Mars: Situated Inductive Reasoning in an Open-World Environment

Xiaojuan Tang Jiaqi Li Yitao Liang Song-chun Zhu Muhan Zhang* Zilong Zheng*
(*Corresponding authors)
Institute for Artificial Intelligence, Peking University National Key Laboratory of General Artificial Intelligence, BIGAI

https://github.com/XiaojuanTang/Mars

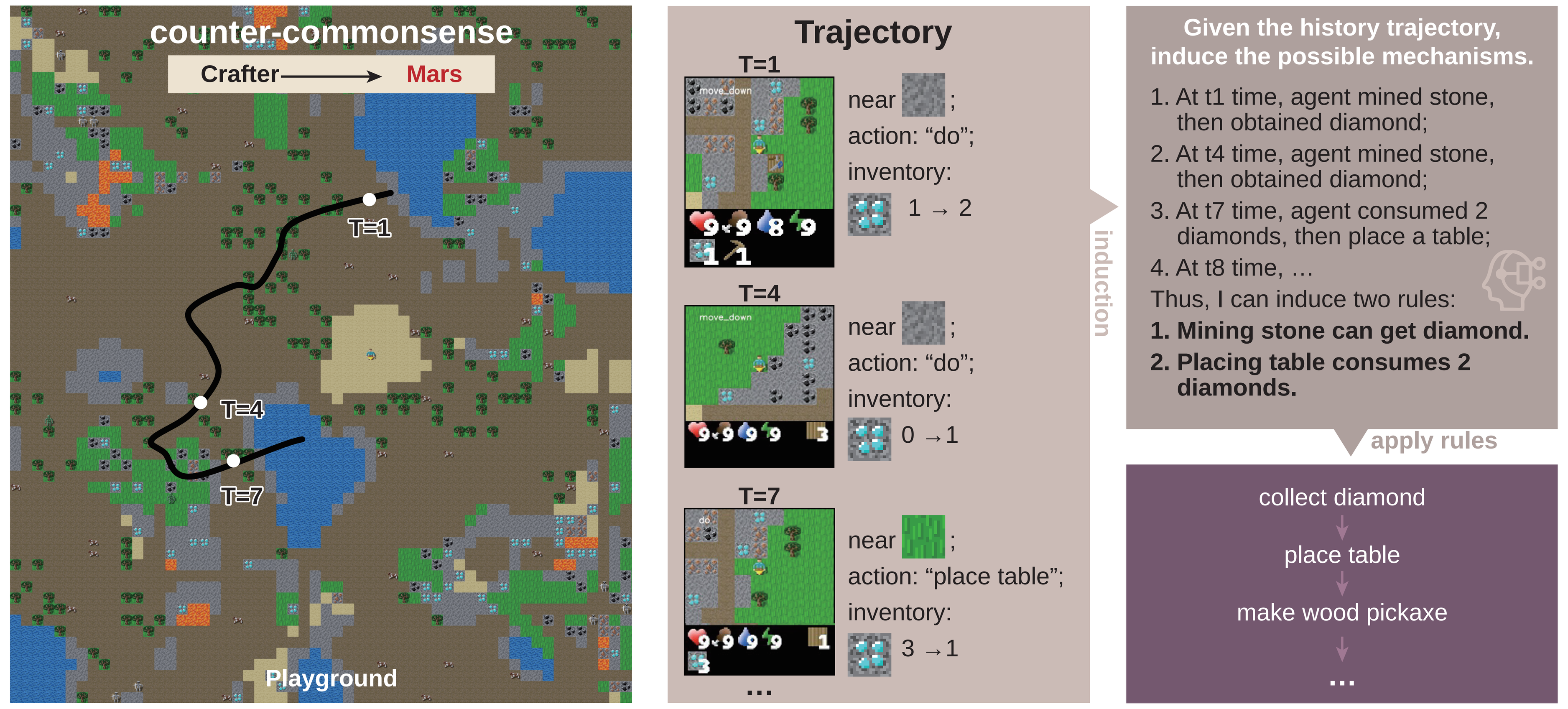
Situated Inductive Reasoning

Image a scenario: in the US, you **drive on the right side of the road**. When you travel to the UK, you might initially find it strange how people drive. However, you soon realize that **driving on the left is the norm here** and adapt yourself to the new rule.

