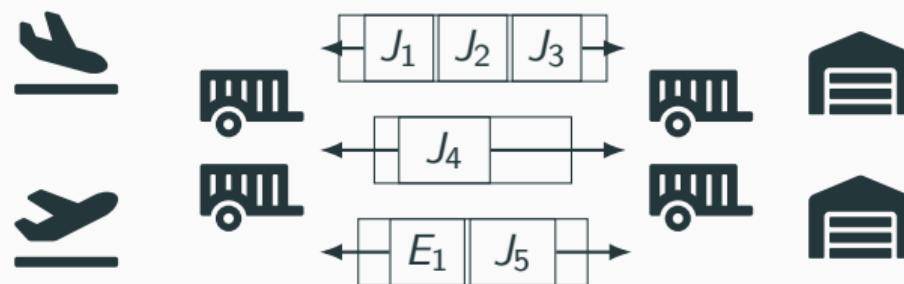


image source: <https://github.com/TUPLES-Trustworthy-AI/Beluga-AI-Challenge>

The Beluga Explainability Challenge

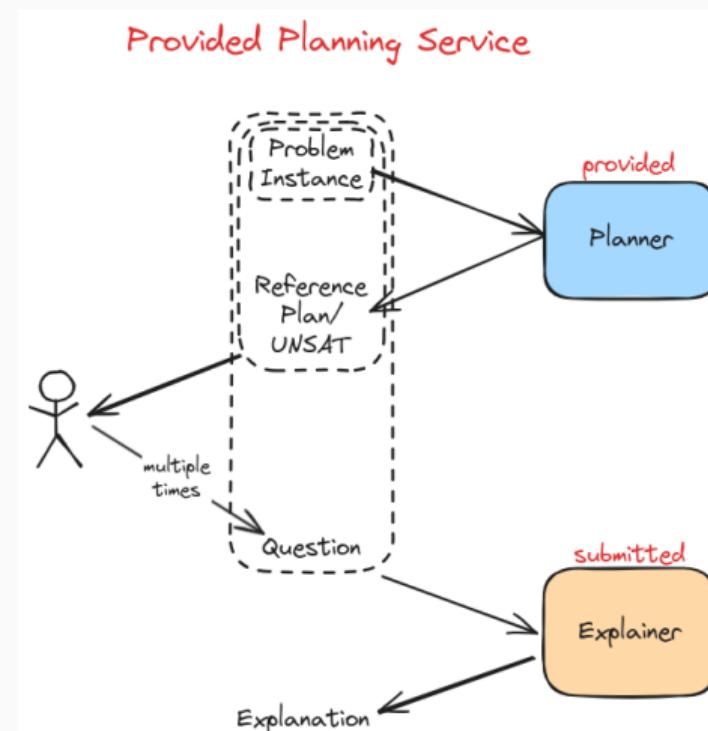


image source: <https://github.com/TUPLES-Trustworthy-AI/Beluga-AI-Challenge/tree/main-competition/explainability>

