将 LLMs 与基于逻辑的框架结合以解释 MCTS

扩展摘要

Ziyan An
Vanderbilt University
Nashville, US

Xia Wang
Vanderbilt University
Nashville, US

Hendrik Baier
Eindhoven University of Technology
Eindhoven, Netherlands

Zirong Chen
Vanderbilt University
Nashville, US

Abhishek Dubey
Vanderbilt University
Nashville, US

Taylor T. Johnson
Vanderbilt University
Nashville, US

Jonathan Sprinkle
Vanderbilt University
Nashville, US

Ayan Mukhopadhyay
Vanderbilt University
Nashville, US

Meiyi Ma
Vanderbilt University
Nashville, US

摘要

针对在顺序规划中对人工智能（AI）缺乏信任的问题，我们设计了一个由计算树逻辑指导的基于大型语言模型（LLM）的自然语言解释框架，该框架旨在服务于蒙特卡罗树搜索（MCTS）算法。由于其搜索树的复杂性，MCTS 通常被认为是难以解读的，但我们的框架足够灵活，能够处理围绕 MCTS 和应用程序领域中的马尔可夫决策过程（MDP）为中心的一系列自由形式的事后查询和知识型查询。通过将用户查询转化为逻辑和变量声明，我们的框架确保从搜索树中获得的证据与底层环境动态以及实际随机控制过程中存在的任何约束保持事实一致。我们通过定量评估严格评估了该框架，并且在准确性及事实一致性方面表现出色。

KEYWORDS

顺序规划，可解释人工智能，大型语言模型，蒙特卡洛树搜索

ACM Reference Format:
Ziyan An, Xia Wang, Hendrik Baier, Zirong Chen, Abhishek Dubey, Taylor T. Johnson, Jonathan Sprinkle, Ayan Mukhopadhyay, and Meiyi Ma. 2025. 将 LLMs 与基于逻辑的框架结合以解释 MCTS: 扩展摘要. In *Proc. of the 24th International Conference on Autonomous Agents and Multiagent Systems (AAMAS 2025), 底特律，密歇根州，美国, May 19 – 23, 2025*, IFAAMAS, 4 pages.

1 介绍

人工智能（AI）算法通常作为黑箱系统运行，对其输出背后的推理提供很少或根本没有洞察。因此，领域专

Proc. of the 24th International Conference on Autonomous Agents and Multiagent Systems (AAMAS 2025), Y. Vorobeychik, S. Das, A. Nowe（编）, May 19 – 23, 2025, 底特律，密歇根州，美国.

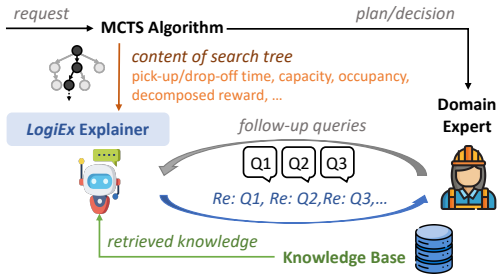

图 1: 我们通过结合领域知识、搜索过程和逻辑推理来解释顺序规划。

家由于对透明度、可理解性和问责制的担忧而犹豫在现实世界环境中部署这些算法，这使得他们无法清楚地了解这些 AI 模型所做的决策背后的影响或理由 [5, 12–16]。

这样一类广泛用于复杂顺序规划问题（如制造工程 [17] 和交通路线规划 [20] ）的人工智能方法是蒙特卡洛树搜索 (MCTS) [10] 。由于它们源自庞大的基于采样的搜索树，即使是专家也难以理解 MCTS 的结果和决策 [2, 4] 。因此，我们开发了一个逻辑增强型大规模语言模型（LLMs）框架，该框架将知识和符号推理与自然语言结合在一起，创建了一个强大且表达力强的 xAI 系统来解释像 MCTS 这样的规划算法（图 1 ）。

旨在解决一系列灵活的自由形式用户查询，我们的框架利用了先进的大语言模型（LLM），这使得基于自然语言 [6, 8, 19] 的 xAI 系统得以开发。更具体地说，它通过将用户通过聊天界面提交的自然语言询问转换为参数化变量和逻辑表达式，提供了处理查询的广泛灵活性。然后根据这些逻辑表达式的指定标准评估搜索树，并在最终解释中以自然语言形式呈现结果。该框架还支持无限数量的后续查询，促进了与用户的互动、来回沟通。

2 方法

背景. 作为我们框架的测试平台，我们使用了一个作为马尔可夫决策过程（MDP）定义的辅助运输规划场景。我们定义了 MDP 的状态空间、动作空间、约束和奖励。状态转换由一个模拟的需求模型驱动，该模型用于辅助运输行程请求。我们利用 MCTS 生成车辆分配决策，在每个“决策时点” [9] 开始。

查询类别和类型. 第一类查询，称为事后查询，在算法完成执行后寻求对返回计划的解释，并专注于解释特定的 MCTS 决策。第二类，称为基于背景知识的查询，则一般关注于 MCTS 的决策过程。用户提交查询后，Query-Classification LLM 组件会解析新查询并尝试将其意图归类为两类之一。用户的查询在内容或叙述方面没有限制。然而，为了战略性地处理这些查询，我们根据第一类用户的基本意图预定义了 26 种具体的查询类型。相比之下，对于可以用背景知识回答的第二类查询，并没有具体类型，因为一条知识可以解答多个查询。