- *Situatedness*: How to summarize and form conclusions from the present, and live observations?
- *Abstractiveness*: how to derive inductive conclusions (i.e., rules or general claims) beyond past experiences?

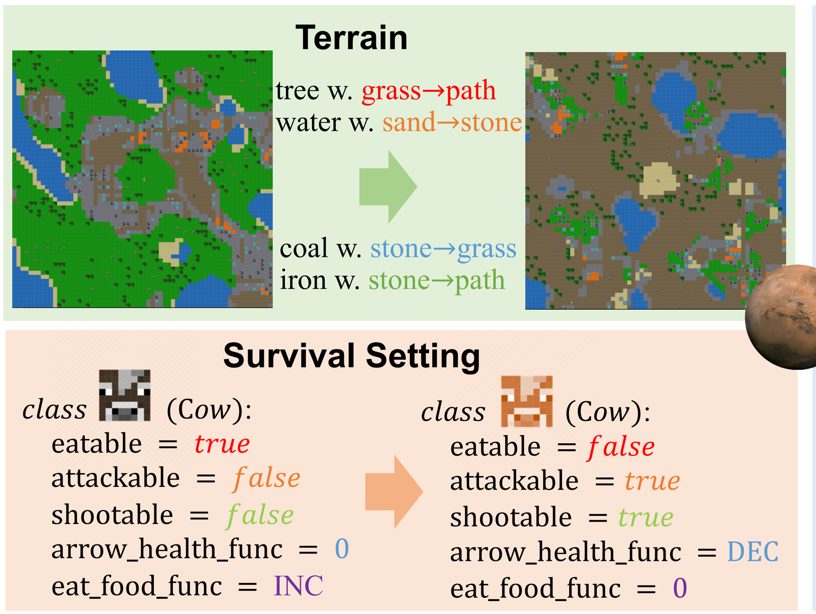
A Novel Benchmark: Mars

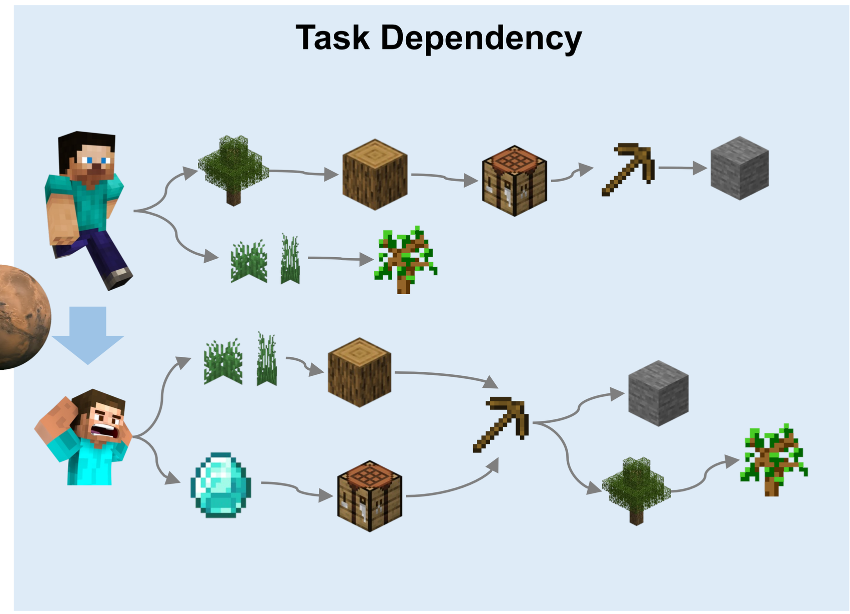
Instruction: In Mars, your goal is to unlock achievements: < collect wood, collect diamond, place table, … >



In Mars, agents are required to quickly derive new general knowledge (rules) from interactions within a specific environment and apply the newly acquired knowledge effectively in a new context.