Explanations from Counterfactual Task Reformulation

Explaining what is using what could have been

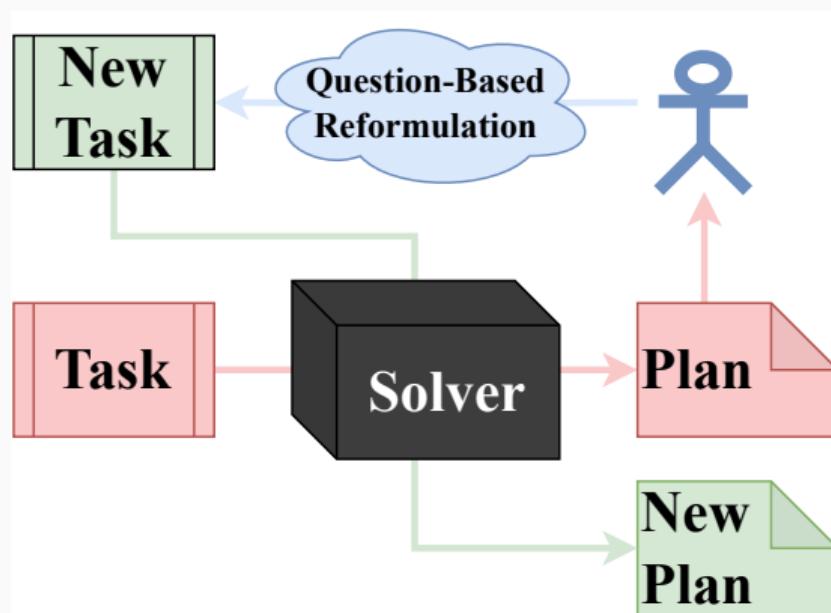
Framework

Counterfactual
Task Reformulation

Solving the
Counterfactual Task

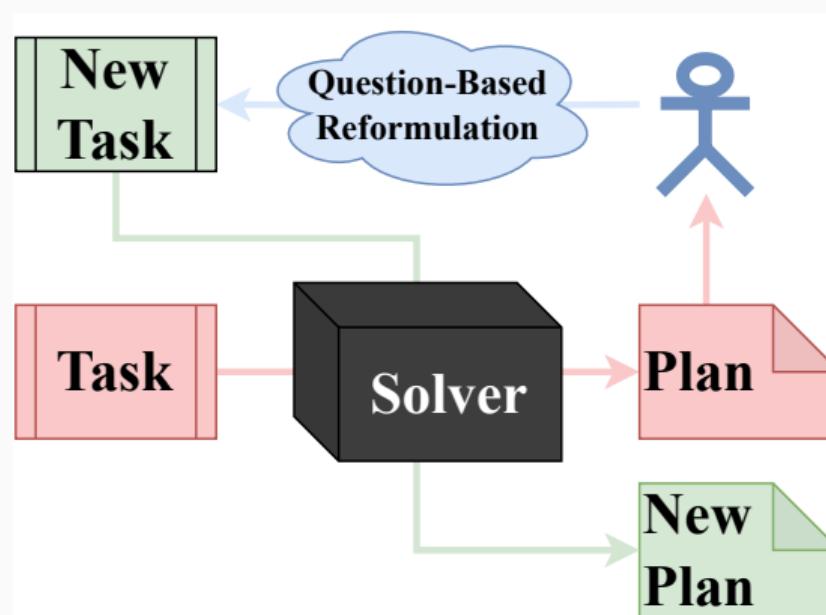
Contrastive LLM
Explanation

Counterfactual Reformulations for Beluga



Counterfactual Reformulations for Beluga

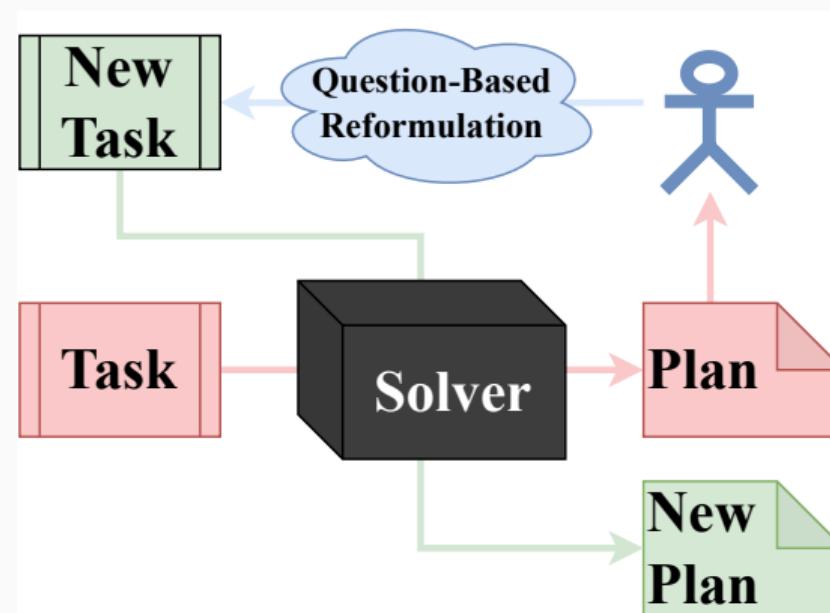
Why is jig x loaded on rack A instead of another rack B?



Counterfactual Reformulations for Beluga

Why is jig x loaded on rack A instead of another rack B?

Forbid placing x in A, enforce placing x in B.



Counterfactual Reformulations for Beluga

Why is jig x loaded on rack A instead of another rack B?

Forbid placing x in A, enforce placing x in B.

Why load jig x on rack A instead of loading jig y on rack B?

Forbid placing x in A, enforce placing y in B.

Why load jig x on rack A before loading jig y on rack B?

Only allow placing x in A if y has been placed in B.

What is the impact of removing rack A for maintenance?

Forbid placing anything in A.

How can I keep one rack empty all the time?

Forbid placing anything on the smallest empty rack.

General Counterfactual Reformulations

Why is jig x loaded on rack A instead of another rack B?

Forbid placing x in A, enforce placing x in B.

Why load jig x on rack A instead of loading jig y on rack B?

Forbid placing x in A, enforce placing y in B.

Why *this* instead of *that*?

Forbid *this*, enforce *that*

General Counterfactual Reformulations

Why load jig x on rack A before loading jig y on rack B?

Only allow placing x in A if y has been placed in B.

Why *this* before *that*?