逻辑生成器和解析器. 每种预定义的查询类型都与一个少样本提示相关联，其中包含输入查询和输出逻辑的示例对。当一个新的查询被分类为特定的查询类型后，相应的提示将用于引导逻辑生成 LLM 组件来制定该查询的逻辑陈述。我们根据回答问题所需证据的类型对所有用户问题进行分类：那些可以用基础级别的证据解决的问题，指的是直接从树节点提取的信息；那些依赖于衍生证据的问题，需要考虑不同深度或分支上的多个节点；以及那些需要逻辑比较证据的问题，涉及多级计算和使用计算树逻辑 (CTL) [7] 在两个分支之间进行比较。变量被组织成一个三级层次结构，每一级都建立在前一级定义的变量和逻辑之上。

逻辑评分器. 为了获得定量和定性证据，我们定义了评分器函数，这些函数将包括状态和动作的 MCTS 树作为输入，并根据特定标准的评估返回数值或布尔值 [3] 。对于基础变量，结果是通过遍历树来识别与该变量对应的节点而获得的。对于派生证据变量，我们进一步定义公式来计算搜索树中所有相关节点的整体平均定量结果。最后，我们利用 CTL 模型检测算法来获取逻辑比较证据，其中输入是 MCTS 树。

知识检索. 为了为第二类查询提供领域知识驱动的解释，我们准备了一个包含大约 3,000 个单词的轻量级知识库，分为 34 个部分。该知识库涵盖了关于辅

表 1: 定量评估结果。

方法	度量	@1↑	@3↑
Llama3.1	FactCC / BERT	25.77% / 06.15%	34.62% / 12.31%
Ours (Llama)	FactCC / BERT	67.88% / 86.54%	**83.27%** / **97.50%**
GPT-4o	FactCC / BERT	42.31% / 40.00%	51.15% / 55.77%
Ours (GPT)	FactCC / BERT	**72.12%** / **88.46%**	81.35% / 94.81%

助交通服务和 MCTS 算法的背景信息，以及 MDP 的详细组件，包括预定义约束、算法目标和奖励函数。我们利用 RAG 技术与 OpenAI 文本嵌入-3-小型模型获得前 k 个结果，仅在相关信息块的相关性得分超过预定义阈值时才将它们传递给 LLM。

生成解释. 一旦计算出的证据列表或检索到的专业知识获得后，框架将与一个问答式的大语言模型互动以生成最终响应。提供给大语言模型的关键信息包括原始用户查询、所使用的证据变量、前一步骤中得分函数的结果以及检索到的知识。

3 评估

我们定量评估该框架以回答研究问题(RQ)：我们的框架是否在生成事实准确且相关的解释方面优于现有的大语言模型？我们在评估中考虑了三种大语言模型：GPT-4 [1], GPT-4o [1] 和 Llama3.1 [18] 模型。我们手动准备了 620 个不同的查询作为输入，每个查询重复三次。对于每个查询，我们手动准备与其对应的类别 ID、正确的证据变量和逻辑以及参考叙述段落。我们使用两个指标比较生成的解释：BERTScore [21] 和 FactCC [11]。

事实一致性结果与讨论. 如表 1 所示，基础 LLM 取得的最佳 FactCC 分数为 51.15%，最高 BERTScore 为 55.77%。这两个结果都表明基础 LLM 在直接生成相关且事实准确的解释方面存在困难。然后我们将其与以 GPT-4 和 Llama3.1 作为骨干模型的我们的框架进行了比较，观察到了显著的改进。我们的框架在所有类别中始终优于基础 LLM。具体来说，使用 Llama3.1 时 FactCC 分数提高了 2.40×，而使用 GPT-4 模型则提高了 1.59×。BERTScore 的改善更加明显，Llama3.1 骨干模型总体增加了 7.92×，GPT-4 骨干模型则增加了 1.70×。

4 结论

我们提出了一种用于 MCTS 顺序规划的可解释性框架。在辅助运输规划场景中进行了测试，我们的框架可以通过三级层次证据提供事后解释(搜索后)和基于 RAG 的解释来解决各种用户查询。我们全面定量评估了该框架的表现。结果显示，与传统的 LLMs 相比，我们的框架总体上表现更优。

致谢

本材料基于由国家科学基金会（NSF）在奖助金编号 2028001、2220401、2151500、CNS-2238815 和 CNS-1952011，AFOSR 在 FA9550-23-1-0135，DARPA 在 FA8750-23-C-0518，NWO 在 NWA.1389.20.251 以及 Horizon Europe 在 101120406 资助下完成的工作。本文仅反映作者的观点，并不一定代表资助机构的意见。

REFERENCES

[1] Josh Achiam, Steven Adler, Sandhini Agarwal, Lama Ahmad, Ilge Akkaya, Florencia Leoni Aleman, Diogo Almeida, Janko Altenschmidt, Sam Altman, Shyamal Anadkat, et al. 2023. Gpt-4 technical report. *arXiv preprint arXiv:2303.08774* (2023).