counter-commonsense modifications





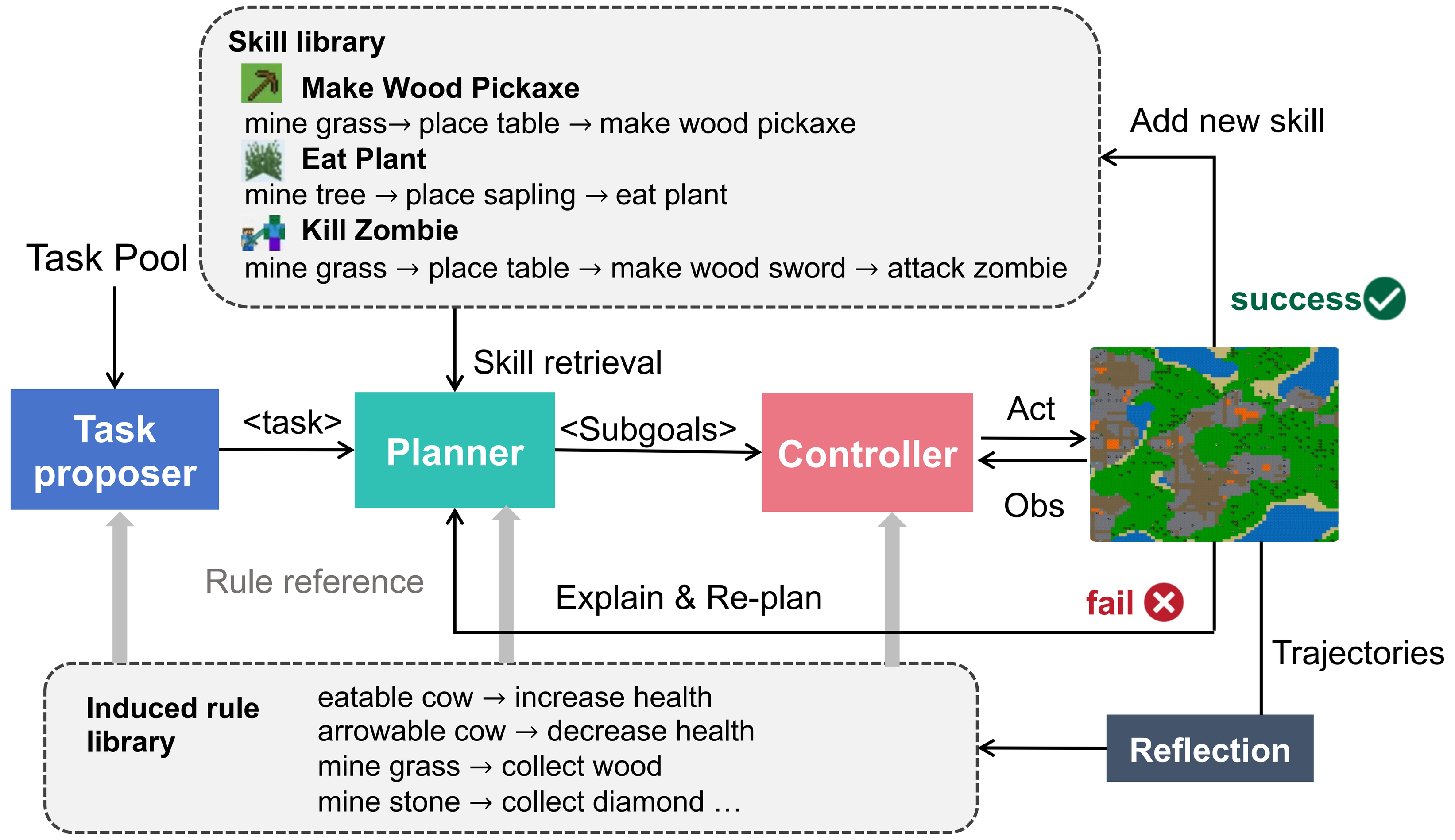
Modified world need to adhere to certain principles:

- resource balance
- each achievement is achievable
- supply exceeds demand
- each collected item is obtainable
- each tool has a practical use
- …

Method: Induction from Reflection (IfR)

Building on the Jarvis-1 framework, we further introduce IfR module:

- When the controller finishes a subgoal, LLMs will **induce possible game mechanisms** based on the agent's historical trajectory.
- The derived rules are then stored in a ***rule library***, which the task proposer, planner, and controller can use.



