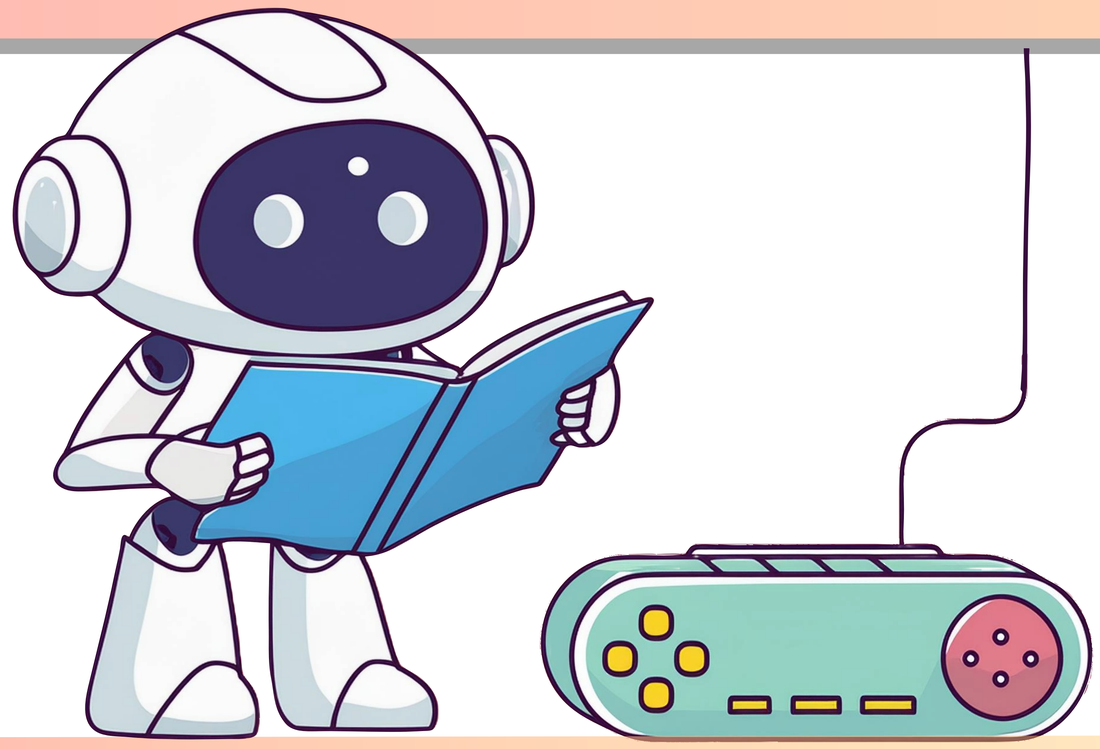
OCALM: Object-Centric Assessment with Language Models

Timo Kaufmann*,[1,2,3] Jannis Blüml*,[4,5] Antonia Wüst*,[4]
Quentin Delfosse*,[4,6] Kristian Kersting[4,5,7] Eyke Hüllermeier[1,2]

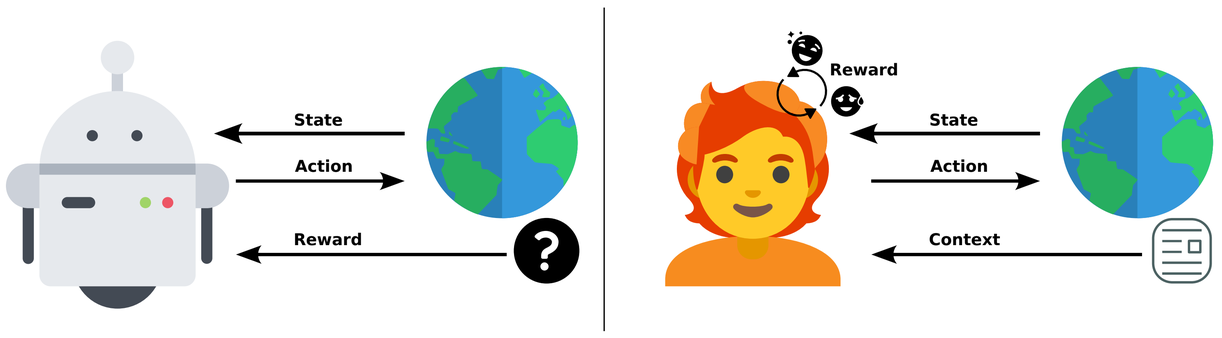
Reinforcement Learning Conference

Relational concepts improve **LLM-generated** interpretable **rewards** from task descriptions.

