AUTOMATED GENERATE REWARD FUNCTION BY LLMS IN REINFORCEMENT LEARNING

Phuc Pham Huynh

University of Information Technology - UIT (VNU-HCM)

What ?

We're revolutionizing Reinforcement Learning with our approach combines:

- Advanced prompt engineering with few-shot examples and chain-of-thought reasoning.
- Iterative refinement through evolutionary search.
- Comprehensive evaluation against human-designed benchmarks.

Why ?

The Problem: Creating effective reward functions is a major bottleneck in RL deployment:

- Requires deep expertise in both the domain and RL techniques
- Involves extensive trial-and-error, consuming valuable time and resources
- Often results in sub-optimal designs that limit agent performance.

Overview

Input → Method → Output

Input

- Natural Language Task Descriptions
- Technical Environment Information
- Example Samples (For few-shot methods)
- Training Feedback (For iterative methods)

Method

Figure 1. Illustration for a large language model.

Output

- Complete source code of reward functions that can be directly integrated into RL frameworks

Description

1. Experimental Setup

- Deploy RL environments (Gymnasium, MuJoCo, Isaac Gym)
- Select tasks of varying complexity from control to robot manipulation

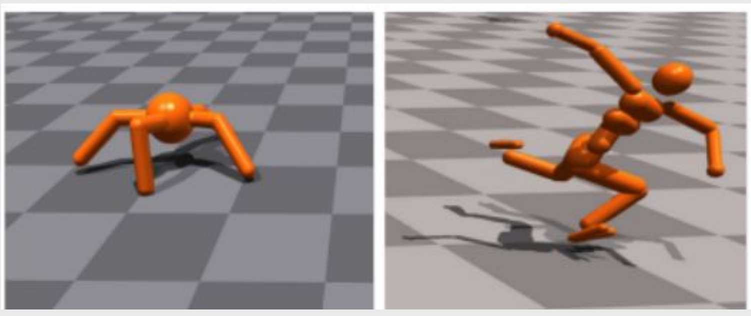

Figure 2. Illustration for a dataset.

2. LLM-Based Reward Generation

- Integrate coding-capable LLMs (GPT-o1, Claude, DeepSeek-R1)
- Design prompts using zero-shot, few-shot, and chain-of-thought approaches
- Implement evolutionary search to iteratively improve reward functions

3. Training and Evaluation

- Train agents using standard RL algorithms (PPO, SAC, TD3)
- Measure performance metrics including success rate and convergence speed

4. Comparative Analysis

- Benchmark against human-designed reward functions
- Analyze strengths and limitations of automatically generated rewards
- Identify failure cases and propose improvements

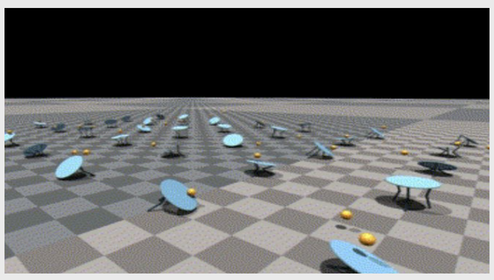

Figure 3. Illustration for result analysis.

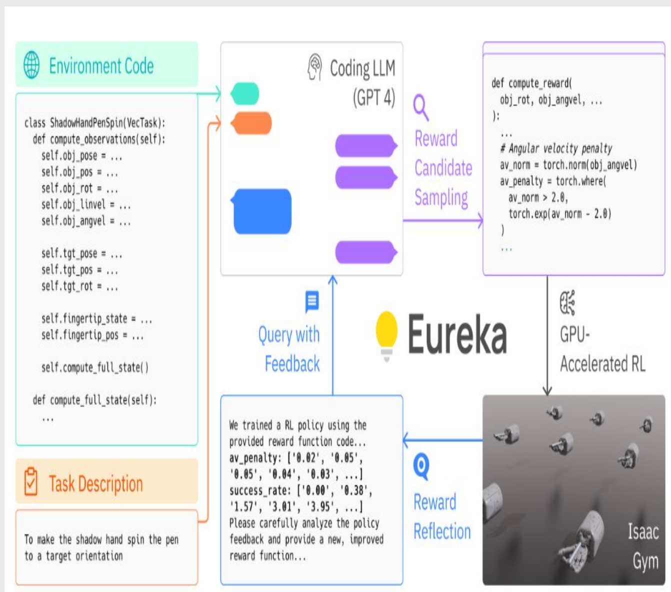

Figure 4. Illustration for the method

NII

Tên tác giả chính– Trường Đại học

TEL : *******   Email : *********