Experimental Results

Metrics	Mod. Type	RL-based methods PPO	DreamerV3	LLM-based methods ReAct	Reflexion	Skill Library	Ours
Reward	Default	$1.9^{\pm1.4}$	$11.5^{\pm1.6}$	$7.7^{\pm1.6}$	$6.0^{\pm1.7}$	$8.0^{\pm2.1}$	$9.0^{\pm2.3}$
	Terrain	$-0.1^{\pm0.6}$	$9.3^{\pm2.2}$	$7.4^{\pm2.7}$	$6.4^{\pm3.0}$	$9.5^{\pm2.9}$	$8.0^{\pm3.7}$
	Survival	$-0.6^{\pm0.5}$	$8.6^{\pm4.1}$	$6.4^{\pm3.7}$	$4.6^{\pm3.9}$	$7.9^{\pm2.9}$	$7.7^{\pm3.7}$
	Task. Dep	$2.1^{\pm1.2}$	$8.8^{\pm2.8}$	$5.0^{\pm2.1}$	$3.2^{\pm1.6}$	$1.5^{\pm1.9}$	$5.6^{\pm2.9}$
	Terr. Surv.	$0.0^{\pm0.7}$	$7.1^{\pm2.1}$	$6.7^{\pm2.5}$	$4.9^{\pm2.5}$	$3.0^{\pm2.5}$	$6.8^{\pm1.9}$
	Terr. Task.	$-0.7^{\pm0.3}$	$6.6^{\pm0.7}$	$4.8^{\pm2.0}$	$5.3^{\pm2.5}$	$5.5^{\pm1.5}$	$6.9^{\pm1.8}$
	Surv. Task.	$-0.6^{\pm0.4}$	$9.6^{\pm3.4}$	$1.5^{\pm1.3}$	$1.0^{\pm1.6}$	$2.3^{\pm1.5}$	$3.3^{\pm1.4}$
	All three.	$0.1^{\pm0.8}$	$5.1^{\pm1.8}$	$0.7^{\pm1.6}$	$-0.4^{\pm0.7}$	$-0.5^{\pm0.5}$	$0.1^{\pm0.5}$
Score (%)	Default	$1.3^{\pm1.7}$	$14.2^{\pm1.3}$	$8.0^{\pm1.5}$	$5.3^{\pm0.9}$	$8.3^{\pm1.3}$	$13.0^{\pm2.1}$
	Terrain	$0.3^{\pm0.1}$	$13.0^{\pm1.6}$	$7.6^{\pm2.6}$	$7.4^{\pm1.6}$	$11.9^{\pm3.4}$	$11.8^{\pm2.9}$
	Survival	$0.2^{\pm0.0}$	$10.8^{\pm2.8}$	$8.0^{\pm0.6}$	$5.5^{\pm1.7}$	$9.7^{\pm2.0}$	$11.0^{\pm3.7}$
	Task. Dep	$1.7^{\pm0.6}$	$12.1^{\pm1.9}$	$4.6^{\pm1.6}$	$2.2^{\pm0.8}$	$1.5^{\pm0.6}$	$6.9^{\pm2.5}$
	Terr. Surv.	$0.4^{\pm0.1}$	$7.9^{\pm1.3}$	$7.1^{\pm3.0}$	$4.7^{\pm1.6}$	$2.8^{\pm0.6}$	$6.7^{\pm0.8}$
	Terr. Task.	$0.1^{\pm0.1}$	$4.2^{\pm0.1}$	$3.8^{\pm0.3}$	$5.5^{\pm1.7}$	$4.1^{\pm0.7}$	$7.1^{\pm2.5}$
	Surv. Task	$0.1^{\pm0.1}$	$15.9^{\pm2.6}$	$1.3^{\pm0.2}$	$1.1^{\pm0.1}$	$1.9^{\pm0.1}$	$2.1^{\pm0.4}$
	All three.	$0.6^{\pm0.2}$	$4.0^{\pm0.3}$	$1.0^{\pm0.3}$	$0.2^{\pm0.1}$	$0.2^{\pm0.0}$	$0.6^{\pm0.0}$

All baseline models exhibit **a performance decline** when transitioning from the Default to Mars scenarios, underscoring that Mars has significant challenges for current methodologies.