CSTA++

Unifying State, Action Temporal Abstractions and Causal Information for Contrastive Policy Explanations

Xiaowei Liu[1], Yining Yuan[2], Weiru Liu[2]

August 16, 2026

[1]University of Vienna, [2]University of Bristol

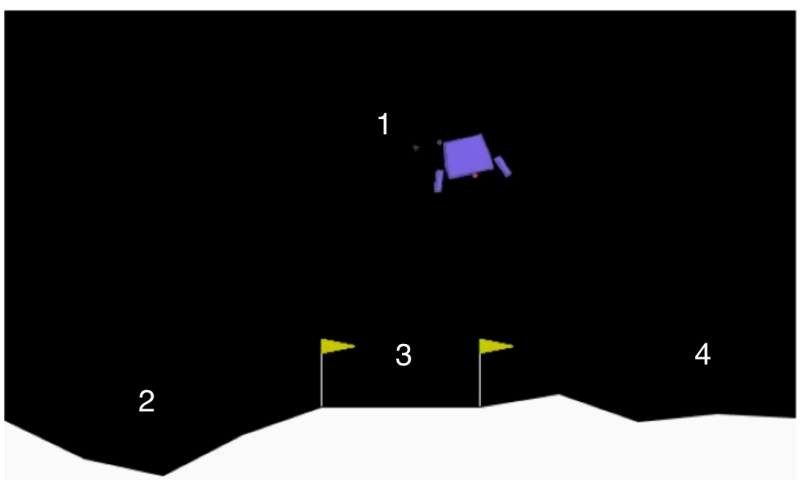
Example: Lunar Lander

State space: $\mathbb{R}^8$. Action space: $\mathbb{R}^2$.

Users want to know:

What if the agent reaches `area 2` rather than `area 3`?

- The `"areas"` are abstractions of states/observations.



Example: Lunar Lander

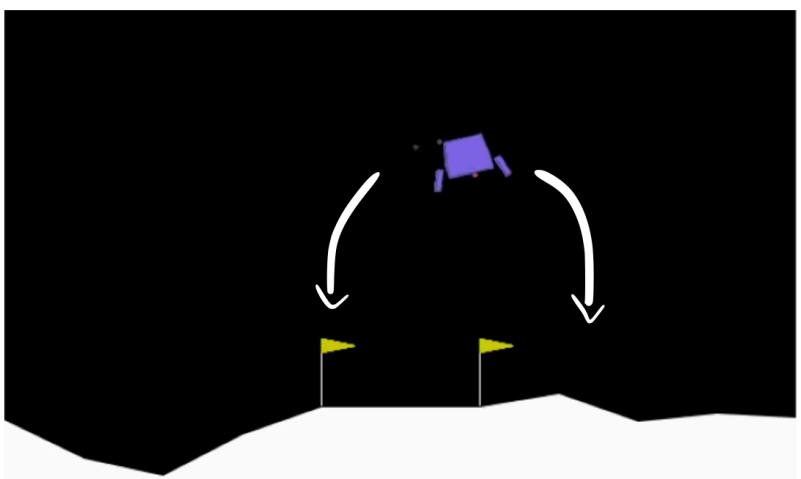
Users want to know:

What if the agent takes **`leftward-descent`** rather than **`rightward-drift`**?

- The descriptions are abstractions of actions (or macro actions, maneuvers).

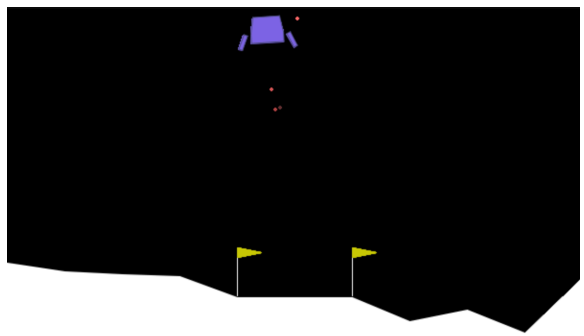
Answers could be, for example:

"This alternative path would have caused time to complete the partition to increase by 22.37 while also causing the reward to increase by 59.76."



Introduction

- **Continuous** state/action spaces (e.g., robotic arms).



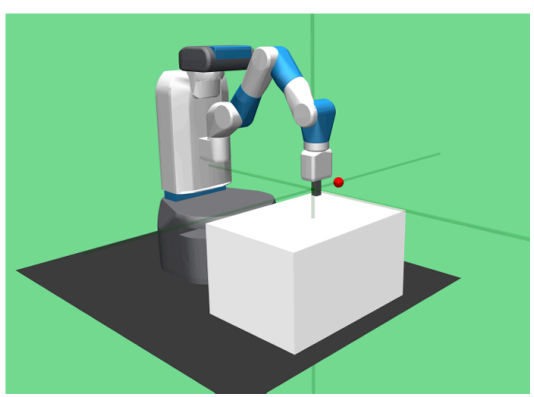
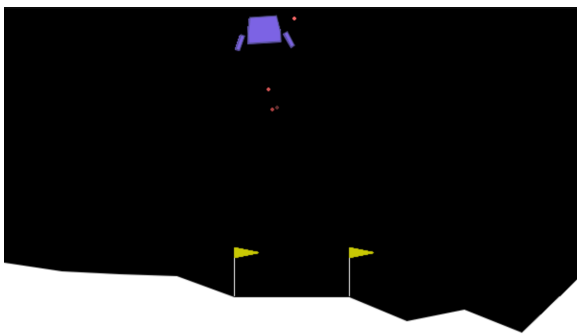


Figure 2: **`LunarLander`** and **`FetchReach`** from Gymnasium.

Introduction

- **Continuous** state/action spaces (e.g., robotic arms).
- With **limited offline trajectories**, learning a model is often not feasible.
- Decisions need **abstraction + temporal structure** together.