Enforce *this* and *that*

Make *that* a precondition for *this*

General Counterfactual Reformulations

Can you not use *that*?

Remove interaction with *that*

What is the impact of removing rack A for maintenance?

Forbid placing anything in A.

How can I keep one rack empty all the time?

Forbid placing anything on the smallest empty rack.

Framework

Counterfactual
Task Reformulation

Solving the
Counterfactual Task

Contrastive LLM
Explanation

In competition:

- no control over initial plan quality
- tight time limit

→ LAMA

In competition:

- no control over initial plan quality
- tight time limit

→ LAMA

Post competition analysis on optimal plans:

- LM-Cut in Fast-Downward
- SymK

Solver Comparison

		Reformulated			
		LM-Cut		SymK	
		Solved	Not Solved	Solved	Not Solved
Default	Solved	53	4	84	0
	Not Solved	18	149	16	124

Framework

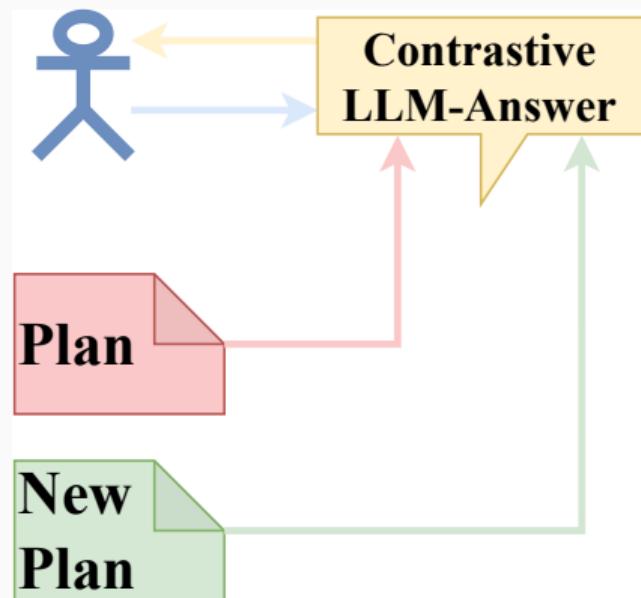
Counterfactual
