

MCTS-UCT with LLM - Assignment Report

Paper Referenced:

"Large Language Models as Commonsense Knowledge for Large-Scale Task Planning," NeurIPS 2023, Zirui Zhao et al.

Objective:

Implement Monte Carlo Tree Search with Upper Confidence Bound applied to Trees (MCTS-UCT) guided by a simulated Large Language Model (LLM) scoring heuristic in a simple sequential decision task.

Implementation:

- Created a toy environment with numeric states incremented by +1 or +2 actions.
- Defined goal state as 10 with reward 1.
- MCTS algorithm balances exploration-exploitation using UCT formula.
- An LLM proxy function heuristically scores states during rollout to bias action selection.

Results:

The algorithm successfully planned a sequence of actions reaching the goal state efficiently. This demonstrates the concept of combining MCTS with LLM-guided scoring in a runnable Python implementation.

Limitations:

- Proxy scoring is a heuristic, not a real LLM.
- Toy environment is simple; more complex tasks needed for full validation.
- Hallucination detection and knowledge grounding from the original paper are not included.

```
Starting state: 0
Step 1, choosing action: 2
New state: 2
Step 2, choosing action: 2
New state: 4
Step 3, choosing action: 2
New state: 6
Step 4, choosing action: 2
New state: 8
Step 5, choosing action: 2
New state: 10
Reached terminal state.
```