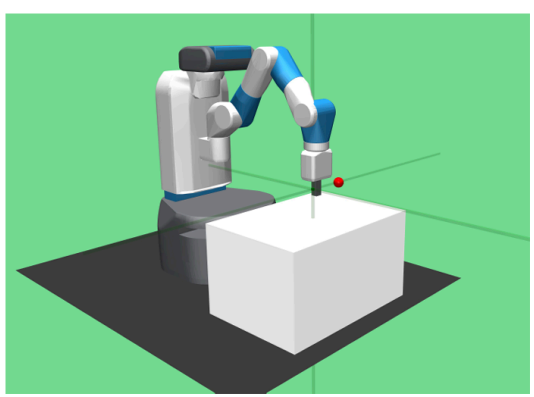
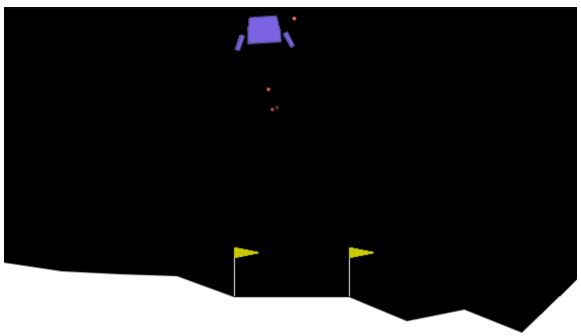


Figure 2: `LunarLander` and `FetchReach` from Gymnasium.

Introduction

- **Continuous** state/action spaces (e.g., robotic arms).
- With **limited offline trajectories**, learning a model is often not feasible.
- Decisions need **abstraction + temporal structure** together.
- Provide **evidence** regarding the decisions (Miller's Evaluative AI proposal.)



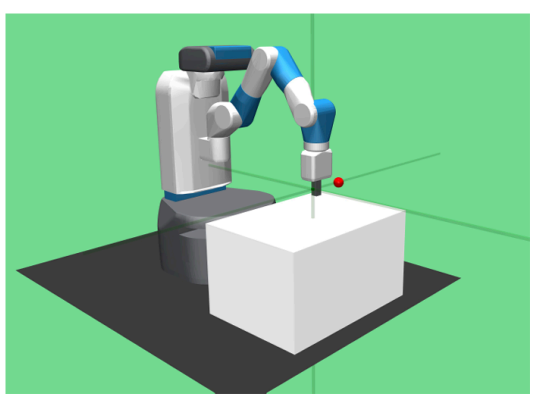
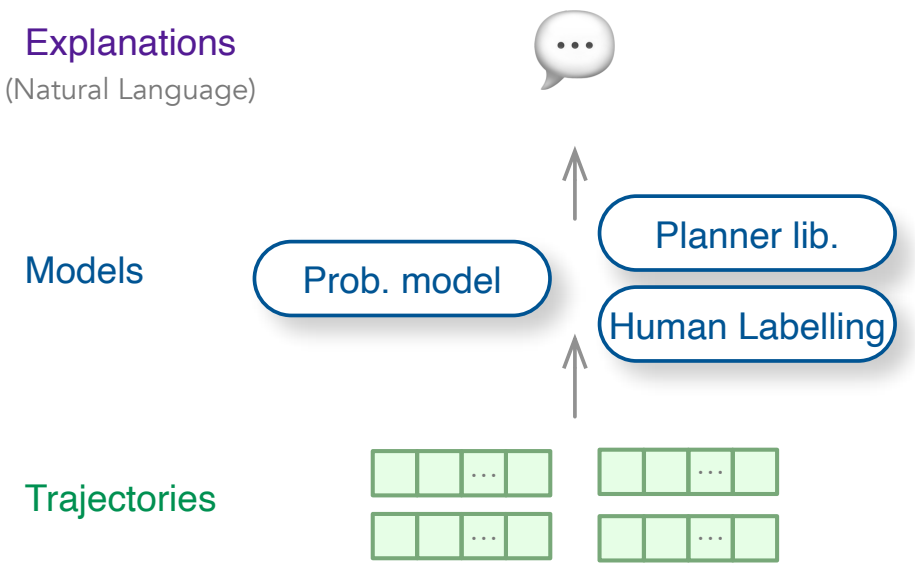


Figure 2: `LunarLander` and `FetchReach` from Gymnasium.

Previous Work: CEMA

CEMA: Causal Explanations in Multi-Agent systems (Gyevnar et al., 2024).

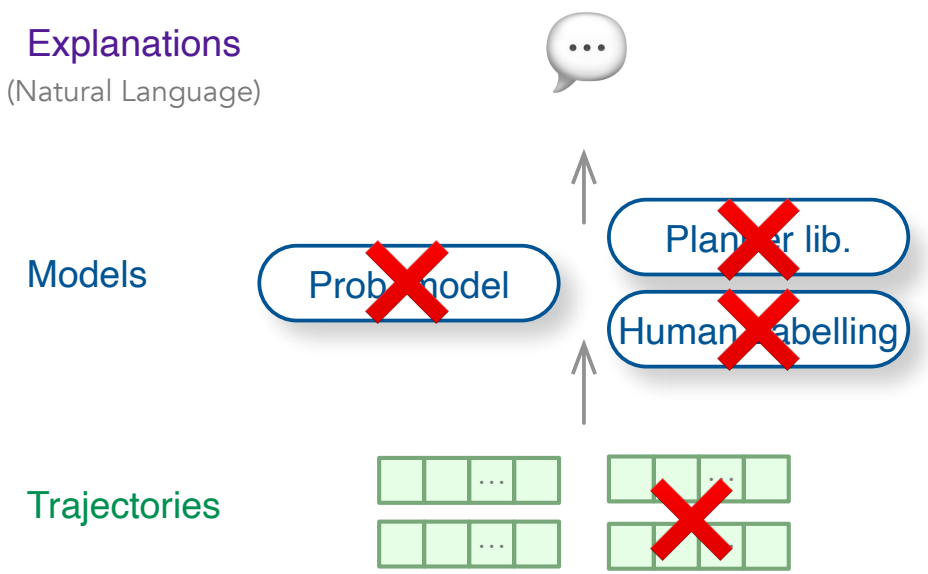
- 👍 Probabilistic model for rollouts.
- 👍 Natural language for counterfactual explanation.
- 😢 Requires a forward probabilistic model.
- 😢 Depends on a planning library.