[2] Ziyan An, Hendrik Baier, Abhishek Dubey, Ayan Mukhopadhyay, and Meiyi Ma. 2024. Enabling MCTS Explainability for Sequential Planning Through Computation Tree Logic. *arXiv preprint arXiv:2407.10820* (2024).

[3] Ziyan An, Taylor T Johnson, and Meiyi Ma. 2024. Formal Logic Enabled Personalized Federated Learning through Property Inference. In *Proceedings of the AAAI Conference on Artificial Intelligence*, Vol. 38. 10882–10890.

[4] Hendrik Baier and Michael Kaisers. 2020. Explainable search. In *2020 IJCAI-PRICAI Workshop on Explainable Artificial Intelligence*. 178.

[5] Hendrik Baier and Michael Kaisers. 2021. Towards explainable MCTS. In *2021 AAAI Workshop on Explainable Agency in AI*. 178.

[6] Zirong Chen, Elizabeth Chason, Noah Mladenovski, Erin Wilson, Kristin Mullen, Stephen Martini, and Meiyi Ma. 2024. Sim911: Towards Effective and Equitable 9-1-1 Dispatcher Training with an LLM-Enabled Simulation. *arXiv preprint arXiv:2412.16844* (2024).

[7] Edmund M Clarke and E Allen Emerson. 1981. Design and synthesis of synchronization skeletons using branching time temporal logic. In *Workshop on logic of programs*. Springer, 52–71.

[8] Neel Guha, Julian Nyarko, Daniel Ho, Christopher Ré, Adam Chilton, Alex Chohlas-Wood, Austin Peters, Brandon Waldon, Daniel Rockmore, Diego Zambrano, et al. 2024. Legalbench: A collaboratively built benchmark for measuring legal reasoning in large language models. *Advances in Neural Information Processing Systems* 36 (2024).

[9] Waldy Joe and Hoong Chuin Lau. 2020. Deep reinforcement learning approach to solve dynamic vehicle routing problem with stochastic customers.

In *Proceedings of the international conference on automated planning and scheduling*, Vol. 30. 394–402.

[10] Levente Kocsis and Csaba Szepesvári. 2006. Bandit based monte-carlo planning. In *European conference on machine learning*. Springer, 282–293.

[11] Wojciech Kryściński, Bryan McCann, Caiming Xiong, and Richard Socher. 2019. Evaluating the factual consistency of abstractive text summarization. *arXiv preprint arXiv:1910.12840* (2019).

[12] Scott M Lundberg and Su-In Lee. 2017. A unified approach to interpreting model predictions. *Advances in neural information processing systems* 30 (2017).

[13] Meiyi Ma, Ji Gao, Lu Feng, and John Stankovic. 2020. STLnet: Signal temporal logic enforced multivariate recurrent neural networks. *Advances in Neural Information Processing Systems* 33 (2020), 14604–14614.

[14] Meiyi Ma, John A Stankovic, and Lu Feng. 2021. Toward formal methods for smart cities. *Computer* 54, 9 (2021), 39–48.

[15] Joao Marques-Silva and Alexey Ignatiev. 2022. Delivering trustworthy AI through formal XAI. In *Proceedings of the AAAI Conference on Artificial Intelligence*, Vol. 36. 12342–12350.

[16] Marco Tulio Ribeiro, Sameer Singh, and Carlos Guestrin. 2016. " Why should i trust you?" Explaining the predictions of any classifier. In *Proceedings of the 22nd ACM SIGKDD international conference on knowledge discovery and data mining*. 1135–1144.

[17] M Saqlain, S Ali, and JY Lee. 2023. A Monte-Carlo tree search algorithm for the flexible job-shop scheduling in manufacturing systems. *Flexible Services and Manufacturing Journal* 35, 2 (2023), 548–571.

[18] Hugo Touvron, Thibaut Lavril, Gautier Izacard, Xavier Martinet, Marie-Anne Lachaux, Timothée Lacroix, Baptiste Rozière, Naman Goyal, Eric Hambro, Faisal Azhar, et al. 2023. Llama: Open and efficient foundation language models. *arXiv preprint arXiv:2302.13971* (2023).

[19] Sean Welleck, Jiacheng Liu, Ronan Le Bras, Hannaneh Hajishirzi, Yejin Choi, and Kyunghyun Cho. 2021. Naturalproofs: Mathematical theorem proving in natural language. *arXiv preprint arXiv:2104.01112* (2021).

[20] Di Weng, Ran Chen, Jianhui Zhang, Jie Bao, Yu Zheng, and Yingcai Wu. 2020. Pareto-optimal transit route planning with multi-objective monte-carlo tree search. *IEEE Transactions on Intelligent Transportation Systems* 22, 2 (2020), 1185–1195.

[21] Tianyi Zhang, Varsha Kishore, Felix Wu, Kilian Q Weinberger, and Yoav Artzi. 2019. Bertscore: Evaluating text generation with bert. *arXiv preprint arXiv:1904.09675* (2019).