PlanU: Large Language Model Decision Making through Planning under Uncertainty

Ziwei Deng[ab*], **Mian Deng**[ab*], **Chenjing Liang**[ab], **Zeming Gao**[ab], **Chennan Ma**[ab], **Chenxing Lin**[ab], **Haipeng Zhang**[ab], **Songzhu Mei**[c], **Cheng Wang**[ab], **Siqi Shen**[ab†]
[a]Fujian Key Laboratory of Sensing and Computing for Smart Cities,
School of Informatics, Xiamen University (XMU), China
[b]Key Laboratory of Multimedia Trusted Perception and Efficient Computing, XMU, China
[c]School of Computer, National University of Defense Technology, China
{dengziwei,miandeng,liangcj,zeminggao,chennanma,lincx1123,zhanghaipeng}@stu.xmu.edu.cn,
{siqishen,cwang}@xmu.edu.cn, {sz.mei}@nudt.edu.cn

Abstract

Large Language Models (LLMs) are increasingly being explored across a range of decision-making tasks. However, LLMs sometimes struggle with decision-making tasks under uncertainty that are relatively easy for humans, such as planning actions in stochastic environments. The adoption of LLMs for decision-making is impeded by uncertainty challenges, such as LLM uncertainty and environmental uncertainty. LLM uncertainty arises from the stochastic sampling process inherent to LLMs. Most LLM-based Decision-Making (LDM) approaches address LLM uncertainty through multiple reasoning chains or search trees. However, these approaches overlook environmental uncertainty, which leads to poor performance in environments with stochastic state transitions. Some recent LDM approaches deal with uncertainty by forecasting the probability of unknown variables. However, they are not designed for multi-step decision-making tasks that require interaction with the environment. To address uncertainty in LLM decision-making, we introduce PlanU, an LLM-based planning method that captures uncertainty within Monte Carlo Tree Search (MCTS). PlanU models the return of each node in the MCTS as a quantile distribution, which uses a set of quantiles to represent the return distribution. To balance exploration and exploitation during tree search, PlanU introduces an Upper Confidence Bounds with Curiosity (UCC) score which estimates the uncertainty of MCTS nodes. Through extensive experiments, we demonstrate the effectiveness of PlanU in LLM-based decision-making tasks under uncertainty.

1 Introduction

Large language models (LLMs) have demonstrated remarkable capabilities in various domains, such as reasoning and coding [1, 2, 3]. The success of LLMs in various domains has motivated researchers to apply them to decision-making tasks [4, 5, 6], where an agent selects actions based on the current state in order to achieve specific goals.

Prior work has explored leveraging LLMs for decision-making in three ways: (1) LLMs are used as policies [4, 7], where the LLM is provided with necessary prompts to generate actions. (2) LLMs are used as world models [5], where the LLM is repurposed with appropriate prompts to generate the next state and reward based on a given action. (3) LLMs are used as both the policy and the world model [8, 6, 9], our work belongs to such category as it mimics human decision-making process,

[*]Equal contribution
[†]Corresponding author

39th Conference on Neural Information Processing Systems (NeurIPS 2025).