Previous Work: CEMA

CEMA: Causal Explanations in Multi-Agent systems (Gyevnar et al., 2024).

- 👍 Probabilistic model for rollouts.
- 👍 Natural language for counterfactual explanation.
- 😢 Requires a forward probabilistic model.
- 😢 Depends on a planning library.



We aim to answer:

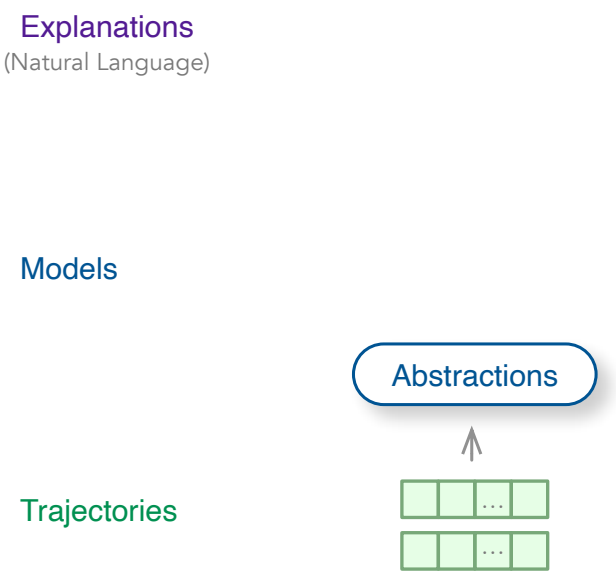
What if agent did (foil) rather than (fact)?

using state and action abstractions from offline trajectory data alone, without explicitly learning a transition model.

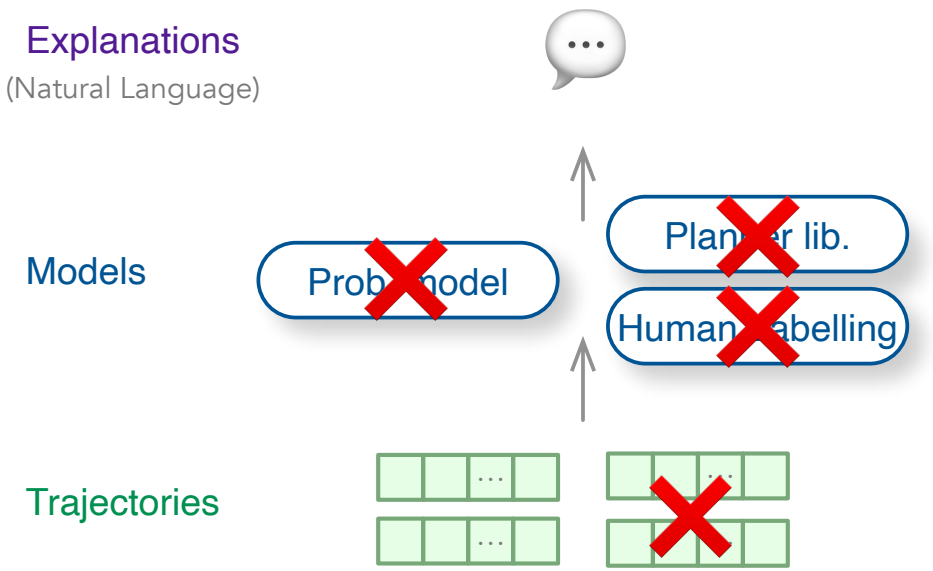
Previous Work: CSTA

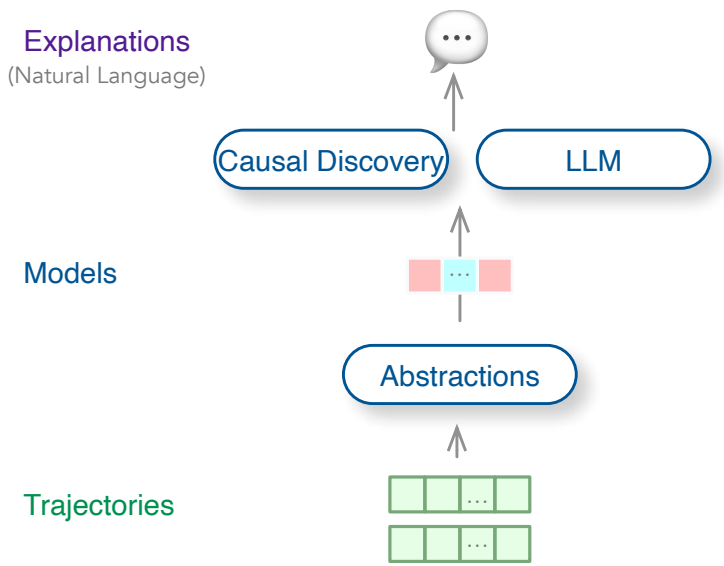
CSTA: Contrastive Spatiotemporal Abstraction (Bewley & Lawry, 2021).

- 👍 Interpretable abstractions of states and rewards with the binary tree.
- 👎 No action abstraction, and no support for interactive queries.