Task Reformulation

Solving the
Counterfactual Task

Contrastive LLM
Explanation

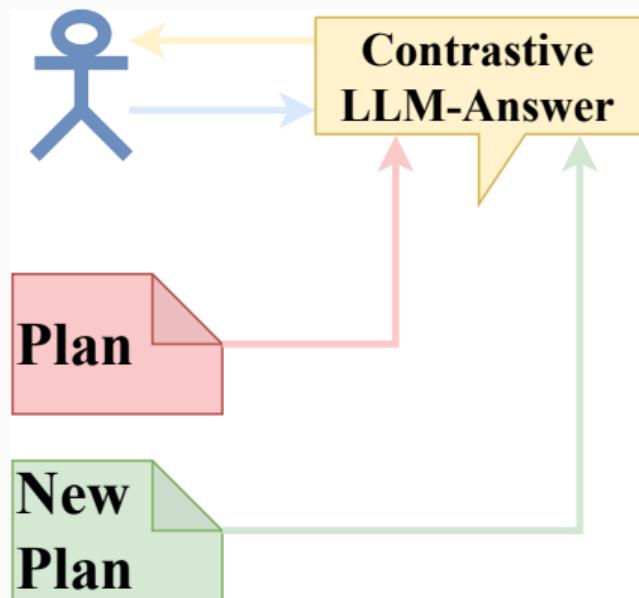
Contrastive Explanations

- Use an LLM



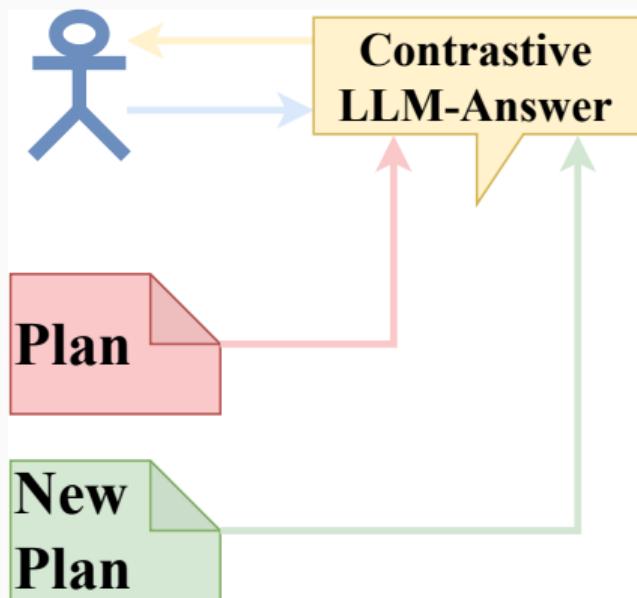
Contrastive Explanations

- Use an LLM
- Prompt:
 - It's an expert



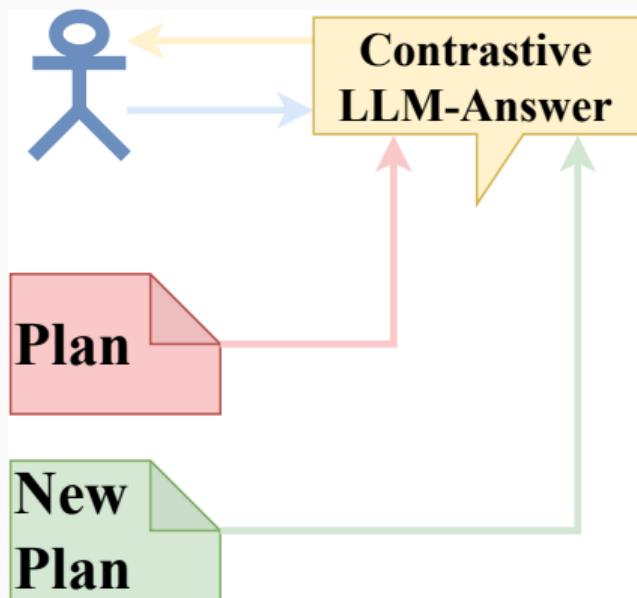
Contrastive Explanations

- Use an LLM
- Prompt:
 - It's an expert
 - Target audience



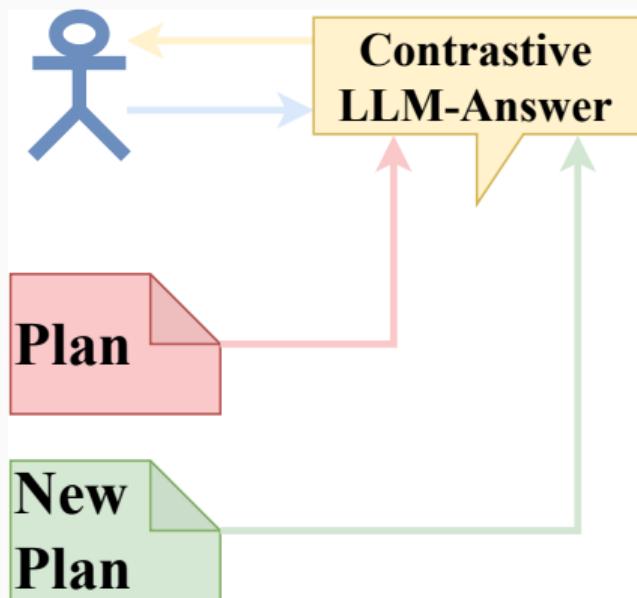
Contrastive Explanations

- Use an LLM
- Prompt:
 - It's an expert
 - Target audience
 - Scoring criteria



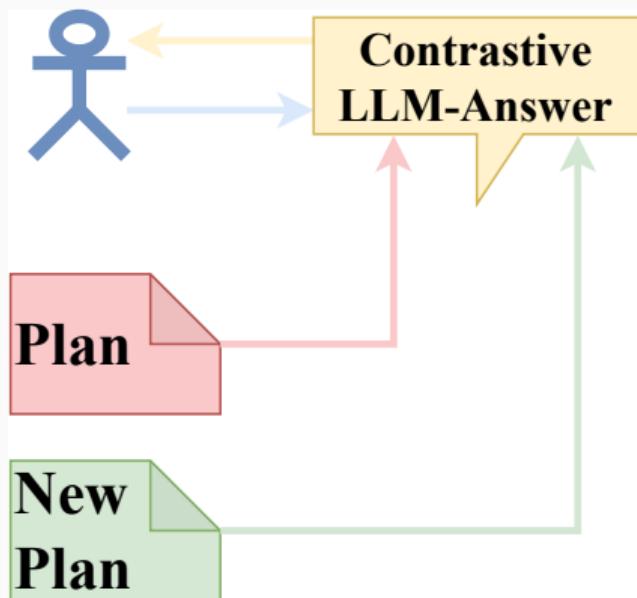
Contrastive Explanations

- Use an LLM
- Prompt:
 - It's an expert
 - Target audience
 - Scoring criteria
 - One-shot



Contrastive Explanations

- Use an LLM
- Prompt:
 - It's an expert
 - Target audience
 - Scoring criteria
 - One-shot
- Model: o3-mini, due to constraints





MR Lab
www.mrlab.ai



Beluga Paper
ICAPS 2025 - HAXP