SCAN ME

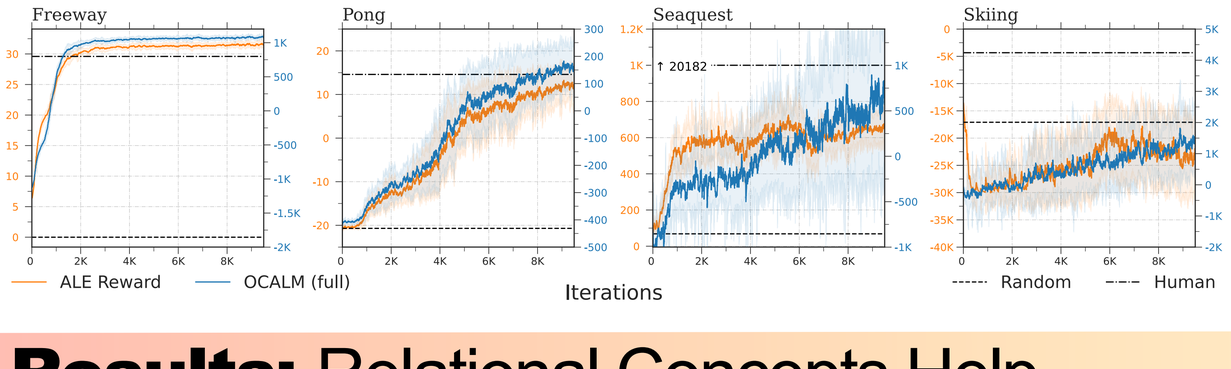


Goal: Context-Based Rewards

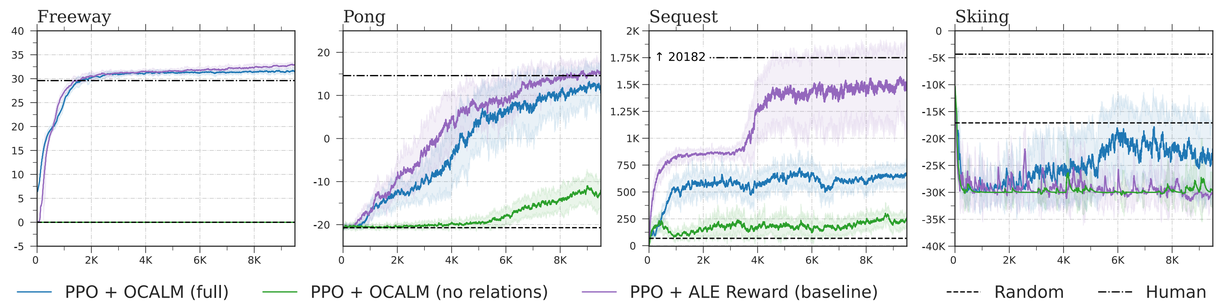


Idea: Contrary to humans, RL agents struggle to derive tasks' objectives from a contextual description.
Aim: Self-generated effective reward signals for RL agents.
OCALM: Deriving reward functions from natural language task descriptions using LLMs and object-centric reasoning.

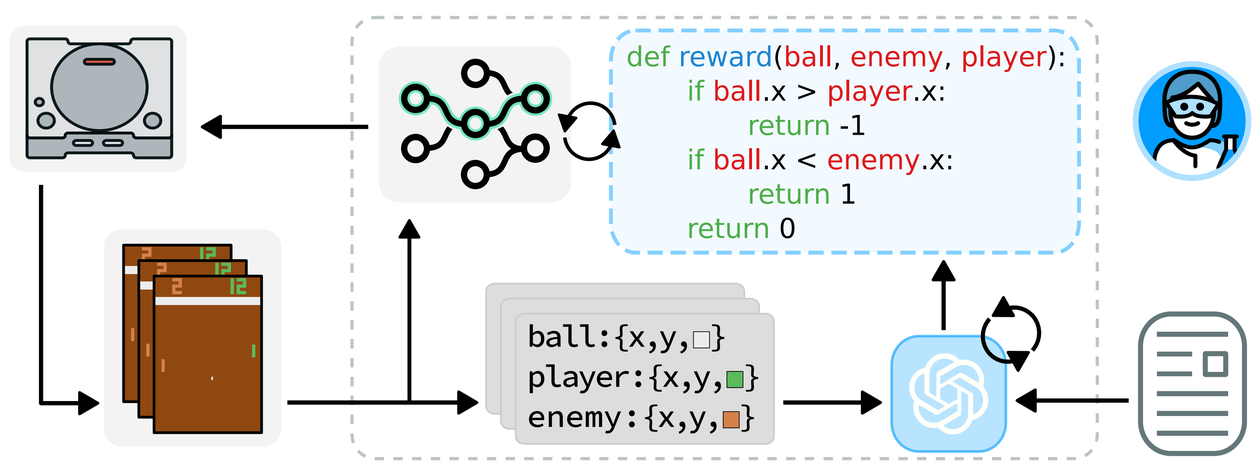
Results: LLMs Generate Useful Rewards



Results: Relational Concepts Help



OCALM: Rewards Based on Task Description



I. Context and Objects:
(i) Using Context: Extract task descriptions from natural language representation.
(ii) Create State Representation: Identify object-centric (or neurosymbolic) state abstractions.
II. LLM-Driven Reward Generation:
(iii) Generating Useful Relational Functions: The LLM is tasked with generating relational functions to describe the relationships between objects in the given environment.
(iv) Reward Generation: Given the task context and the created relational functions, the LLM generates a readable neurosymbolic reward function.
(v) Reward Scaling: Adjust the created reward function in such a way that the rewards are on a scale from -1 to 1.
III. Policy Training:
(vi) Train DRL: Train agents using the generated reward function instead of the one, given by the environment.

Conclusion

We introduce **OCALM** - generating reward functions for games, using NLP task descriptions. We demonstrate that:
(i) LLMs can generate learnable reward functions in one shot,
(ii) the derived signals are effective in guiding the agent to learn the desired behavior,
(iii) incorporating object-centric reasoning significantly improves the quality and applicability of the generated reward functions, and reduce the inference costs.


Timo Kaufmann


Jannis Blüml


Quentin Delfosse

[1] LMU Munich
[2] Munich Center of Machine Learning
[3] ONE Munich
[4] AIML Lab TU Darmstadt
[5] hessian.AI
[6] ATHENE
[7] German Research Center for AI