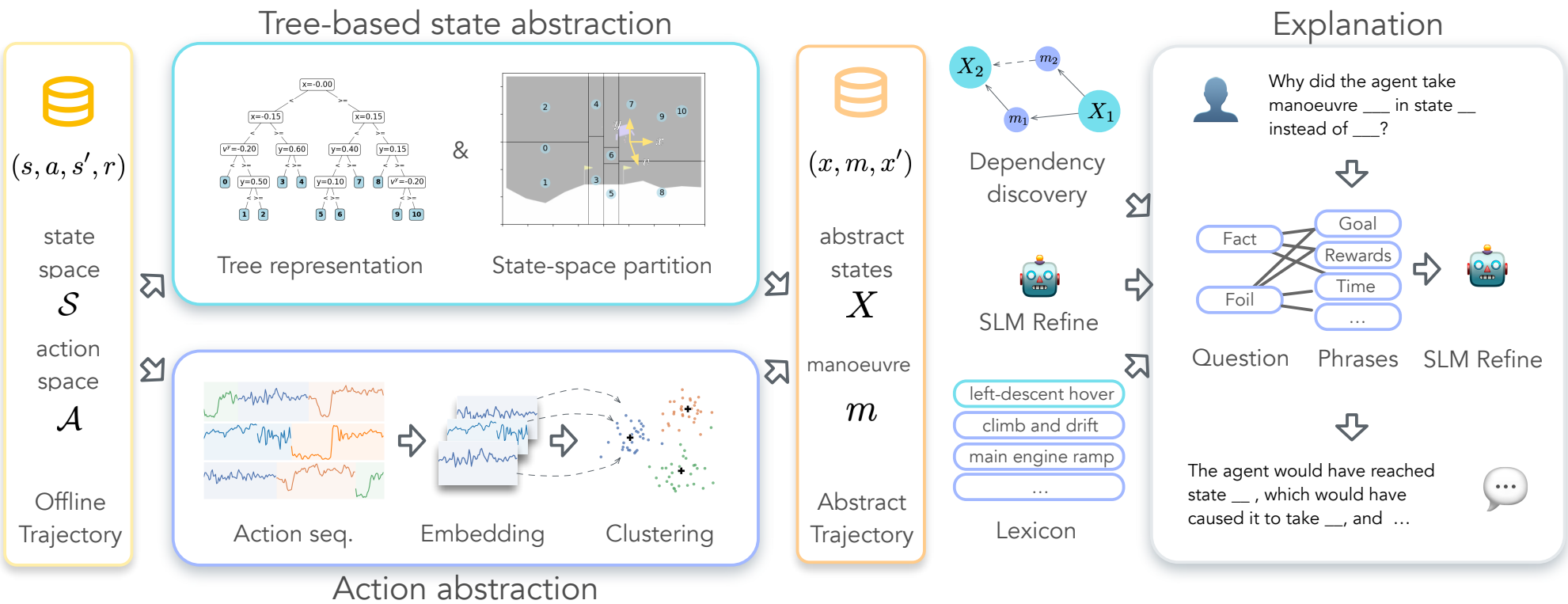


Our Work: CSTA++



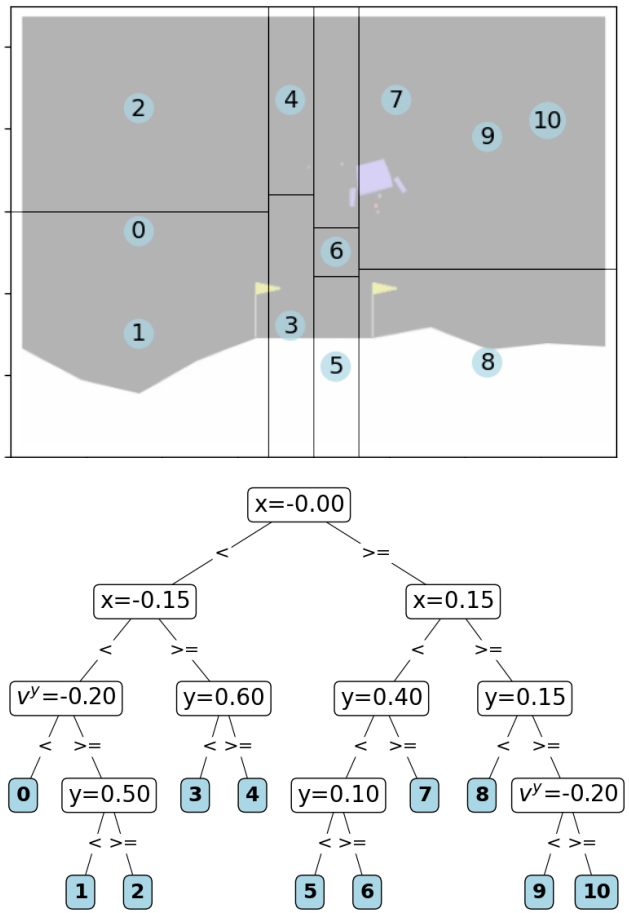


Pipeline of CSTA++



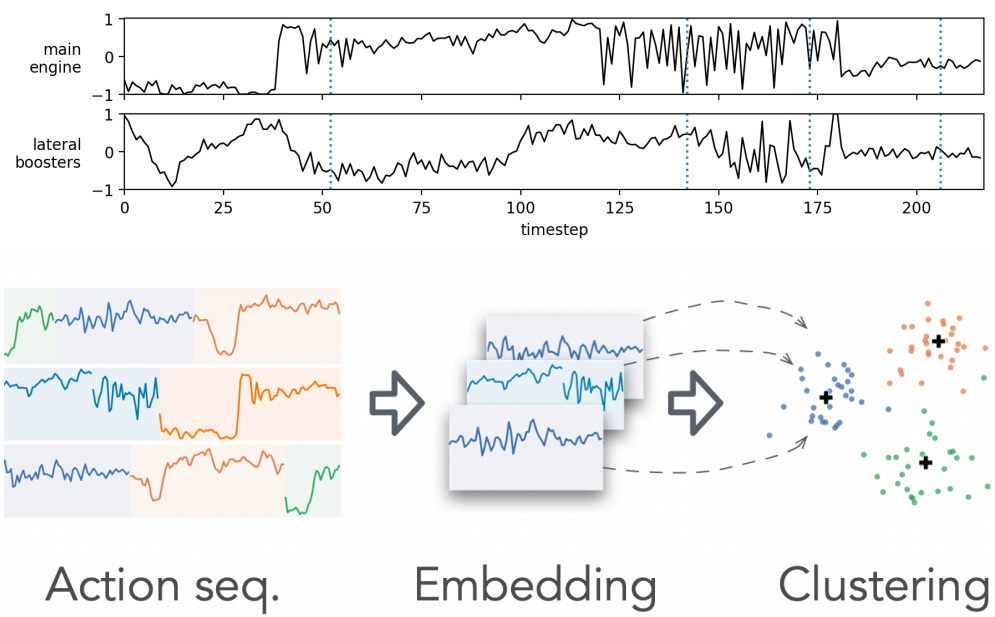
Tree-based State Abstraction

- Maps continuous state space $\mathcal{S} \in \mathbb{R}^D$ onto a finite set of abstract states $\mathcal{X} = \{x_1, x_2, ..., x_m\}$
- Recursively splits the state space with a binary decision tree, maximising the Jensen–Shannon divergence between abstract states.
- Abstract transitions should be as mutually divergent as possible.



Action Abstraction to Manoeuvres

- For each abstract state pair (x_i, x_j), we segment the trajectories using Bayesian changepoint detection.
- Cluster segments across all trajectories with time-series k-means and soft dynamic time warping.
- Name each cluster with a small language model (Lexicon).
- 🚧 Causal dynamics graph between abstract states and manoeuvres.



Lexicon by SLMs

We feed the SLM environment descriptions and action sequences.

It returns `{name, description}` pairs, which are reused in later queries.

Lexicon by SLMs

We feed the SLM environment descriptions and action sequences.

It returns `{name, description}` pairs, which are reused in later queries. For example,

- `manoeuvre 6-5-0, action sequence,...`
- `"left-descent brake"`
- `state 7: x > 0.1, vy ≈ 0, ...`
- `"upper-right hovering"`

Lexicon by SLMs

We feed the SLM environment descriptions and action sequences.

It returns `{name, description}` pairs, which are reused in later queries. For example,

- `manoeuvre 6-5-0, action sequence,...`
- `"left-descent brake"`
- `state 7: x > 0.1, vy ≈ 0, ...`
- `"upper-right hovering"`

Table 1: State lexicon examples

STATES	Lexicon
6	left-centre descent
5	left-centre low-altitude approach
3	left-centre low-altitude hover
1	left low-velocity approach

Table 2: MANOEUVRES lexicon examples

MANOEUVRES	Lexicon
6-5, #0	left-descent brake
6-3, #0	left-descent hover
3-1, #0	leftward-dive initiation
3-5, #0	steady-left-drag

Explanation

Q: What if agent did (foil) rather than (fact)?

- **Mechanistic** explanation (what *causes* it?)
  - 🚧 Causal dynamics graph

Language Template:

- **Mechanistic:** Had the agent taken **a** instead of **b**, it would reached __ instead of __. Reaching __ would caused the agent to __…

Explanation

Q: What if agent did (foil) rather than (fact)?

- **Mechanistic** explanation (what *causes* it?)
  - 🚧 Causal dynamics graph
- **Teleological** explanation (what is it *aiming* for?)
  - Goals, time-to-goal;
  - Change in reward;
  - Estimated final outcome

Language Template:

- **Mechanistic:** Had the agent taken **a** instead of **b**, it would reached __ instead of __. Reaching __ would caused the agent to __…
- **Teleological:** Had the agent taken **a** instead of **b**, it would have caused time to **decrease/increase** by __ and reward to **decrease/increase** by __. This indicates that finally it would __…

Explanation

Q: What if agent did (foil) rather than (fact)?

- **Mechanistic** explanation (what *causes* it?)
  - 🚧 Causal dynamics graph
- **Teleological** explanation (what is it *aiming* for?)
  - Goals, time-to-goal;
  - Change in reward;
  - Estimated final outcome
- **Supportive** explanation (what else *happens*?)

Language Template:

- **Mechanistic:** Had the agent taken **a** instead of **b**, it would reached __ instead of __. Reaching __ would caused the agent to __…
- **Teleological:** Had the agent taken **a** instead of **b**, it would have caused time to **decrease/increase** by __ and reward to **decrease/increase** by __. This indicates that finally it would __…
- **Supportive: b** involves…, __ means…

Explanation Example

What if take *steady left push* instead of *main engine cut left push*?

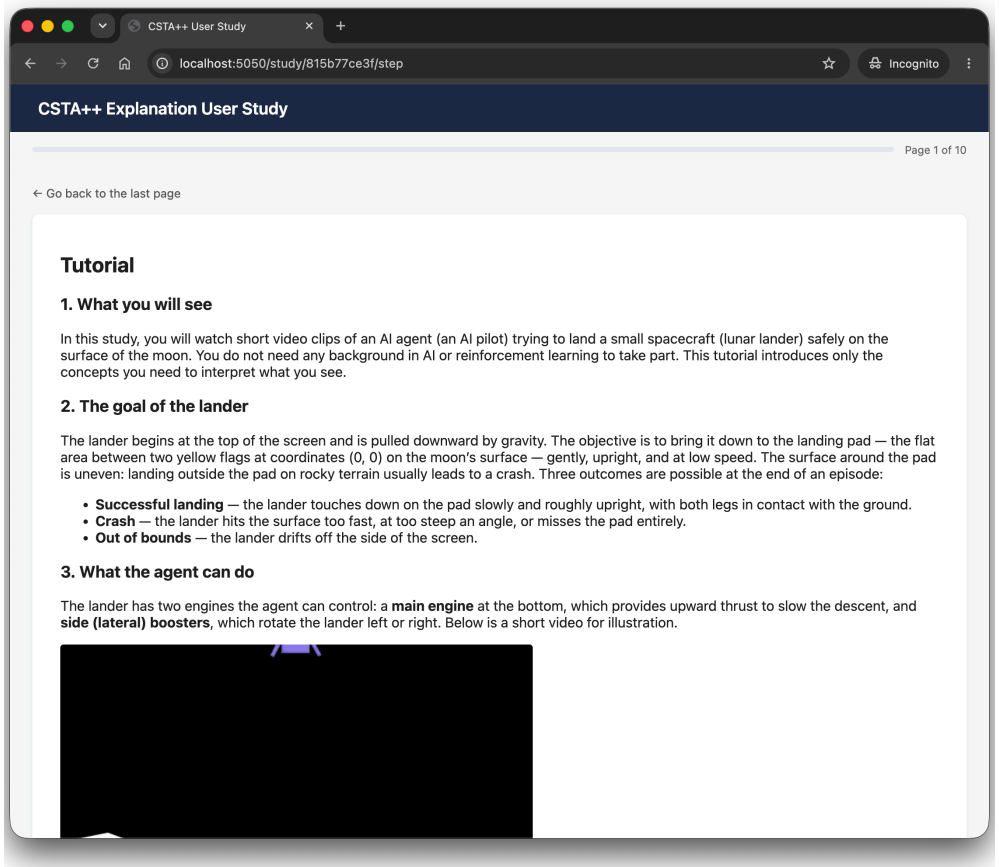
M Had the agent taken *steady left push*, it would have reached *left of pad rising slowly*, instead of *falling above the pad*. *Reaching left of pad rising slowly* would have caused the agent to *reach low above the pad left of center*, instead of *left of pad falling*.

T This alternative path would have caused time to *decrease by 10.00* while also causing the reward to *increase by 15.34*. This indicates that finally it would *land successfully*.

S *Main engine cut left push* involves *a brief leftward push from the side thruster while the main engine throttles down from near full power to off*, *steady left push* involves *a sustained push from the right side thruster to drift the craft leftward, maintained over two timesteps.*

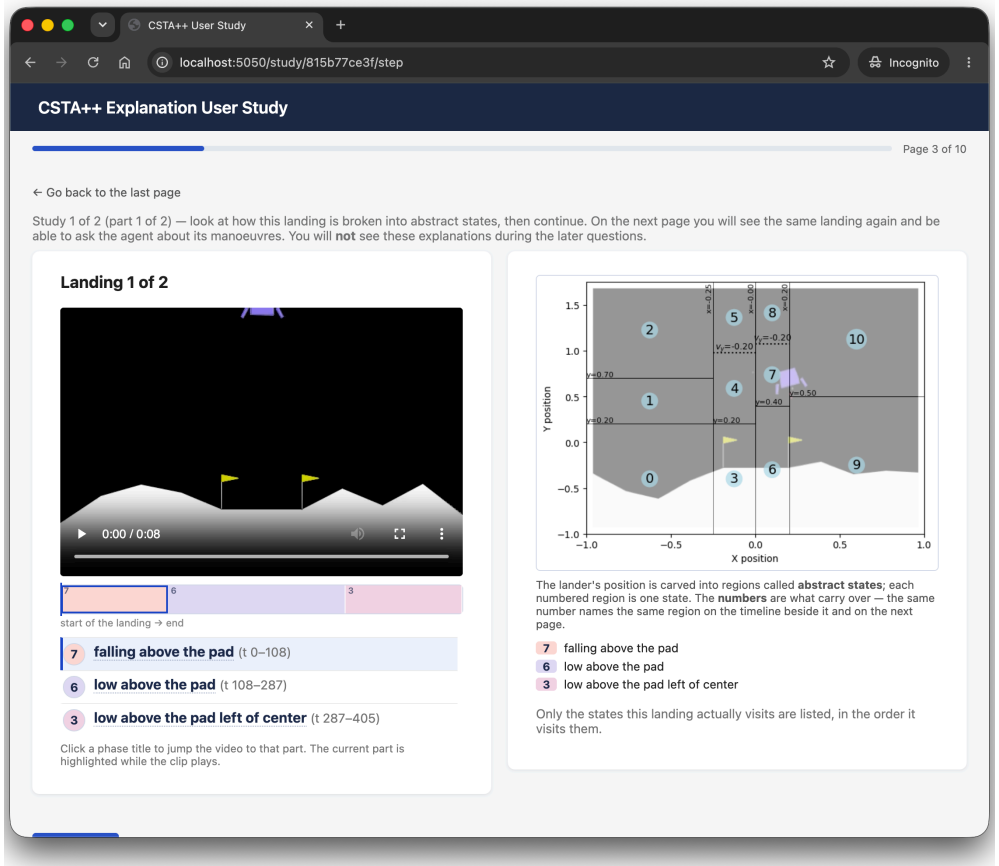
User Study Design and Interface

- Introduction
  - Consent, tutorial and attention check



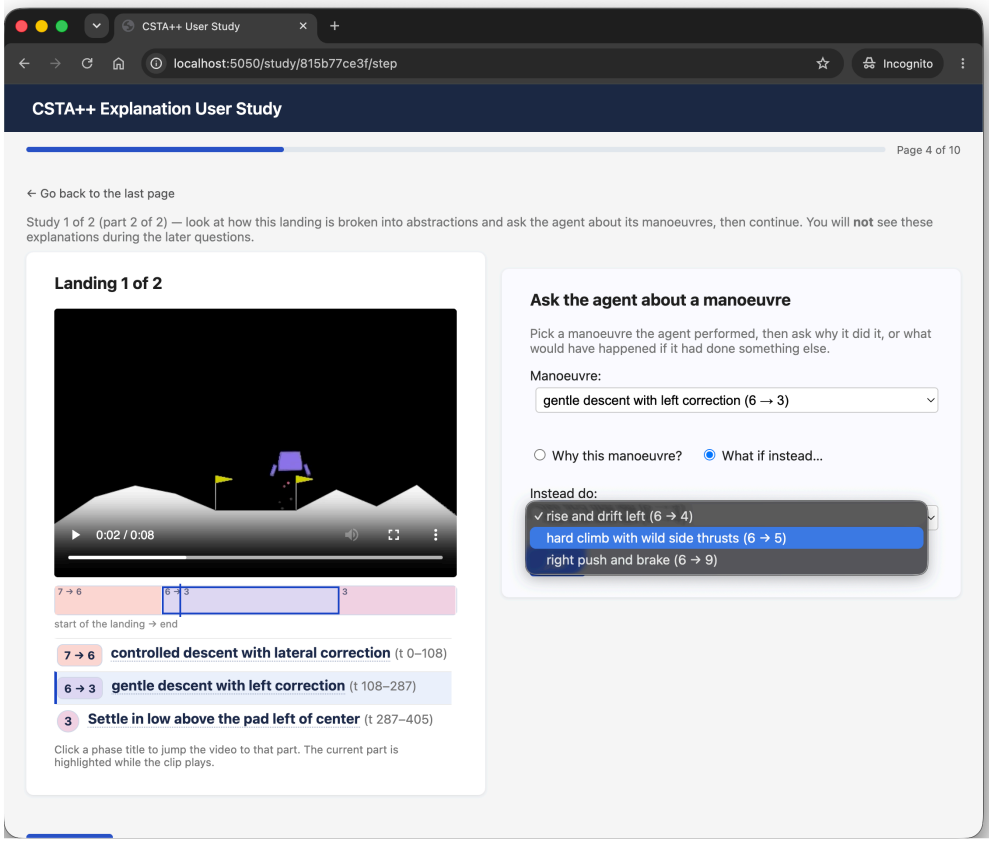
User Study Design and Interface

- **Introduction**
  - Consent, tutorial and attention check
- **Study phase**: two cases *with* explanation
  - Control group: CSTA
  - Experimental group: CSTA++



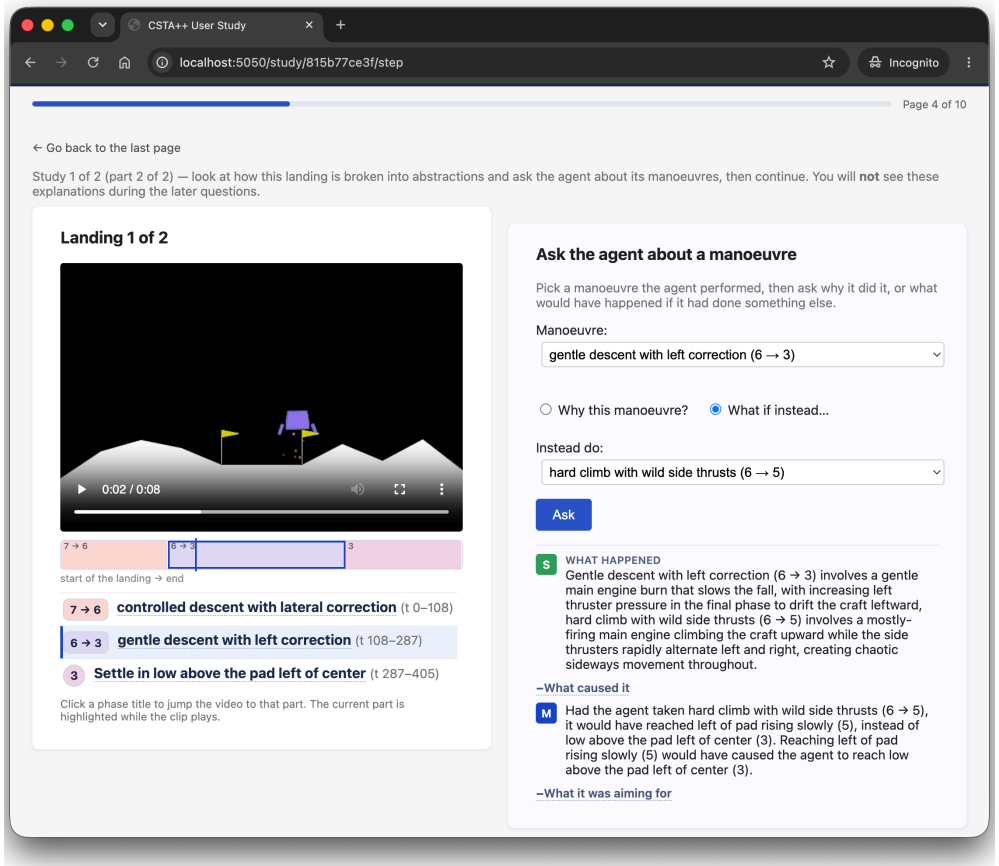
User Study Design and Interface

- **Introduction**
  - Consent, tutorial and attention check
- **Study phase**: two cases *with* explanation
  - Control group: CSTA
  - Experimental group: CSTA++



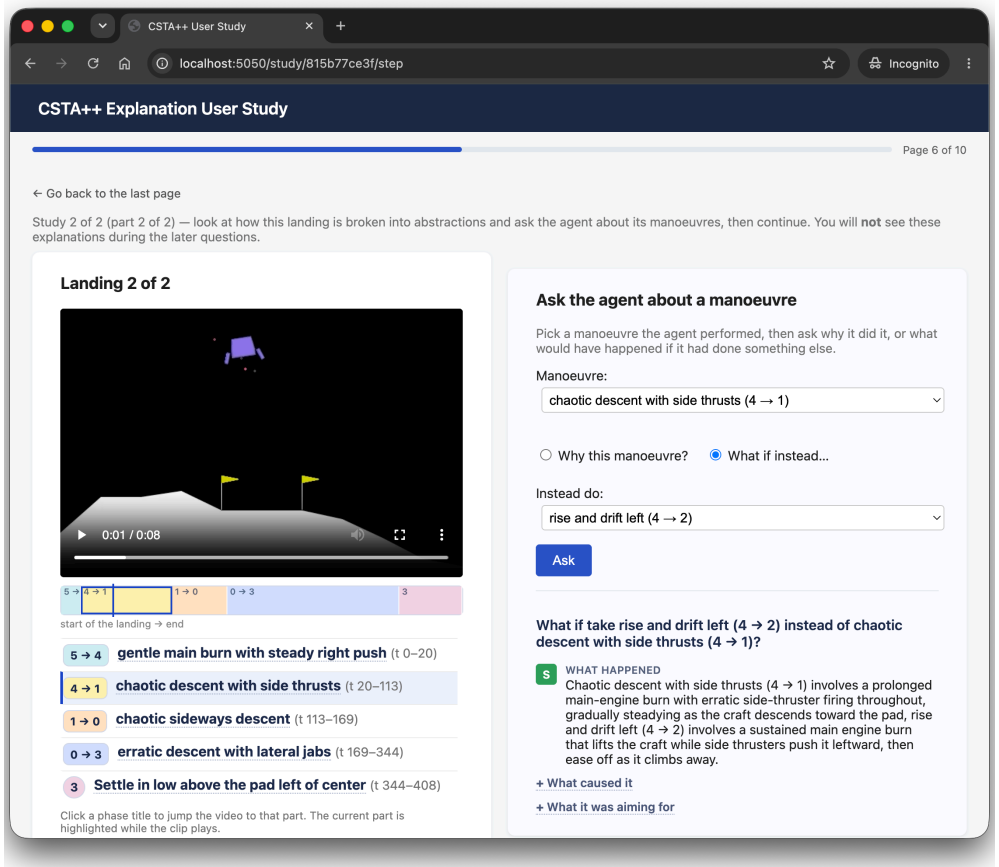
User Study Design and Interface

- **Introduction**
  - Consent, tutorial and attention check
- **Study phase**: two cases *with* explanation
  - Control group: CSTA
  - Experimental group: CSTA++



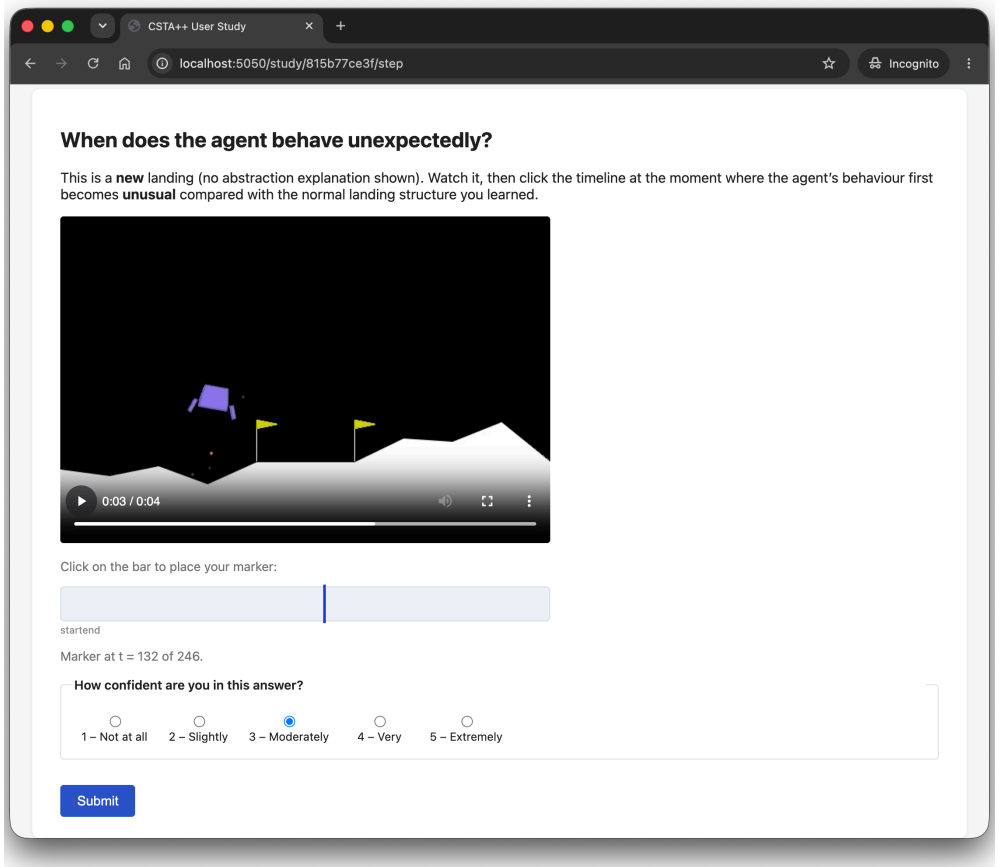
User Study Design and Interface

- Introduction
  - Consent, tutorial and attention check
- **Study phase**: two cases *with* explanation
  - Control group: CSTA
  - Experimental group: CSTA++



User Study Design and Interface

- Introduction
  - Consent, tutorial and attention check
- **Study phase**: two cases *with* explanation
  - Control group: CSTA
  - Experimental group: CSTA++
- **Test phase** *without* explanation
- Questionnaire



Limitations & Future Work

- CSTA++ cannot explain outcomes that never occur in the data.
- Recovered causal structure may not reflect true mechanisms
  - We therefore phrase explanations as “must-do” / “could-do”.
- Benchmark SLMs (9B-20B), not very small ones.
- Shorten the final explanation.
- Extend to goal-conditioned domains (e.g., Pick-and-Place) and robotics applications.

Takeaways and Acknowledgements

- **Setting**: continuous state/action spaces with offline data only.
- **Key ideas**:
  - **Discretise**: the continuous state space.
  - **Manoeuvres**: Bayesian changepoints + time-series clustering.
  - **Lexicon & Refine**: SLM assigns labels into templates.

Paper

Data (WIP)

We thank the anonymous reviewers at XAI@IJCAI'26 and Tom Bewley for comments!

This project was partially funded by the Royal Society UK International Exchange Programme (IES\R2\242045) and the University of Bristol.

Bibliography

Bewley, T., & Lawry, J. (2021). TripleTree: A Versatile Interpretable Representation of Black Box Agents and their Environments. *AAAI 2021*, 11415–11422.

Gyevnar, B., Wang, C., Lucas, C. G., Cohen, S. B., & Albrecht, S. V. (2024). Causal Explanations for Sequential Decision-Making in Multi-Agent Systems. *AAMAS 2024*, 771–779.